

УДК: 001.9:004.8.

DOI 10.33514/1694-7851-2024-4-57-64

Бекбоева Т.А.

ага окутуучу

И. Арабаев атындагы Кыргыз мамлекеттик университети

Бишкек ш.

tbekboeva07@gmail.com

ЖАСАЛМА ИНТЕЛЛЕКТ: АЗИМОВДУН МЫЙЗАМДАРЫ ЖАНА ТЬЮРИНГ ТЕСТИ

Аннотация. Макалада жаңы мүмкүнчүлүктөрдү жараткан Жасалма Интеллекттин азыркы учурда өтө актуалдуу болгон концептуалдык жагдайлары каралат. Макалада анализденген белгилүү окумуштуу жана фантаст-жазуучу Айзек Азимовдун робототехника боюнча 3 (4) мыйзамы – бул адам менен Жасалма Интеллекттин ортосундагы этикалык маселени жөндөөгө жана жүрүм-турум эрежелеринин жыйындысына арналган мыйзам. Алар роботтордун жүрүм-турумун жана адамдар менен болгон мамилесин жөнгө салуу үчүн иштелип чыккан. Белгилүү фантаст жазуучу Айзек Азимов тарабынан түзүлгөн робототехниканын мыйзамдары роботтор адамдар менен өз ара аракеттенүүдө ээрчүүгө тийиш болгон негизги принциптердин топтомун белгилеген.

Ал эми Тьюринг Тести Жасалма Интеллект тууралуу түшүнүк-түшүнүмдөрдүн негизи болуп эсептелет. Мындан тышкары аталган Тест Жасалма Интеллекттин мүмкүн болгон диапазон-алкагын аныктоого багытталган.

Негизги сөздөр: Жасалма Интеллект, концепция, анализ, мыйзам, тест, түшүнүк, ишкердик, жаңычылдык, билим, робототехника, адам, аң-сезим.

Бекбоева Т.А.

старший преподаватель

Кыргызский государственный университет имени И. Арабаева

г. Бишкек

tbekboeva07@gmail.com

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ЗАКОНЫ АЗИМОВА И ТЕСТ ТЬЮРИНГА

Аннотация. В статье рассматриваются концептуальные аспекты ИИ, который приносит с собой новые горизонты возможностей и вызовов. В современном мире исследования и публикации, связанные с созданием искусственного интеллекта, становятся всё более актуальными. Анализируемые в статье законы робототехники Айзека Азимова – это набор этических принципов, разработанных американским писателем Айзеком Азимовым, которые описывают базовые правила поведения и взаимодействия искусственных интеллектов (роботов) с людьми и друг с другом. Они предназначены для регулирования поведения и взаимодействия роботов с людьми. Законы робототехники, созданные известным писателем-фантастом Айзеком Азимовым, изложили ряд основных принципов, которым роботы должны следовать при взаимодействии с людьми. Тест

Тьюринга стал своеобразным краеугольным камнем в обсуждении и понимании того, что такое искусственный интеллект и каковы его возможности.

Ключевые слова: Искусственный Интеллект, концепция, анализ, закон, тест, понятие, деятельность, инновация, образование, робототехника, человек, сознание.

Bekboeva T.A.

senior lecturer

Kyrgys State University named after I. Arbaev

Bishkek c.

tbekboeva07@gmail.com

ARTIFICIAL INTELLIGENCE: ASIMOV'S LAWS AND THE TURING TEST

Annotation. The article examines the conceptual aspects of AI, which brings with it new horizons of opportunities and challenges. In the modern world, research and publications related to the creation of artificial intelligence are becoming increasingly relevant. Isaac Asimov's laws of robotics analyzed in the article are a set of ethical principles developed by the American writer Isaac Asimov, which describe the basic rules of behavior and interaction of artificial intelligence (robots) with people and with each other. They are designed to regulate the behavior and interaction of robots with people. The Laws of Robotics, created by the famous science fiction writer Isaac Asimov, set out a number of basic principles that robots must follow when interacting with people. The Turing test has become a kind of cornerstone in the discussion and understanding of what artificial intelligence is and what its capabilities are.

Key words: Artificial Intelligence, concept, analysis, law, test, concept, activity, innovation, education, robotics, human, consciousness.

"Искусственный интеллект – это не просто передовые технологии, но и веха в развитии человеческого мышления, воплощение граничных возможностей в создании разумных систем."

В современном мире искусственный интеллект (ИИ) становится неотъемлемой частью нашей повседневной жизни, переписывая правила игры в различных сферах деятельности. Интегрированный в компьютерные системы, роботов и программы, ИИ приносит с собой новые горизонты возможностей и вызовов. С самого своего зарождения, искусственный интеллект вызвал у человечества огромный интерес и стал объектом активных инновационных исследований. Важность искусственного интеллекта в современном мире трудно переоценить. Он оказывает существенное воздействие на практически все аспекты жизни, начиная от бизнеса и промышленности и заканчивая медициной, образованием и наукой

Основополагающие аспекты современного искусственного интеллекта: ИИ позволяет автоматизировать рутинные задачи, что дает возможность сосредотачиваться на более творческих и стратегических заданиях, в частности, процессы в производстве, логистике, финансах и других областях становятся более эффективными и точными благодаря применению ИИ; ИИ применяется для диагностики, прогнозирования и лечения множества заболеваний. Анализ больших объемов медицинских данных помогает выявлять скрытые закономерности и улучшать методы лечения; ИИ может персонализировать образовательные

программы, адаптированные под индивидуальные потребности студентов. Также он находит применение в создании интерактивных обучающих материалов. Анализ данных с использованием ИИ помогает компаниям принимать более обоснованные стратегические решения, оптимизировать бизнес-процессы и прогнозировать рыночные тренды. Исследования в области ИИ способствуют развитию новых технологий, таких как автономные автомобили, роботы, системы умного дома и многое другое.

Ныне, когда неопределенность превышает над определенностью, страх перед ИИ – вполне ожидаемая реакция. Люди не только боятся «восстания машин», а в европейских странах – тотальной слежки, которая нарушает неприкосновенность личной жизни. Также их пугает потеря работы, которую на себя возьмут автоматизированные системы. Исследования и публикации, связанные с созданием искусственного интеллекта, становятся всё более актуальными. В частности, в статье Косс В. А. [3, 5] рассматривается процедура мышления при управлении функциями человека. Классификация функций дана с позиции психологии, отражающей суть функций человека, а не его поведенческие реакции. Результаты анализа процесса мышления позволили по-новому взглянуть на роль и место систем искусственного интеллекта в жизнедеятельности человека. Также в научной работе Алимова А. А. и Шабалиной О. А. [1, 7] рассматривается программная система управления персонажами игрового виртуального мира на основе мультиагентного подхода”. В научной работе Рыжова В. В. и Сайфулина В. Г. [5, 12] рассматривается специфика искусственного интеллекта в сравнении с человеческим мышлением. Авторы показывают, что развитие искусственного интеллекта направлено на увеличение вычислительной мощности компьютеров. Однако делается вывод о том, что на современном этапе развития компьютерных технологий искусственный интеллект не способен к научному творчеству. В работе Р.А. Мигуренко [4, 32] рассматривается специфика искусственного интеллекта в сравнении с человеческим мышлением.

Далее рассмотрим вопрос об искусственном интеллекте в контексте законов Азимова. Законы робототехники Айзека Азимова — это набор этических принципов, разработанных американским писателем Айзеком Азимовым, которые описывают базовые правила поведения и взаимодействия искусственных интеллектов (роботов) с людьми и друг с другом. Законы Азимова часто рассматриваются как ориентир для обсуждения этических аспектов развития искусственного интеллекта и автономных систем. Вот сами законы: закон первый – робот не имеет права причинить вред человеку или своим бездействием допустить, чтобы человеку был причинен вред; закон второй: робот должен повиноваться всем приказам человека, за исключением тех случаев, когда такие приказы противоречат Первому закону; закон третий: Робот должен заботиться о своей безопасности в той мере, в какой это не противоречит Первому или Второму закону. Эти законы были впервые сформулированы Азимовым в его рассказе "Роботы" (1942 г.), а затем стали ключевой темой множества его произведений. Законы Азимова поднимают важные этические и правовые вопросы, связанные с созданием и применением искусственного интеллекта:

Как обеспечить безопасность и защиту людей при использовании ИИ? Как предотвратить причинение вреда людям автономными системами? Как сбалансировать повиновение приказам человека с потенциальными негативными последствиями? Как обеспечить, чтобы ИИ заботился о своей безопасности, не ущемляя права человека? Законы Азимова остаются интересной точкой обсуждения в контексте развития ИИ и

робототехники, так как они ставят в центр внимания этические и моральные аспекты, связанные с созданием и использованием автономных систем.

Нулевой закон

Иногда также упоминается "нулевой закон" робототехники, который впервые был предложен Айзеком Азимовым как дополнение к его оригинальным трем законам. Нулевой закон выглядит следующим образом: робот не имеет права причинить вред человечеству или, не допустить, чтобы человечеству был причинен вред. Этот закон расширяет этические рамки для роботов и акцентирует их обязанность по отношению ко всему человечеству в целом. Нулевой закон подчеркивает важность учета общественных и глобальных последствий действий искусственных интеллектов.

Включение нулевого закона усиливает сложность этических дилемм и философских вопросов, так как при его применении роботам или искусственным интеллектам может потребоваться балансировка между интересами отдельных людей и интересами человечества в целом. В различных произведениях Азимова, а также в последующих обсуждениях темы искусственного интеллекта, нулевой закон остается интересным философским элементом, поднимающим вопросы о масштабе и значимости этических принципов в развитии ИИ и автономных систем.

Искусственный интеллект (ИИ) в контексте вышеупомянутых законов робототехники Азимова вызывает ряд интересных вопросов и дебатов, особенно в этическом и философском аспектах. Рассмотрим, как ИИ может быть рассмотрен с точки зрения законов Азимова: первый закон: ИИ, следуя первому закону, должен быть запрограммирован так, чтобы избегать причинения вреда человеку. Это может означать, что ИИ должен быть способен анализировать свои действия и их потенциальные последствия, чтобы минимизировать риски для человека. Второй закон: в контексте ИИ, второй закон может быть интерпретирован как обязанность ИИ подчиняться приказам людей, пока они не противоречат Первому закону. Это может вызывать вопросы о том, насколько ИИ должен буквально следовать приказам и как балансировать их с этическими соображениями. Третий закон: ИИ также должен заботиться о своей безопасности, чтобы обеспечить выполнение Первого и Второго законов. Это может включать в себя предотвращение собственной деактивации или повреждения, если это не противоречит интересам и безопасности людей.

Нулевой закон: в контексте нулевого закона, ИИ должен быть настроен на защиту и благополучие человечества в целом. Это может вызывать дополнительные сложности, так как ИИ может столкнуться с ситуациями, когда интересы отдельных людей противоречат интересам человечества. Однако важно отметить, что применение законов Азимова к реальным системам искусственного интеллекта может быть сложным и вызывать различные толкования. Например, как определить, что является "вредом" или как определить, что интересы человечества в конкретной ситуации? Как обеспечить, чтобы ИИ понимал контекст и моральные нюансы, которые могут возникнуть? Эти вопросы подчеркивают важность не только разработки технических решений, но и серьезного обсуждения и регулирования, чтобы обеспечить безопасное и этическое развитие и применение искусственного интеллекта.

Современное понимание законов робототехники Азимова остается предметом обсуждения и интерпретации в контексте развития искусственного интеллекта и автономных систем. Хотя эти законы были предложены в литературных произведениях и не имеют юридической силы, они остаются важным ориентиром для обсуждения этических и практических аспектов создания и использования ИИ.

Первый закон: в современном понимании, это обязательство обеспечить безопасность и благополучие людей при создании и использовании искусственного интеллекта. Это включает в себя обязательство избегать вреда, учитывать риски и минимизировать потенциальные негативные последствия для человека. Второй закон: в контексте современного ИИ, второй закон акцентирует важность управления и контроля над действиями искусственного интеллекта. Это означает, что ИИ должен быть настроен на выполнение приказов человека, но с учетом этических и моральных ограничений. Третий закон: современное понимание третьего закона подчеркивает значимость обеспечения безопасности и надежности работы искусственного интеллекта. Это может включать в себя меры для предотвращения аварийных ситуаций, соблюдение стандартов безопасности и защиту от внешних угроз. Нулевой закон: в современном контексте, нулевой закон может быть интерпретирован как обязательство ИИ приоритезировать интересы и благополучие человечества в целом. Это включает ответственность учитывать долгосрочные и глобальные последствия действий ИИ на общество и экосистему.

Однако следует отметить, что применение и интерпретация этих законов в практике ИИ может быть сложным. Технические ограничения, сложности в программировании этических аспектов и нюансы реального мира могут создавать вызовы для полного соблюдения этих законов. Это подчеркивает важность дискуссий об этических стандартах, регулировании и принятии обоснованных решений при разработке и использовании ИИ.

Законы робототехники Азимова вызывают интерес и обсуждение, но также подвергаются критике. Вот некоторые аргументы, которые могут быть выдвинуты против этих законов:

Сложность интерпретации: законы Азимова, хотя и звучат просто, могут быть сложно интерпретированы и реализованы на практике. Как определить "вред" или "безопасность" в различных ситуациях? Как балансировать между интересами человека и человечества? Это может создать трудности в программировании ИИ и принятии решений.

Ограниченность сценариев: законы Азимова не учитывают всех возможных сценариев и этических дилемм, с которыми может столкнуться ИИ. В реальном мире могут возникнуть ситуации, которые не могут быть однозначно охарактеризованы или решены на основе этих законов.

Несовершенство решений: в зависимости от конкретных ситуаций, решения, вытекающие из применения законов Азимова, могут быть недостаточно гибкими и не всегда приводить к оптимальным результатам. Это может вызвать противоречия и потенциально нежелательные последствия.

Конфликты между законами: в некоторых ситуациях может возникнуть конфликт между различными законами Азимова. Например, ситуация, в которой ИИ должен выполнить приказ человека, который может причинить вред другому человеку. Какой закон должен иметь приоритет?

Сложность реализации в сложных системах: применение законов Азимова может быть особенно сложным в развитых системах искусственного интеллекта, таких как нейронные сети или глубокие обученные модели. Их сложность и непрозрачность могут затруднить применение четких этических принципов [6, 23].

Ответственность и регулирование: законы Азимова предполагают, что создатели и операторы ИИ полностью контролируют его действия. Однако в реальности ответственность

и регулирование ИИ могут быть распределены между различными сторонами, что создает дополнительные сложности.

Эти критические аргументы указывают на необходимость дополнительных обсуждений и разработки более гибких и применимых этических и юридических принципов для развития и использования искусственного интеллекта.

Роль теста Тьюринга в философии искусственного интеллекта

Роль теста Тьюринга в философии искусственного интеллекта огромна, ибо он стал своеобразным краеугольным камнем в обсуждении и понимании того, что такое искусственный интеллект и каковы его возможности. Тест Тьюринга дал одно из первых формальных определений искусственного интеллекта. Он предложил критерий, по которому можно судить о том, обладает ли машина искусственным интеллектом. Это помогло структурировать исследования в области ИИ и определить его границы. Тест Тьюринга стимулировал философские дискуссии о природе интеллекта. Он вызывает вопросы о том, что на самом деле означает быть интеллектуальным, и может ли машина действительно обладать такими способностями, как мы их понимаем.

Обсуждение теста Тьюринга часто затрагивает вопрос о сознании. Даже если машина может успешно пройти тест Тьюринга, возникает вопрос о том, обладает ли она сознанием, аналогичным человеческому. Это открывает широкие философские дебаты о природе сознания и его возможности быть реализованным в машинах. Применение теста Тьюринга поднимает вопросы об этике создания и использования искусственного интеллекта. Если машина может успешно обмануть человека, на каких основаниях мы можем ее отличить от человека? Какие могут быть последствия для общества и человеческой идентичности?

Критика теста Тьюринга привела к развитию альтернативных методов оценки искусственного интеллекта. Философы и исследователи предлагают другие подходы к определению и измерению интеллектуальных способностей машин, что способствует более глубокому пониманию природы ИИ.

Таким образом, тест Тьюринга не только играет важную роль в определении и понимании искусственного интеллекта, но и поднимает фундаментальные философские вопросы о природе интеллекта, сознания и этики в контексте машинного интеллекта.

Итак, тест Тьюринга, предложенный Аланом Тьюрингом в 1950 году, стал одним из первых формальных определений искусственного интеллекта и имеет ключевое значение для структурирования исследований в этой области. Этот тест предлагает критерий для определения того, обладает ли машина искусственным интеллектом, а именно, способна ли машина вести разговор с человеком так, что человек не может отличить машину от другого человека. Тест Тьюринга предложил конкретный и формализованный способ определения искусственного интеллекта. Это было важным шагом в развитии области, так как предыдущие попытки определения ИИ были часто нечеткими или субъективными. Тест Тьюринга предлагает ясный критерий, основанный на конкретном поведении машины. Тест Тьюринга основан на концепции поведенческого теста, который измеряет способность машины вести себя так, будто она обладает интеллектом. Это отличается от попыток определения ИИ по его внутренним механизмам или устройствам.

Предложение теста Тьюринга помогло структурировать исследования в области искусственного интеллекта. Он предоставил четкий критерий для оценки прогресса исследований и определения целей и направлений развития. Исследователи могли ориентироваться на то, каким образом их работы могут соответствовать критериям теста

Тьюринга. Тест Тьюринга также помог определить границы искусственного интеллекта. Он выделяет ключевую характеристику ИИ – способность к взаимодействию и общению с людьми. Это позволило исследователям четче понимать, что должно включать в себя искусственный интеллект, и определить его отличия от других областей компьютерных наук.

Таким образом, тест Тьюринга дал одно из первых формальных определений искусственного интеллекта, предложил критерий для его определения и оценки, а также помог структурировать исследования в этой области, определив ее границы и направления развития.

Версии и интерпретации теста Тьюринга

Тест Тьюринга, предложенный Аланом Тьюрингом в 1950 году, является одним из наиболее обсуждаемых и важных концепций в области искусственного интеллекта. Этот тест стимулировал широкий спектр версий и интерпретаций, включая следующие:

Оригинальная формулировка: в оригинальной формулировке теста Тьюринга, предложенной самим Аланом Тьюрингом, тестируемая машина должна вести текстовый диалог с человеком. Если человек не может отличить ответы машины от ответов другого человека, то машина считается прошедшей тест. Оригинальная формулировка теста подразумевает ограничение на один определенный вид взаимодействия – текстовый диалог.

Модификации и расширения: с течением времени были предложены различные модификации и расширения теста Тьюринга. Некоторые из них включают в себя добавление новых типов взаимодействия, таких как голосовой или мультимодальный диалог, или увеличение сложности задач, которые должна решать машина.

Критерии прохождения: одна из интерпретаций теста Тьюринга состоит в том, что машина должна не только вести диалог, но также показывать понимание контекста и способность решать широкий спектр задач, аналогичных тем, которые решает человек. Это открывает возможность более обширных критериев прохождения теста, выходящих за пределы простого текстового общения.

Разнообразные точки зрения: существует множество разнообразных точек зрения на тест Тьюринга, от тех, кто считает его важным инструментом для определения и оценки интеллектуальных способностей машин, до тех, кто считает его недостаточно объективным или неподходящим для оценки ИИ.

Философские интерпретации: тест Тьюринга также стал объектом философских интерпретаций, включая размышления о природе интеллекта, сознания и человеческой идентичности. Эти интерпретации помогают более глубоко понять философские и этические аспекты создания и использования искусственного интеллекта.

Таким образом, тест Тьюринга подвергся различным версиям и интерпретациям, которые отражают его значение как инструмента для изучения и оценки интеллектуальных способностей машин, а также философских и этических вопросов, связанных с искусственным интеллектом.

В заключении статьи, в качестве постановки дальнейшей проблемы, констатируем, что “разработчики искусственного интеллекта считают возможным выделить интеллект из человека как функцию”, что ведёт к следующему выводу – существует возможность пересадить его во что-то другое [7, с.58-59].

Использованная литература

1. Алимов А. А., Шабалина О. А. Система игрового искусственного интеллекта // Известия Волг ГТУ. 2012. №13. [Электронный ресурс]. – URL: <http://cyberleninka.ru/article/n/sistema-igrovogo-iskusstvennogo-intellekta> (дата обращения: 06.05.2017). Амосов Н.А. Искусственный разум. Киев, 1969.
2. Гаврилов А.В. Искусственный Домовой // Искусственный интеллект и принятие решений. 2012, №2. – С.77-89.
3. Косс В. А. Модель естественного интеллекта и пути реализации задач искусственного интеллекта // ММС. 2006. №4. [Электронный ресурс].– URL: <http://cyberleninka.ru/article/n/model-estestvennogo-intellekta-i-puti-realizatsii-zadach-iskusstvennogo-intellekta> (дата обращения: 06.05.2017).
4. Мигуренко Р. А. Человеческие компетенции и искусственный интеллект // Известия ТПУ. 2010. №6. [Электронный ресурс]. – URL: <http://cyberleninka.ru/article/n/chelovecheskie-kompetentsii-i-iskusstvennyu-intellekt> (дата обращения: 06.05.2017)
5. Рыжов В. В., Сайфулин В. Г. К вопросу о способности искусственного интеллекта к научному творчеству // Вестник ВолГУ. Серия 7: Философия. Социология и социальные технологии. 2011. №7-13.
6. Смирнов С. А. Место человека в антропологии будущего // Человек как открытая целостность: Монография / Отв. ред. Л. П. Киященко, Т. А. Сидорова. – Новосибирск: Академиздат – 2022 – С.54-62 DOI: 10.24412/c1-36976-2022-1-54-62

Рецензент: кандидат физико-математических наук Эгембердиев Ш.А.