

DOI: 10.15593/perm.kipf/2020.4.01

УДК 004.896

## ПРИНЦИПЫ КРЕАТИВНОЙ РОБОТОТЕХНИКИ

А.Ю. Алексеев<sup>1</sup>, Т. Пожарев<sup>2</sup>

<sup>1</sup>Московский государственный университет имени М.В. Ломоносова, г. Москва, Россия

<sup>2</sup>Московский государственный университет имени М.В. Ломоносова, г. Новый Сад, Сербия

### О СТАТЬЕ

Получена: 01 октября 2020 г.  
Принята: 17 ноября 2020 г.  
Опубликована: 19 января 2021 г.

#### Ключевые слова:

аргумент Лавлейс, контраргумент Лавлейс, тест Лавлейс, тест Лавлейс 2.0, тест Лавлейс 3.0, комплексный тест Тьюринга, креативный робот, машина Тьюринга, машина Корсакова, машина Корсакова-Тьюринга.

### АННОТАЦИЯ

Креативный робот – это робототехническая реализация комплексного теста Тьюринга на предмет компьютерной имитации творческой деятельности. Исследованы концептуальные и технические принципы креативной робототехники. Концептуальные принципы включают аргумент Лавлейс (1843), контраргумент Тьюринга (1950), тест Лавлейс (2000), тест Лавлейс 2.0 (2014) и тест Лавлейс 3.0 (2015).

Согласно аргументу А. Лавлейс, компьютер может создать высокохудожественное произведение, однако самостоятельно творить не может: машина лишь пассивно исполняет инструкции разработчика-программиста. А. Тьюринг, напротив, полагает, что машина может творить и, следовательно, мыслить. Контраргумент Тьюринга делится на две части: символическое и коннекционистское возражение. Первое неубедительно: разработчик пытается обмануть наблюдателя (тьюрингового судью), подчеркивая неотличимость творчества от его программной имитации. Оригинальность продукта определяется неожиданным ходом выполнения программы, например машинным сбоем. Коннекционистский (нейросетевой) ответ правдоподобнее: в общих чертах поведение робота предсказуемо, однако частные продукты машинного обучения могут казаться оригинальными.

Тест Лавлейс отрицает символический ответ Тьюринга путем постулирования аксиом, запрещающих обман наблюдателя и сбой программы. Уникальность артефакта ограничена тем, что продукты известны разработчику: множество продуктов может быть потенциально бесконечным, тем не менее оно декларативно или процедурально задано, следовательно, конечно и предсказуемо.

Тест Лавлейс 2.0 отрицает коннекционистский ответ Тьюринга, так как применим к нейронным сетям глубокого обучения. Судье сложно оценить креативность артефакта вне сферы его творческих компетенций. Аксиоматика ужесточается специализацией предметных областей и введением роли оценщика креативности продуктов.

Предложен тест Лавлейс 3.0. Он дополняет предыдущую версию аксиомой обнаружения оценщиком осмысленного производства артефакта. Машина не способна продуцировать смысл, поэтому по-настоящему творить не может. Новый тест пройти нельзя, однако возможны относительные успехи креативной робототехники. Для этого нужны технические принципы интеграции символического и коннекционистского подходов. Предлагается машина Корсакова-Тьюринга как формальное определение алгоритма функционирования креативного робота: стопка перфокарт коннекционистской машины Корсакова (1832) характеризует связи субсимвольных составляющих ленты и таблицы символической машины Тьюринга (1936).

© ПНИПУ

© **Алексеев Андрей Юрьевич** – доктор философских наук, ведущий научный сотрудник кафедры философии и методологии науки

© **Пожарев Тодор** – аспирант кафедры философии и методологии науки

© **Andrey Yu. Alekseev** – Doctor of Science (Philosophy), Leading Researcher, Department of Philosophy and Methodology of Science

© **Todor Pozarev** – Postgraduate, Department of Philosophy and Methodology of Science



Эта статья доступна в соответствии с условиями лицензии Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

## THE PRINCIPLES OF CREATIVE ROBOTICS

Andrey Yu. Alekseev<sup>1</sup>, Todor Pozarev<sup>2</sup>

<sup>1</sup>Lomonosov Moscow State University, Moscow, Russia

<sup>2</sup>Lomonosov Moscow State University, Novi Sad, Serbia

### ARTICLE INFO

Received: 01 April 2020  
Accepted: 03 September 2020  
Published: 19 January 2021

#### Keywords:

Lovelace argument, Lovelace test, Lovelace Test 2.0, Lovelace Test 3.0, Comprehensive Turing Test, creative robot, general artificial intelligence, computer creativity, Turing machine, Korsakov machine, Korsakov-Turing machine.

### ABSTRACT

A creative robot is a robotic implementation of the Comprehensive Turing test for computer simulation of creative activity. The article explores the conceptual and technical principles of creative robotics. Conceptual principles include Lovelace's argument (1843), Turing's counter-argument (1950), Lovelace's test (2000), Lovelace's test 2.0 (2014) and Lovelace's test 3.0 (2015).

According to A. Lovelace's argument a computer can create a highly artistic work, but it cannot create on its own: the machine only passively executes the instructions of the developer-programmer. A. Turing, on the contrary, believes that the machine can create and, therefore, think. Turing's counter-argument falls into two parts: the symbolic objection and the connectionist objection. The first is unconvincing: the developer is trying to deceive the observer (the Turing Interrogator) into the indistinguishability of creativity and its software imitation. The originality of a product is determined by an unexpected program flow, such as a machine failure. The connectionist (neural network) answer is more plausible: in general terms, the behavior of a robot is predictable, but particular machine learning products can seem original.

Lovelace's test denies Turing's symbolic response by postulating axioms that deceive the observer and prevent program failure. The uniqueness of the artifact is limited by the fact that the products are known to the developer: the set of products can be potentially infinite, nevertheless, it is declaratively or procedurally specified, and therefore, it is actually finite and predictable.

The Lovelace 2.0 test rejects Turing's connectionist answer as it applies to deep learning neural networks. It is difficult for the Interrogator to ascribe the creativity of an artifact outside the scope of his creative competence. Axiomatics is being tightened by the specialization of subject areas and the introduction of the role of evaluator of product creativity.

In the article it is proposed the Lovelace 3.0 test. It complements the previous version with the axiom of the evaluator detecting meaningful artifact production. The machine is incapable of producing meaning, so it cannot truly create. The new test cannot be passed, but the relative successes of creative robotics are possible. This requires technical principles for integrating symbolic and connectionist approaches. The Korsakov-Turing machine is proposed as a formal definition of the algorithm for the functioning of a creative robot: a stack of punched cards of the Korsakov connectionist machine (1832) characterizes the connections between the sub-symbolic components of the tape and the table of the Turing symbolic machine (1936).

© PNRPU

### Введение: дефиниция креативного робота

Может ли робот творить? Слово сочетание «креативный робот» интуитивно понятно. Функционирование такого робота может быть сравнимо с творческой деятельностью человека, неотлично от нее и порой ее превосходить. Однако что такое творчество и как можно (и можно ли) его запрограммировать? Ведь компьютерная имитация творчества сомнительна: симуляция креативности акреативна. В воспроизведении нет ни новизны, ни оригинальности. Если имитация творчества невозможна в силу противоречия, то имеется ли возможность программировать причинные основания креативности? Каковы критерии креативности функций компьютерной системы? Кто устанавливает эти критерии и оценивает творческие способности робота? Какова техническая организация этого «робота», которому можно обоснованно приписывать креативные компетенции? Насколько важна роль креативной робототехники как инфраструктуры построения и развития таких роботов? Какова роль среды в производстве креативной машины?

Вопрос «Может ли робот творить?» методологически неотличим от основного вопроса искусственного интеллекта «Может ли машина мыслить?». Ровно семьдесят лет назад путь его решения предложил Алан Тьюринг в статье «Компьютерная технология и интеллект» [1] (статья опубликована в октябре 1950 г.). А. Тьюринг не утруждался дефинициями «интеллекта» и «машины». Он предложил игру в имитацию пола игрока и полагал, что в этой игре,

в условиях скрытия физических особенностей игроков средствами телекоммуникации, проявляется диалоговый интеллект игрока. Игроком может быть человек или машина. И отвечать на вопросы судьи (наблюдателя, исследователя), присуждающего пол игроку, способен не только человек, но и компьютерная программа.

Интеллектуальная машина, согласно такому подходу, – это компьютерная реализация игры в имитацию. Мы предложим аналогичное определение: креативный робот – это робототехническая реализация комплексного теста Тьюринга на творчество. Если робот способен проходить такой тест, то тогда он может творить.

Журнал Mind выделил А. Тьюрингу два авторских листа. Поэтому сегодняшний юбиляр в своей эпохальной статье [1] представил достаточно много: тестовые вопросы игры в имитацию, полемику по поводу мыслящих машин, футурологические прогнозы, описания технических, программных, информационных средств мыслящих машин. Журнал «Технологос» выделяет много меньше места. По этой причине мы вынуждены раскрыть лишь базовые принципы креативной робототехники. К ним относятся *принципы* концептуальной (логической) и компьютерной (технической) организации роботов. То есть в работе будет изучаться *технологос креативного робота*.

### **Концептуальные принципы креативной робототехники: комплексный тест Тьюринга на творчество**

Концептуальная организация креативной робототехники рассматривается в контексте комплексного теста Тьюринга (см. подробнее в [2–4])<sup>1</sup>. Дело в том, что сегодня когнитивная феноменология компьютерного мира не может быть раскрыта путем применения оригинального теста Тьюринга, точнее игры в имитацию, в которую играют и люди, и компьютеры. За семьдесят лет осмысления философской, научной и инженерной роли теста Тьюринга произошло много событий, созданы сотни версий теста, версий этих и последующих версий. Поэтому при изучении когнитивных феноменов компьютерного мира целесообразно пользоваться комплексным тестом Тьюринга [2].

Комплексный тест обеспечивает концептуальный анализ возможности компьютерной реализации (репликации, репрезентации и репродукции) таких феноменов, как интеллект, сознание, понимание, творчество, жизнь, личность, общество, любовь и многое другое. Тест выполняет *дефинитную функцию*, сочетая в едином формате смыслы слов, которые характеризуют естественные когнитивные феномены и искусственные, компьютерно имитируемые феномены. *Критическая функция* обеспечивает рациональную форму дискуссий относительно ответа на вопросы: может ли компьютер мыслить, творить, понимать, сознавать, чувствовать, любить и пр.? *Конструирующая функция* способствует проектированию когнитивных компьютеров, так как раскрывает фундаментальные принципы их работы. *Конститутивная функция* подчеркивает праксеологические особенности субъекта компьютерных технологий, его свободу и зависимость от мира электронной культуры. *Коммуникативная функция* позволяет организовать междисциплинарную методологию изучения и развития искусственного интеллекта: тьюринговые мысленные эксперименты просты и потому понятны многим. Возможность проходить тесты определяется *интеррогативной функцией* комплексного теста и зависит от умения судьи ставить и анализировать вопросы выявления когнитивных компетенций изучаемых систем.

Идея комплексного использования тестов для изучения когнитивных феноменов компьютерного мира проста. Предполагается тесная взаимосвязанность и взаимообусловленность одного феномена и других феноменов. Чтобы изучить *искусственную креативность*, необхо-

<sup>1</sup> Далее в статье используются в том числе материалы предыдущей работы одного из авторов. См. [5].

димо запустить оригинальный тест (тест Тьюринга) и сотни других тестов. Необходим и особенно важен, конечно, специальный тест на творчество, креативный тест (тест Лавлейс). Он оценивает творческие компетенции компьютерной системы и отвечает на основной вопрос данной статьи. Важен субкогнитивный тест (тест Френча). Он раскрывает особенности компьютерной имитации в среде многоуровневых ассоциаций. Тест на понимание (тест Серля «Китайская комната») оценивает вклад, который вносит компьютерная обработка в осмысление когнитивного феномена. Тест на здравый смысл (тест Маккарти-Деннета) изучает возможности компьютерного моделирования знаний высококвалифицированного эксперта наряду с представлением многообразия иррелевантных сведений. Инвертированный тест (тест Ватта) оценивает тестирование с позиции судьи-машины. НБИКС-тест (тест Харнада) изучает когнитивную робототехнику в конвергентной системе, построенной на базе интегративного комплекса нано-, био-, инфо-, когни- и социотехнологий. Тест зомби предлагает исследовать возможности компьютерной имитации систем, выказывающих полноту когнитивных компетенций, но фактически не обладающих этими способностями. В [2] предложена двадцать одна версия теста Тьюринга. Но их гораздо больше. Все они влияют на оценку креативности компьютерной системы. Главная роль в этой оценке отводится специальным тестам компьютерного творчества: аргумент Лавлейс (предложен в 1843 г.), контраргумент А. Тьюринга (1950), тест Лавлейс (2000), тест Лавлейс 2.0 (2014) и тест Лавлейс 3.0 (2015). В целом в этих тестах отражается почти двухвековая история изучения искусственной креативности.

### **Аргумент Ады Лавлейс: компьютер творить не может**

Аргумент Ады Лавлейс («компьютер творить не может») был сформулирован в примечаниях к переводу статьи итальянского военного инженера Л. Ф. Менабреа (впоследствии ставшего премьер-министром Италии) «Очерк аналитической машины, изобретенной Чарльзом Бэббеджем». Статья на французском языке появилась в октябре 1842 года, а перевод, выполненный Адой Лавлейс, – в августе 1843 года [6; 7, с. 63–54]. Благодаря примечанию к этой статье Ада Лавлейс по праву считается первой в мире программисткой. Впервые в мировой истории в 1843 году научному сообществу был представлен «листинг» программы. Программа А. Лавлейс использовала цикл для вычисления чисел Бернулли и применяла для этого не «двадцать тысяч перфокарт», на которых последовательно фиксировались бы результаты рекуррентных вычислений, а всего три перфокарты. Эти перфокарты отвечали за хранение в «оперативной памяти» машины трех параметров итерации: начала цикла, шага итерации и условия завершения. Знаменитый «цикл» А. Лавлейс так и не был реализован: машина Бэббиджа заработала лишь в начале XX века в роли музейного экспоната. Тем не менее проект оказался востребованным в современной компьютерной технологии: по сути, все компьютеры массового применения построены на принципах работы этой машины.

А. Лавлейс впервые предложила использовать машину для работы с нечисловыми объектами: «...легко вообразить, что операционный механизм машины может работать с иными объектами, кроме чисел. Главное, чтобы абстрактные операции, которые применяются к абстрактным объектам той или иной науки, имели отображение в системе операций работы машины...» [6]. То есть А. Лавлейс является автором алгебраической парадигмы построения интеллектуальных компьютерных систем, в соответствии с которой семантика языка программирования задается средствами универсальных алгебр. Базовые множества, операции и отношения алгебр получают содержательную интерпретацию в приложениях самого различного назначения, например в робототехнике с недоопределенными закономерностями физических манипуляций или, скажем, в поэтике. Подчеркнем, что алгебраическая парадигма наряду с логической образует фундамент со-

временной инженерии знаний интеллектуальных компьютерных систем. Поэтому очевидна значимость идей Ады Лавлейс в развитии современного искусственного интеллекта.

Далее следует аргумент «компьютерного творчества»: «Легко предположить, что инструкции и механизмы машины приспособляются к воспроизведению звуков в соответствии с фундаментальными закономерностями, которые изучаются в науке о гармонии и музыкальной композиции. Благодаря этому машина способна создавать сложные и теоретически обоснованные музыкальные произведения на любой вкус» [6]. Вроде бы все замечательно: вычислительная машина способна сочинять высокохудожественные произведения, стоит только придумать соответствующую программу. Однако это не так.

В самом последнем примечании к работе о машине Бэббиджа (примечание «G» [6]) Ада Лавлейс отказывается от излишнего оптимизма: «Желательно оберегать себя от преувеличения возможностей аналитической машины. Ведь часто при возникновении нового, во-первых, переоценивается то, что представляется интересным и замечательным, и, во-вторых, своеобразной естественной реакцией является недооценка истинного положения дел, когда новые понятия заслоняют те понятия, которые ранее уже выдерживали критику» [6], т.е. вначале оцениваются технические новации (или, как сейчас сказали бы, определяются «technique assessment»).

Далее, основное положение негативной части аргумента Ады Лавлейс: «Аналитическая машина не претендует на создание чего-то нового. Она только делает то, что мы можем ей приказывать выполнить. Машина способна анализировать, однако от нее нельзя ожидать формирования нового аналитического отношения или установления новой аналитической истины. Она помогает нам наглядно представить то, с чем мы уже знакомы» [6].

А. Лавлейс наносит сильный удар по представителям так называемых «творческих» профессий: музыкантов, литераторов, художников и др. Ведь за изящным антуражем может стоять творческое бессилие автора «шедевра». Возможно, Ада мстит своему отцу, лорду Байрону, который один раз в жизни видел свою единственную законорожденную дочь. Хотя великого поэта вряд ли можно обвинить в отсутствии креативности. Кстати, в работах Р.С. Гутера и Ю.Л. Полунова раскрывается иная версия: Ада не любила свою мать, а отца боготворила [7].

Интересен факт выбора теоремы Бернулли как предмета программирования. Это одна из главных теорем комбинаторики. Первые компьютерные приложения комбинаторики связываются с машиной Р. Луллия и его проектом «Великого искусства» (Ars Magna). Машина Луллия обеспечивала юридическую постгlossаторскую практику вывода полисиллогизмов, формализующих правила естественного и позитивного права. Если Ада Лавлейс продолжает историю построения интеллектуальных машин, начатую в 1308 году Раймондом Луллием, то программа искусственной креативности насчитывает более семи веков. Исследователи креативности пытаются решить вопрос, как комбинаторика влияет на креативность.

Документально подтвержденным в [1] является скептический вывод А. Лавлейс о том, что, хотя компьютер может многое, к примеру создавать высокохудожественные произведения, на самом-то деле он творить не может и, следовательно, не способен мыслить. Это представлялось А. Тьюрингу наиболее веским контрударом по его позиции искусственного интеллекта.

### **Контраргумент А. Тьюринга: компьютер может творить**

Опровержению аргумента Ады Лавлейс (машина творить не может) посвящена наиболее важная часть статьи – ее завершающий параграф [1]. Разделим контраргумент А. Тьюринга на две части, которые назовем «символьным ответом» и «коннекционистским ответом», в соответствии с названиями двух базовых парадигм современной методологии искусственного ин-

теллекта. В рамках символического опровержения утверждается: «Под солнцем ничего нового не происходит. Все, что мы считаем оригинальным, на самом деле представляется лишь результатом обучения или следствиями общих принципов» [1]. Человеческая способность к творчеству непонятна. Невозможно найти подлинного автора идеи, так как помимо конечного автора имеется бесконечный ряд его учителей и соучителей, которых обучили другие авторы и соавторы. Совершенно не очевидны критерии оригинальности. К тому же результаты работы компьютера способны поразить любое воображение.

Это неубедительно. Логико-атомистический каркас символической парадигмы не обеспечивает свободы развития сущностей и связей моделируемой предметной области, ведь любая система формализованных знаний генерирует результаты в контексте известного формализма (аргумент Геделя-Лукаса-Пенроуза). Для продуцирования оригинального продукта диспозиционные знания программиста обязаны превышать актуальные знания судьи. Символический компьютер предлагает два варианта достижения этого превосходства: декларативный и процедурный. Для декларативных «знаний» такой вариант ограничен простыми системами с конечным множеством актуально или потенциально обозримых «знаний». Разработчик не способен угадать все возможные ситуации сложной предметной области, так как она «испещрена» уникальными вариациями, не подпадающими под общие дефиниции. Для процедурных «знаний», генерирующих многообразие предметной области из аксиом по правилам вывода, ограниченность универсального подхода менее очевидна. Однако и здесь границы задаются большими системами, бесконечное множество «знаний» которых потенциально перечислимо. То есть ни для малых, декларативно представимых систем, ни для больших, процедурально формируемых систем, не нужны методы и средства искусственного интеллекта. Для них достаточно традиционных компьютерных средств. Символистский ответ А.Тьюринга не убедителен перед натиском аргумента Ады Лавлейс: словами креативный процесс невозможно описать и тем более запрограммировать.

Иное дело – коннекционистский ответ. Он раскрывается в заключение статьи [1] и поэтому является, по сути, итогом тьюринговых исследований в области искусственного интеллекта. Предлагается новый проект машины. Это не УЦВМ (собственно, ее прообразом и являлась аналитическая машина Ч. Бэббиджа). Это – нейрокомпьютер в виде обучаемой нейронной сети. Имитация коммуникации заменяется моделированием обучения.

Интерпретируя не совсем ясные идеи обучения нейронной сети на основе [1, 8], предложим следующий проект тьюрингового нейрокомпьютера: 1) построение исходной системы знаний и интеллектуальных способностей, которыми обладает взрослый человек; 2) изучение способа роста знаний в ходе обучения; 3) изучение опыта, который не может быть результатом обучения; 4) очистка «знаний» от их содержания для формирования структурно-процедурного скелета интеллектуальных способностей; 5) представление «чистых» знаний как модели интеллекта ребенка («машины-ребенка»); 6) обучение машины-ребенка и формирование так называемой «машины-взрослого»; 7) сравнение «знаний» и «способностей» машины-взрослого со знаниями и способностями взрослого человека; 8) корректировка п. 4; 9) переход к п. 2.

Согласно коннекционистскому опровержению аргумента Ады Лавлейс, учитель, как правило, не осведомлен о внутренних процессах машины-взрослого. Однако он в некоторой степени и общих чертах может направлять и предсказывать поведение «ученика» (сегодня такой процесс многие называют предиктивным кодированием). Напомним, что основной аргумент Лавлейс, что машина может делать только то, что программист может ей приказывать, согласуется с тем, что некоторую часть работы машины он может прописать, а другую может не учитывать.

В случае с обучаемой машиной программист не знает наверняка о ее поведении, но лишь прогнозирует это поведение и, возможно, проверяет свои гипотезы для корректировки после-

дующего пути обучения. То есть машина-взрослый – это уже иная машина, не УЦВМ. Принципы УЦВМ повторяют проект аналитической машины Бэббеджа. То есть аргумент Лавлейс надо бы снять более совершенным нейросетевым проектом компьютера. И именно к обучаемым компьютерам, а не к УЦВМ относится заключительный и часто цитируемый вывод А. Тьюринга относительно роли и места интеллектуальных машин в обществе как соперников и помощников человека во всех сферах деятельности [1, р. 460].

Коннекционистское опровержение аргумента Лавлейс намного убедительнее символистского, если подкреплять его правдоподобным проектом нейросетевой машины. К этому проекту (машине Корсакова-Тьюринга) мы вернемся в завершение работы при описании технических принципов креативного робота.

### Тест Лавлейс

Одной из специальных работ, которую целесообразно отнести к креативной робототехнике, является статья С. Брингсйорда, П. Беллоу, Д. Феруччи «Творчество, тест Тьюринга и (улучшенный) тест Лавлейс» [9]. В работе много примеров из сферы изобразительного искусства, литературы, поэзии. Аргумент Ады Лавлейс – наиболее сильное из всех возражений тесту Тьюринга. Авторы используют перефразированную формулировку аргумента: компьютер не может самостоятельно творить, так как творчество требует как минимум изобретения чего-либо нового. Но компьютеры не изобретают ничего нового; они всего лишь делают то, что программист «заставляет» их делать посредством программ.

Как имитировать креативность некреативным? Ведь использование УЦВМ, «самое высшее достижение которой – решение арифметических задач» [9], это примерно то же самое, когда глупый пытается имитировать умного. Обратное, конечно, возможно [10].

Человек-разработчик  $P$  создает робота  $K$  (в формате искусственного агента), который продуцирует артефакт  $A$  как уникальную, неповторимую, полезную вещь.  $P$  не способен определить, каким образом  $K$  продуцирует  $A$ , несмотря на то, что  $P$  полностью осведомлен о внутреннем устройстве, функциях и поведении  $K$ .

Робот  $K$  проходит тест Лавлейс тогда и только тогда, когда:

- 1) робот  $K$  продуцирует артефакт  $A$ ;
- 2) артефакт  $A$  не является результатом обмана, случайного стечения обстоятельств или сбоя машины, т.е.  $K$  всегда может  $A$  воспроизвести (репродуцировать, повторить);
- 3) судья  $I$  не способен объяснить, как  $K$  создал  $A$ , хотя имеет полное представление о  $K$ : структуре баз данных, алгоритмах функционирования, техническом обеспечении и пр.

Предложенная аксиоматика приводит к следующим соображениям:

1) Креативный тест невозможно пройти путем обмана. Подтасовка ответов характерна для прохождения оригинального теста Тьюринга, здесь же по условию 2 это невозможно.

2) Тест Лавлейс всем современным компьютерным системам присваивает статус некреативных. Все эти модели запрограммированы разработчиками, которые знают особенности их функционирования. Ничего «от себя» такие роботы добавить не могут.

3) Необходимым условием прохождения теста Лавлейс является способность робота к программированию собственного программного кода. Но это невозможно из-за аргумента Геделя-Лукаса-Пенроуза: механическая система, которой является робот, не способна к самоформализации. Робот может порождать программы методами «автогенеративного программирования». Но он не способен к автогенерации принципов автогенерации, т.е. не способен самостоятельно порождать автогенеративные аксиомы и правила вывода. Креативный робот должен быть самопрограммируемым на уровне принципов программирования.

4) Но и самопрограммируемости недостаточно для прохождения теста Лавлейс. Дело в том, что все изменяемые правила носят «переходящий» характер. Результат обучения – это, к примеру, новая структура базы данных или измененный программный код. Но такие модификации не оригинальны, они происходят в контексте известного, по крайней мере со стороны учителя.

5) Моделирование «свободной воли», имитация мотивационно-волевых механизмов человека, стремящегося к созданию чего-то нового не результативно, так как результат представляется в некотором пространстве решений.

Таким образом, тест Лавлейс показывает невозможность компьютерной реализации феномена творчества. Однако возможно компьютерное продуцирование квазикреативных («как бы» творческих) продуктов. И от компетентности тьюрингового судьи зависит приписывание роботу креативности.

### Тест Лавлейс 2.0

Не так давно, в 2014 году, тест Лавлейс был дополнен небольшой статьей Марка Ридла «Тест Лавлейс 2.0 на искусственные креативность и интеллект» [11]. Была предложена новая версия исходных предпосылок:

- 1)  $K$  продуцирует артефакт  $A$  типа  $T$ ;
- 2)  $A$  должен соответствовать набору ограничений  $C$ , где  $c_i \in C$  любой критерий, выражаемый на естественном языке;
- 3) человек-оценщик  $H$ , выбравший  $T$  и  $C$ , удовлетворен тем, что  $A$  является допустимым экземпляром  $T$  и соответствует  $C$ ;
- 4) человек-судья  $I$  определяет комбинацию  $T$  и  $C$  как нереальную для обычного человека.

В предложенной аксиоматике количество людей увеличилось. Четко распределены их роли. Помимо человека, выполняющего роль тьюрингового судьи, появляется человек ( $H$ ), компетентно оценивающий гениальность, талантливость, креативность артефакта  $A$  или, напротив, бесталанность, обыденность этого произведения. Происходит четкая спецификация критериев оцениваемого продукта (п. 2). Человек-оценщик и человек-судья должны быть экспертами в достаточно ограниченной области, т.е. нужен конкретный специалист, а не «человек вообще», оценивающий, может или не может творить машина.

Тест «Лавлейс 2.0» возник на пике интереса к моделям глубокого машинного обучения. Эти модели помимо многофакторного взвешивания нейронных сетей позволяют комбинировать фон и рисунок, накладывая, например, на выявленный стиль одного художника «рисунок» другого художника. Например, полотно «Портрет Эдмонда Белами», нарисованное компьютером, было создано путем смешения стилей многих художников. В октябре 2018 года этот рисунок был продан с аукциона Christie's (Нью-Йорк, США) почти за полмиллиона долларов.

«Творчество не уникально для человеческого интеллекта. Но это одна из его отличительных черт. Многие формы творчества требуют интеллекта», – полагает М. Ридл [11]. На наш взгляд, это суждение дополняет тьюринговый подход, в котором интеллектуальность определяется креативностью среди всего прочего. Креативность и интеллектуальность – это взаимопологающие компетенции робота. Интеллектуальная робототехника основана на технологиях креативной робототехники. И креативная робототехника должна быть интеллектуальной.

### Тест Лавлейс 3.0

Продолжая нумерацию версий теста, предложим тест Лавлейс 3.0. По сути, идея нового теста на компьютерное творчество прозвучала в нашей ранней статье [12]. Оценивать компьютерную креативность можно с двух позиций: физикалистской и менталистской. Физикалистский



(натуралистический) подход обеспечивает имитацию творческой деятельности за счет сложных и необычных комбинаций известных объектов. Объекты могут быть представлены явно. Но могут быть завуалированными, неявными, скрытыми, то есть выглядеть своеобразными ассамбляжами субобъектов. Нейронные сети глубокого обучения – а именно эти технологии вызвали появление теста Лавлейс 2.0 – интересны тем, что позволяют манипулировать с субобъектами, например со «стилями» художников, вызывая у людей эстетические переживания и выманивая у них «хайповые» деньги, как было в приведенном выше примере. Второй, когнитивный (менталистский), подход ориентирован на работу с феноменами субъективной реальности, с идеальными объектами, со смыслами. Это абстракции, интенции, концепции. Эти когнитивные феномены семиотически закрепляются, выражаются и используются в объективном мире. Они значат, но их в нем нет.

Аксиоматика теста остается прежней. Немного уточняется п. 3:

3\*) оценщик  $H$  выявляет факт осмысленного продуцирования артефакта  $A$ .

Например, какая операция использовалась при создании  $A$ : абстрагирование или обобщение? Если оценщик  $H$  обнаружил абстрагирование, то возникает повод утверждать креативность. Если  $H$  обнаружил обобщение, то это повод отрицать креативность. Нейронные сети не формируют идеальные объекты, не оперируют ими. Они способны обобщать и ограничивать. Формально-логически это означает то, что машина работает исключительно с объемами понятий, экстенционально. С содержанием понятий, со смыслами, интенционально, работает исключительно человек. Человеку – интенциональное, роботу – экстенциональное.

По сути, вопрос обнаружения когнитивных признаков продуцирования артефакта  $A$  – это вопрос следующего плана: компьютер  $K$  – это роболичность или робозомби? Вопрос демаркации роболичности и робозомби представляется актуальным и сложным [13]. В первом случае мы имеем возможность проиграть аксиоматику всех версий теста Лавлейс и удостоверить то, что робот, сотворивший  $A$ , – это суть искусственная личность. Мы можем понять, кто, как, чем и зачем сотворил шедевр. Во втором случае этого мы не можем сделать. Возможно, мы можем оценить  $A$  как высококреативный продукт. Но мы не можем восстановить траекторию производства «шедевра» и, самое главное, узнать, насколько задействована когнитивная сфера.

Глубокие нейронные сети никаких революций в «искусственном интеллекте», помимо пиара и пропаганды, не совершали. Для креативной робототехники нейроморфная методология ведет к глубочайшей стагнации. Это доведенный до предела физикалистский подход, пытающийся элиминировать человека из современного робототехнического мира. Однако креативная робототехника возможна лишь при активном и постоянном включении человека в технологический процесс. По крайней мере только человек способен определять степень креативности этой технологии. Но нужна новая компьютерная семиотика естественно-искусственной коммуникации «человек-робот». Следует обратить внимание на технические принципы креативных роботов. Эти принципы должны способствовать репрезентации многообразия смыслов креативной коммуникации. Имеется ли путь, избегающий пленения феномена творчества комбинаторикой нейроморфных субобъектов?

### **Технические принципы креативной робототехники: машина Корсакова – Тьюринга**

Воспользуемся исторической находкой последнего десятилетия и введем машину Корсакова в состав теоретико-алгоритмических оснований креативной робототехники. Машина Корсакова – это компьютерный проект российского инженера С.Н. Корсакова, предложенный за одиннадцать лет до «первой» программы Ады Лавлейс, в 1832 году (!). Текст статьи [14] (на французском языке) был обнаружен Г.Н. Поваровым и получил вторую жизнь в начале

2000-х годов [15, 16]. Имеются две независимые альтернативные интерпретации машины Корсакова, предложенные в 2010 году: символическая, основанная на теоретико-множественной парадигме [17] и коннекционистская интерпретация [18]. Авторы настоящей статьи предлагают воспользоваться второй интерпретацией.

Машина Корсакова состоит из пяти механизмов. Они работают как со словами, обозначающими частные идеи, из которых формируется сложная идея, так и с частными признаками идеи, из которых данная идея «конструируется». С.Н. Корсаков впервые в мире применил в информатике перфокарты, поэтому можно полагать, что все современные электронные носители информации являются электронными версиями его перфокарт. Принципы хранения информационного разнообразия не меняются, будь это отверстия в картоне перфокарты, либо, скажем, перепады напряжения, или различия частоты несущего сигнала и пр.

На перфокартах в общем случае представлены не инструкции программы, что характерно для символического подхода. Иголки табуляторов, попадая в отверстия, сделанные перфоратором, фиксируют связи (коннекции) между признаками частных идей. Табулятор проводится по стопке перфокарт, иглы по-разному проваливаются в отверстия, образуя сложные коннекции. Это параллельные сети связей ассоциативных признаков сложной логической конструкции. Таким образом и «вычисляется» сложная идея. Машина Корсакова – это первая в мире интеллектуальная машина; первый персональный (и работающий) суперкомпьютер; формальное определение коннекционистского алгоритма; компьютерная метафора деятельности художника; гибридная коннекционистско-символическая вычислительная система. Поэтому не в современной интеллектуальной робототехнике, но в далекие 1840-е годы сложились две парадигмы компьютерных систем: 1) коннекционистская (С.Н. Корсаков): машина одновременно выдает результаты анализа связей между признаками сложных «знаний»; 2) символическая (Ч. Бэббидж – А. Лавлейс – А. Тьюринг): машина выполняет программу, кодированную специальными символическими конструкциями [19].

Принципиальные отличия машин Корсакова и Тьюринга, соответственно таковы: коннекционизм/символизм; усиление/имитация человеческого интеллекта; активность/пассивность тьюрингового судьи; интерактивность/априорность теста; открытость/замкнутость компьютерной системы; произвольный/ведомый алгоритм; креативность/реактивность действий; формализация работы художника/математика и пр. В этих оппозициях подчеркивается, что машина Корсакова двойственно дополняет машину Бэббиджа – Лавлейс – Тьюринга. Поэтому ее применение подводит к более универсальным теоретико-алгоритмическим принципам построения компьютерных систем, чем те, которые, по сути, не менялись почти два века с момента описания Адой Лавлейс машины Ч. Бэббиджа.

Эти машины по отдельности не способны решить реформаторские задачи построения принципов креативной робототехники. В машине Корсакова отсутствует идея автоматных переходов. Машина Тьюринга, существующая благодаря этим переходам, тем не менее примитивно целочисленно задает лингвистические конструкции. Это не поддерживает представления об индетерминированности, континуальности, многозначности, неформальности творческого процесса. Совместно же машины способны на многое. Самый простой способ комплексного использования машин состоит в следующем: вместо ленты Тьюринга использовать стопку перфокарт Корсакова. Тогда программируемые переходы наполнятся богатым репрезентативным содержанием. Важно, что собственная инструкция автоматного перехода формируется путем обобщения либо ограничения наборов субсимволических дескрипций этой инструкции.

Такая комплексная машина Корсакова – Тьюринга может выступить принципиальным проектом креативного робота. Она способна программировать богатое репрезентативное содержание и охватывать широкую экстенциональную сферу робототехники.

## Заключение

Отсчет проблематики креативной робототехники предлагается вести от машины Р. Луллия, т.е. с конца XIII – начала XIV века. Идея машины, способной отвечать на все возможные запросы благодаря методам компьютерной силлогистики, включала важное положение о соотношении логической комбинаторики, мышления и творчества. Этот вопрос веками обсуждался в узких научных кругах. «Машина творить не может, настоящим творцом является программист», – такой вердикт возможностям машинного творчества был вынесен Адой Лавлейс в 1843 году. А. Тьюринг в 1950 году выдвинул антитезис и предложил символьный и нейросетевой проекты машин, способных имитировать интеллект. В связи с развитием и повсеместным внедрением символьных интеллектуальных систем в конце XX века стало казаться, что А. Тьюринг прав и можно, как это ни парадоксально, симитировать креативность. Тест Лавлейс эту ситуацию исправил, предъявив достаточно жесткие требования к процессу продуцирования креативного артефакта. Однако нейросетевые методы глубокого обучения на базе суперкомпьютерных технологий в 2010-е годы вновь продемонстрировали правоту Тьюринга. Компьютеры стали как бы успешно соперничать с людьми в культурных и гуманитарных сферах: беседовать и рекомендовать, писать картины, сочинять романы, поэмы, пьесы, работать менеджерами, журналистами, преподавателями и пр. Тест Лавлейс 2.0 эту ситуацию исправил – человек, оценивающий креативность, должен специализироваться в очень узком домене креативных артефактов. Тогда этому эксперту следует доверять. Такой эксперт способен воспроизвести аксиоматику теста Лавлейс в части его касающейся. Получается, что Ада Лавлейс вновь права. Ее правоту мы попытались усилить версией 3.0. Новый тест выявляет когнитивные параметры робота, производящего артефакт. Робот может выступать в роли роболичности, если ему приписывается возможность компьютерной имитации мысли, смысла, свободы, личности, творчества. Такой робот интересен на предмет обнаружения креативности произведенных им артефактов. Если, однако, обнаружится отсутствие осмысленности в ходе продуцирования артефакта (например, когда оценщик или судья заметили подмену абстрагирования фактической операцией обобщения), то тогда необходимо заключить, что творец-то артефакта – это на самом деле робозомби. Следовательно, ни о какой робокреативности речи не идет. Здесь надо говорить о человекоглупости того, кто продолжает настаивать на креативности робототехнического продукта в условиях акогнитивного способа его производства.

Вряд ли такой тест (тест Лавлейс 3.0) можно преодолеть. Он задает абсолютную точку измерения «творчества» робота. Теперь можно говорить лишь о степенях робототехнической креативности.

Сегодня, кстати, не реализованы и технические принципы совмещения символьного и коннекционистского подходов. Один из возможных подходов такой методологической интеграции – это теоретико-алгоритмический проект машины Корсакова–Тьюринга. Пока этому проекту далеко до завершения. Поэтому рано говорить об экстенциональном богатстве предметной области креативной робототехники.

Концептуальные принципы креативной робототехники основаны на комплексном применении тьюринговых тестов. Тест Лавлейс – один из подобных тестов. В решении проблем компьютерного творчества он играет определяющую роль. Мы не утверждаем, что концептуальные основы креативной робототехники следует базировать именно на таком тесте Лавлейс в формате его различных версий. Но этот тест или ему подобный необходим в методологическом инструментарии любой креативной технологии. Он предельно обостряет дистинкцию творческого и нетворческого в целостной человеческой деятельности, которая в весьма недалеком будущем будет осуществляться совместно с креативными роботами.

### Список литературы

1. Turing A.M. Computing Machinery and Intelligence // *Mind*. – 1950. – No. 59 (236). – P. 433–460.
2. Алексеев А.Ю. Комплексный тест Тьюринга: философско-методологические и социокультурные аспекты. – М.: ИИнтелЛ, 2013. – 304 с.
3. Алексеев А.Ю. Проблема творчества в исследованиях искусственного интеллекта// Эпистемология креативности / отв.ред. Е.Н. Князева. – М.: «Канон+» РООИ «Реабилитация», 2013. – С. 374–414.
4. Алексеев А.Ю., Янковская Е.А. Феноменология компьютерного творчества: тест Ады Лавлейс // *Философия творчества: материалы всерос. науч. конф.*, 8–9 апреля 2015 г., Институт философии РАН, г. Москва / под ред. Н.М. Смирновой, А.Ю. Алексеева. – М.: ИИнтелл, 2015. – С. 376–394.
5. Алексеев А.Ю., Янковская Е.А. Феноменология компьютерного творчества: тест Ады Лавлейс // *Философия творчества: материалы всерос. науч. конф.*, 8–9 апреля 2015 г., Институт философии РАН, г. Москва / под ред. Н.М.Смирновой, А.Ю.Алексеева. – М.: ИИнтелл, 2015. – С. 376–394.
6. Menabrea L.F. Sketch of the Analytical Engine. Translated and with extensive commentary by Ada Augusta, Countess of Lovelace. *Bibliothèque Universelle de Genève*, October, 1842. – No. 82. – URL: <https://www.fourmilab.ch/babbage/contents.html>.
7. Гутнер Р.С., Полунов Ю.Л. Августа Ада Лавлейс и возникновение программирования // *Кибернетика и логика. Математико-логические аспекты становления идей кибернетики и развития вычислительной техники*. – М.: Наука, 1978. – С. 57–101.
8. Turing A. Intelligent Machinery. In *ced. Machine Intelligence 5*. – Edinburgh; Edinburgh University Press, 1969. – P. 3–23. Originally, a National Physics Laboratory Report, 1948.
9. Брингсйорд С., Беллоу П., Феруччи Д. Творчество, тест Тьюринга и улучшенный тест Лайвлейс / пер. с англ. А. Ласточкина // *Тест Тьюринга. Роботы. Зомби*, под ред. А.Ю. Алексеева. – М.: МИЭМ, 2006. – С. 62–85.
10. Block N. Psychologism and Behaviorism // *Philosophical Review*. – 1981. – No. 90. – P. 5–43.
11. Mark O. Riedl. The Lovelace 2.0 Test of Artificial Intelligence and Creativity. *Proceedings of the AAAI Workshop: Beyond the Turing Test*, Austin. – Texas, 2015. – URL: <https://arxiv.org/abs/1410.6142>.
12. Пожарев Т., Алексеев А.Ю. Креативные мультимедиа: физикалистский и менталистский подходы // *Философия творчества: материалы всерос. науч. конф.*, 8–9 апреля 2015 г., Институт философии РАН, г. Москва / под ред. Н.М. Смирновой, А.Ю. Алексеева. – М.: ИИнтелл, 2015. – С. 418–424.
13. Алексеев А.Ю. Мультиагентные категории робофилософии: роболичность и робо-зомби [Электронный ресурс] // *Искусственные общества*. – 2020. – Т. 15. Вып. 2. – URL: <https://artsoc.jes.su/s207751800009761-6-1/> (дата обращения: 12.10.2020). DOI: 10.18254/S207751800009761-6
14. Karsakof S. Apercu d'un procédé nouveau d'investigation au moyen de machines à comparer les idées. – St. Petersburg, 1832. – 22 p., 2 pl.
15. Поваров Г.Н. С.Н. Корсаков – русский пионер искусственного разума// *Философия искусственного интеллекта: материалы всерос. междисципл. Конф.*, г. Москва, МИЭМ, 17 – 19 января 2005 г. – М.: Изд-во ИФ РАН, 2005. – С. 198.

16. Povarov G.N. Semen Nikolayevich Korsakov: Machines for the Comparison of Ideas. Computing in Russia. Eds. G. Trogemann, A.Y. Nitussov, W. Ernst. – Wiesbaden: Vieweg & Son, 2001.

17. Корсаков С.Н. Начертание нового способа исследования при помощи машин, сравнивающих идеи / пер. с фр. под ред. А.С. Михайлова. – М.: Изд-во МИФИ, 2009. – 44 с.

18. Корсаков С.Н. (1832 г.). Очерк о новом способе исследования посредством машин для сравнения идей / пер. с фр. А.В. Сыромятина // Электронная культура: трансляция в социокультурной и образовательной среде / под ред. А.Ю. Алексеева, С.Ю. Карпук; Моск. гос. ун-т культуры и искусства. – М., 2009. – С. 15–26.

19. Turing A.M. On Computable Numbers, with an Application to the Entscheidungsproblem, 1936 (Received 28 May, 1936. – Read 12 November, 1936). – URL: [http://www.thocp.net/biographies/papers/turing\\_oncomputablenumbers\\_1936.pdf](http://www.thocp.net/biographies/papers/turing_oncomputablenumbers_1936.pdf). (accessed 12 October 2020).

## References

1. Turing A.M. Computing Machinery and Intelligence. *Mind*, 1950, no. 59(236), pp. 433–460.
2. Alekseev A.Iu. Kompleksnyi test T'uringa: filosofsko-metodologicheskie i sotsiokul'turnye aspekty [Complex Turing test: Philosophical, Methodological and Socio-Cultural Aspects]. Moscow, IIntell, 2013, 304 p.
3. Alekseev A.Iu. Problema tvorchestva v issledovaniikh iskusstvennogo intellekta [The Problem of Creativity in Research of Artificial Intelligence]. *Epistemologiya kreativnosti*. Ed. E.N. Kniazeva. Moscow, «Kanon+» Regional'naiia obshchestvennaia organizatsiia invalidov «Reabilitatsiia», 2013, 512 p., pp. 374–414.
4. Alekseev A.Iu., Iankovskaia E.A. Fenomenologiya komp'iuternogo tvorchestva: test Ady Lavleis [Phenomenology of Computer Creativity: Ada Lovelace's Test]. *Filosofiya tvorchestva: Materialy Vserossiiskoi nauchnoi konferentsii, 8-9 apreliia 2015 g.* Institut filosofii RAN, g. Moskva. Eds. N.M.Smirnova, A.Iu.Alekseev. Moscow, IIntell, 2015, 452 p., pp. 376–394.
5. Alekseev A. Iu., Iankovskaia E.A. Fenomenologiya komp'iuternogo tvorchestva: test Ady Lavleis [The Phenomenology of Computer Creativity: Ada Lovelace's Test]. *Filosofiya tvorchestva: Materialy Vserossiiskoi nauchnoi konferentsii, 8-9 apreliia 2015 g.*, Institut filosofii RAN, g. Moskva. Ed. N.M.Smirnova, A.Iu.Alekseev. Moscow, IIntell, 2015, 452 p., pp. 376–394.
6. Menabrea L.F. Sketch of the Analytical Engine. Translated and with extensive commentary by Ada Augusta, Countess of Lovelace. Bibliothèque Universelle de Genève, October, 1842, No. 82, available at: <https://www.fourmilab.ch/babbage/contents.html>.
7. Gutner R.S., Polunov Iu.L. Avgusta Ada Lavleis i vozniknovenie programmirovaniia [Augusta Ada Lovelace and the Emergence of Programming]. *Kibernetika i logika. Matematiko-logicheskie aspekty stanovleniia idei kibernetiki i razvitiia vychislitel'noi tekhniki*. Moscow, Nauka, 1978, 334 p., pp.57 – 101.
8. Turing A. Intelligent Machinery. In *ced. Machine Intelligence 5*. Edinburgh, Edinburgh University Press, 1969, pp. 3–23. Originally, a National Physics Laboratory Report, 1948.
9. Bringsjord S., Bello P., Ferrucci D. Tvorchestvo, test T'uringa i uluchshennyi test Lavleis [Creativity, the Turing test, and the (Better) Lovelace Test]. *Test T'uringa. Roboty. Zombi*. Ed. A.Iu.Alekseev. Moscow, Moskovskii institut elektroniki i matematiki, 2006, pp. 62–85.
10. Block N. Psychologism and Behaviorism. *Philosophical Review*, 1981, no. 90, pp. 5–43.
11. Mark O. Riedl. The Lovelace 2.0 Test of Artificial Intelligence and Creativity. Proceedings of the AAAI Workshop: Beyond the Turing Test, Austin, Texas, 2015, available at: <https://arxiv.org/abs/1410.6142>
12. Pozharev T., Alekseev A.Iu. Kreativnyye mul'timedia: fizikalistskii i mentalistskii podkhody [Creative Multimedia: Physicalist and Mentalist Approaches]. *Filosofiya tvorchestva: Materialy Vserossiiskoi nauchnoi konferentsii, 8-9 apreliia 2015 g.*, Institut filosofii RAN, g. Moskva. Ed. N.M.Smirnova, A.Iu.Alekseev. Moscow, IIntell, 2015, 452 p., pp. 418–424, available at: <https://aintell.info/21/book.pdf>.
13. Alekseev A. Iu. Mul'tiagentnye kategorii robofilosofii: robolichnost' i robozombi [Multiagent Categories of Robophilosophy: Robolism and Robozombies]. *Iskusstvennye Obshchestva*, 2020, vol. 15, iss. 2 available at: <https://artsoc.jes.su/s207751800009761-6-1/> (accessed 12 October 2020). DOI: 10.18254/S207751800009761-6
14. Karsakof S. Apercu d'un procédé nouveau d'investigation au moyen de machines à comparer les idées. St. Petersburg, 1832. 22 p., 2 pl.
15. Povarov G.N. S.N. Korsakov – russkii pioner iskusstvennogo razuma [S.N. Korsakov - Russian Pioneer of Artificial Intelligence.] *Filosofiya iskusstvennogo intellekta. Materialy Vserossiiskoi mezhdistsiplinarnoi konferentsii, g. Moskva, MIEM, 17 – 19 ianvaria 2005 g.* Moscow, Institut filosofii Rossiiskoi akademii nauk, 2005, p. 198.
16. Povarov G. N. Semen Nikolayevich Korsakov: Machines for the Comparison of Ideas. Computing in Russia. Eds. G. Trogemann, A.Y. Nitussov, W. Ernst. Wiesbaden, Vieweg & Son, 2001.
17. Korsakov S.N. Nachertanie novogo sposoba issledovaniia pri pomoshchi mashin, sravnivaiushchikh idei [Outline of a New Way of Research Using Machines that Compare Ideas]. Ed. A.S. Mikhailova. Moscow, Moskovskii inzhenero-fizicheskii institut, 2009, 44 p.
18. Korsakov S.N. (1832 g.). Ocherk o novom sposobe issledovaniia posredstvom mashin dlia sravneniia idei [Essay on a New Way of Research by Means of Machines for Comparing Ideas]. *Elektronnaia kul'tura: transliatsiia v sotsiokul'turnoi i obrazovatel'noi srede*. Eds. A.Iu. Alekseev, S.Iu. Karpuk. Moscow, Moskovskii gosudarstvennyi institut kul'tury, 2009, 260 p., pp.15–26.
19. Turing A.M. On Computable Numbers, with an Application to the Entscheidungsproblem, 1936. Received 28 May, 1936.–Read 12 November, 1936, available at: [http://www.thocp.net/biographies/papers/turing\\_oncomputablenumbers\\_1936.pdf](http://www.thocp.net/biographies/papers/turing_oncomputablenumbers_1936.pdf).