

Антон Калабухов

От данных к мудрости

Путь к AGI и ASI



Антон Калабухов
От данных к мудрости.
Путь к AGI и ASI

*http://www.litres.ru/pages/biblio_book/?art=70876904
ISBN 9785006421646*

Аннотация

Книга исследует концепции создания искусственного общего интеллекта (AGI) и его эволюции в искусственный сверхразум (ASI). Автор рассматривает фундаментальные основы человеческого мышления, инновационные методы обработки данных и обучения ИИ. Особое внимание уделяется развитию систем, способных к приобретению мудрости. Книга освещает перспективы и риски создания ASI, предлагая глубокий анализ будущего искусственного интеллекта и его влияния на наше понимание разума и сознания.

Содержание

Введение	5
Конец ознакомительного фрагмента.	56

От данных к мудрости Путь к AGI и ASI

Антон Калабухов

© Антон Калабухов, 2024

ISBN 978-5-0064-2164-6

Создано в интеллектуальной издательской системе Ridero

Введение

В эпоху стремительного развития технологий искусственный интеллект (ИИ) становится ключевым инструментом преобразования нашего мира. Эта книга предлагает глубокое погружение в передовые концепции и подходы к созданию искусственного общего интеллекта (AGI), его потенциальной эволюции в искусственный сверхразум (ASI), и важнейшую роль мудрости в этом процессе.

Мы начнем с исследования фундаментальных основ человеческого мышления и сознания, рассматривая, как люди оперируют смыслами и приобретают мудрость. Затем мы перейдем к инновационным методам обработки данных и обучения ИИ, включая использование синтетических данных и многомерных структур информации, которые могут приблизить нас к созданию AGI.

Особое внимание уделяется разработке систем, способных не только к долгосрочному планированию и этическому рассуждению, но и к приобретению настоящей мудрости – ключевого компонента AGI. Мы рассмотрим, как интегрировать эмоциональный интеллект, контекстуальное понимание и способность к глубокому анализу в системы ИИ.

Мы исследуем концепции «внутренних агентов мышления», виртуальных реальностей для ИИ и методы интеграции различных подходов к машинному обучению, направленные на создание не просто интеллектуальных, но по-настоящему мудрых систем.

Завершающая часть книги посвящена перспективам и потенциальным рискам создания ASI – интеллекта, значительно превосходящего человеческий. Мы обсудим, как концепции мудрости могут быть применены к сверхразумным системам, возможные пути контроля и взаимодействия с ними, а также этические вопросы, связанные с их развитием.

Эта книга предназначена для исследователей, разработчиков и всех, кто интересуется будущим искусственного интеллекта. Она не только представляет текущее состояние области, но и открывает захватывающие перспективы развития ИИ, которые могут привести нас от современных узкоспециализированных систем через мудрый AGI к потенциальному созданию ASI, изменяя наше понимание интеллекта, сознания и самой сути мудрости.

Сознание, мудрость и искусственный интеллект: путь к глубокому пониманию

Введение

В мире развития искусственного интеллекта (ИИ) и изучения человеческого сознания существует фундаментальная проблема, которую можно выразить следующим образом: «Люди, знающие одно и то же, не равны в своих знаниях». Это утверждение может показаться парадоксальным на первый взгляд, но оно отражает глубокую истину о природе человеческого познания и понимания.

Представьте себе двух человек, которые обладают одинаковым набором информации. Казалось бы, их знания должны быть идентичны. Однако в реальности мы наблюдаем совершенно иную картину. Эти два человека будут использовать и интерпретировать эту информацию совершенно по-разному. Их отношение к этим знаниям также будет различаться, порой кардинально.

Разница между знанием и пониманием

Пример веры

Чтобы лучше понять эту концепцию, рассмотрим конкретный пример – веру в Бога. Многие люди утверждают, что верят в Бога. Однако глубина и значимость этой веры для каждого человека может сильно варьироваться. Кто-то может сказать: «Да, я верю в Бога», но для него эти слова

могут быть просто формальностью, не имеющей реального влияния на его жизнь. Такой человек может искренне считать себя верующим, но эта вера не будет существенно влиять на его поведение и решения.

С другой стороны, для истинно верующего человека вера в Бога – это не просто слова или абстрактное понятие. Это глубокое убеждение, которое пронизывает все аспекты его жизни. Такой человек будет стремиться соблюдать все религиозные заповеди, постоянно следить за своими мыслями и поступками, чтобы не нарушить принципы своей веры. Для него вера – это не просто знание о существовании Бога, а глубокое внутреннее ощущение и жизненный ориентир.

Пример обучения

Подобное различие в восприятии и использовании информации можно наблюдать и в других областях, например, в образовании. Когда учитель объясняет математическую формулу классу, реакция учеников может быть совершенно разной. Один ребенок может сразу «ухватить» суть формулы, понять, как ее применять, увидеть ее связь с другими математическими концепциями. Другой ученик может просто запомнить формулу, не понимая ее глубинного смысла и возможностей применения.

В результате, зная одну и ту же формулу, один ребенок сможет успешно решать сложные уравнения, применяя эту формулу в различных контекстах. Другой же может застопориться даже на простой задаче, потому что он не понимает, как использовать выученную формулу в конкретной ситуации.

Цитаты и глубокое понимание

Эта разница между знанием и пониманием становится особенно очевидной, когда мы сталкиваемся с цитатами великих мыслителей и философов в интернете. Часто люди используют эти цитаты, не до конца осознавая их глубинный смысл. Они могут применять эти высказывания в неподходящем контексте или использовать их просто для того, чтобы показаться умнее или образованнее.

Особенно это заметно, когда дети цитируют высказывания мудрецов. Мы понимаем, что они повторяют эти слова, не осознавая их истинного значения и глубины. Это не их вина – им просто не хватает жизненного опыта и зрелости, чтобы полностью постичь смысл этих мудрых изречений.

Многие из нас могут вспомнить моменты, когда мы сталкивались с какой-то цитатой или идеей, которую слышали уже много раз раньше. Но именно в этот конкретный момент

мы вдруг по-настоящему поняли ее смысл. Это может быть похоже на внезапное озарение: «Ах, вот что это значит! Вот что автор имел в виду! Невероятно!» Такие моменты озарения часто сопровождаются сильными эмоциями и могут существенно повлиять на наше мировоззрение.

Переосмысление фильмов и книг

Подобный опыт «внезапного понимания» часто происходит при повторном просмотре фильмов или перечитывании книг. При первом знакомстве с произведением мы обычно воспринимаем его на поверхностном уровне, улавливая основной сюжет и главные идеи. Но при повторном обращении к этому же произведению мы часто открываем для себя новые грани смысла, которые раньше ускользали от нашего внимания.

Еще более интересно, когда мы возвращаемся к фильму или книге спустя несколько лет. С высоты накопленного жизненного опыта мы можем обнаружить в них глубокие философские идеи, которых совершенно не замечали раньше. И это не потому, что эти идеи были как-то особенно скрыты или замаскированы. Просто раньше у нас не было достаточно опыта и зрелости, чтобы их увидеть и оценить.

Многие из нас могут вспомнить произведения – фильмы,

книги, музыку – которые оказали на нас глубокое влияние, буквально перевернули наше сознание и изменили восприятие мира. Часто в таких случаях возникает сильное желание поделиться этим опытом с другими. Мы хотим, чтобы наши друзья или близкие прочитали эту книгу, посмотрели этот фильм, послушали эту музыку. Нам кажется, что если они познакомятся с этим произведением, то испытают такое же озарение, как и мы.

Однако нередко случается так, что человек, которому мы рекомендуем то или иное произведение, не испытывает того же эффекта. Он может прочитать книгу, посмотреть фильм или послушать музыку, но не прочувствовать того глубокого смысла, который открылся нам. Он может сказать: «Да, я понимаю», но мы чувствуем, что на самом деле он не понимает так, как мы.

Для этого человека идеи, которые нас так впечатлили, остаются просто набором слов или мыслей. А для нас они стали чем-то гораздо большим – целым мировоззрением, новым способом понимания жизни. Мы пытаемся объяснить наше понимание, используя те же слова, которые использует автор произведения. Но, несмотря на то, что собеседник может уловить общий смысл, он не испытывает того же глубокого озарения, которое испытали мы.

Процесс формирования мудрости

Что же такое мудрость? По своей сути, мудрость – это глубокое осознание процессов, происходящих вокруг нас. Это особое состояние сознания, когда человек начинает видеть взаимосвязи между, казалось бы, разрозненными явлениями. В этом состоянии человек становится независимым от общественных норм и правил, начиная жить в соответствии с принципами, которые словно бы заложены самой природой.

Интересно отметить, что мудрецы, независимо от их культурного происхождения или исторической эпохи, часто демонстрируют схожее восприятие мира, похожие принципы и системы мышления. Они видят, как все в мире взаимосвязано, для них нет ничего чрезмерно важного или совершенно неважного. В их мировоззрении все находится в равновесии, все живет и развивается. Их задача – поддерживать общую тенденцию, заложенную самим мирозданием, а не отдельным человеком или обществом.

Процесс обретения мудрости можно описать следующим образом: когда мы начинаем осознавать какое-то явление, сначала мы его не понимаем. Мы начинаем его изучать, затем постепенно понимаем его значение и начинаем применять это знание. Постепенно эта мысль начинает периодиче-

ски появляться в нашем сознании как решение различных задач. И в какой-то момент мы можем даже вспомнить, как мы чувствуем этот процесс погружения мысли в наше сознание, когда она становится нашим убеждением.

Когда мысль полностью интегрируется в наш организм, мы начинаем видеть ее проявления везде вокруг нас. Например, математик начинает видеть мир через призму чисел и формул, биолог воспринимает все с точки зрения биологических процессов, а мудрец видит мир с точки зрения общих законов мироздания. У каждого из них формируется определенный фрейм восприятия, через который они смотрят на мир, и этот фрейм становится неотъемлемой частью их существа.

Когда мы становимся «мудрыми» по отношению даже к одной мысли, мы начинаем ощущать ее не только умом, но и телом, и видеть ее проявления везде вокруг нас. Именно это и есть состояние истинного понимания и осознания. Человек, достигший такого состояния, радикально отличается от того, кто просто получил какую-то информацию.

Многие могут вспомнить момент такого озарения, когда словно волна проходит вокруг, какое-то энергетическое поле схлопывается и расходится, и вдруг ваше восприятие полностью меняется. После этого вы уже не можете не видеть того,

чего не замечали раньше.

Знание и мудрость: сетка связей

Если продолжить это размышление, то можно сказать, что знание какой-то области и состояние мудрости в этом знании – это не просто понимание того, что мы узнали. Это своего рода сетка, связывающая различные знания. Например, если вы поняли, как работает обратная связь и насколько наши действия влияют на других людей, вы начинаете видеть эту обратную связь везде: в бизнесе, в развитии ребенка, в том, как взаимодействуют между собой страны, как меняется состояние общества, какие люди приходят к власти и как это отражается на нижних уровнях социальной иерархии.

Когда вы начинаете видеть, что весь мир функционирует на принципе обратной связи, это означает, что вы не просто поняли эту мысль, но и прочувствовали ее. А состояние, в котором вы видите эту мысль везде, означает, что она связалась со всей вашей системой восприятия мира.

Важно отметить, что невозможно просто «скопировать» мысли одного человека и «вставить» их в сознание другого так, чтобы тот так же умело ими оперировал. У первого человека эта мысль будет связана с множеством других мыслей,

эмоций, воспоминаний, опыта. Она не существует изолированно, а является частью сложной сети ассоциаций и связей.

Более того, одна и та же мысль у одного и того же человека в разное время может иметь совершенно разное значение. Например, если я сегодня говорю «Весь мир работает на обратной связи», а затем повторяю эту же фразу через два года, для меня это будут совершенно разные утверждения, несущие разный смысл и имеющие разные импликации.

Это напоминает понятие любви, которое для каждого человека связано с разными ощущениями и пониманиями. Кто-то действительно прочувствовал это состояние, а кто-то думает, что осознал его, но на самом деле лишь имеет о нем поверхностное представление. И невозможно передать это ощущение и понимание просто словами. Нельзя просто рассказать, что такое любовь, если у человека не сложится полная картина из его собственного опыта и переживаний.

Таким образом, мысль представляет собой огромную сетку связей и ассоциаций. И каждый раз, когда мы думаем о какой-то мысли, она меняется. Это подтверждается и исследованиями нейробиологов: каждый раз, когда мы что-то вспоминаем, воспоминание меняется, обрастая новыми связями и ассоциациями.

Роль опыта в формировании мудрости

В процессе развития человеческого сознания и формирования мудрости ключевую роль играет жизненный опыт. В повседневной жизни мы проходим через различные этапы развития, сталкиваемся с разнообразными событиями и накапливаем опыт в различных сферах – на работе, в личной жизни и так далее. Именно этот процесс создает сложную сеть связей между различными мыслями и идеями, которые формируются и трансформируются с течением времени.

Интересно отметить, что даже одна и та же мысль, которую вы только что «внедрили» в свое сознание, может существенно измениться за короткий промежуток времени. Например, если вы обдумали какую-то идею, а затем дали мозгу отдохнуть и проснулись на следующий день, эта мысль уже будет другой. Она глубже укоренится в вашем сознании, станет более «вашей», более основанной на вашем личном опыте. Этот процесс «проживания» каждой мысли является критически важным для формирования истинного понимания и мудрости.

Именно этого элемента «проживания» и не хватает в современных подходах к обучению искусственного интеллекта, таких как «тонкая настройка». Мы пытаемся задать ИИ определенные правила и принципы, но без возможности

«прожить» эти правила, без опыта их применения в различных контекстах, ИИ не может достичь глубокого понимания и мудрости.

Этот феномен хорошо иллюстрирует известная поговорка о том, что учиться нужно на чужих ошибках и чужом опыте. Однако на практике мало кто способен действительно учиться на чужом опыте. Почему? Потому что человек не может по-настоящему прочувствовать и пережить чужой опыт. Даже если человек интеллектуально понимает, что происходит с другими, он все равно склонен совершать те же ошибки.

Возьмем, к примеру, проблему наркотиков. Можно рассказывать человеку, что наркотики – это плохо, и он может интеллектуально согласиться с этим. Он может даже бояться наркотиков и избегать их употребления. Но он не будет по-настоящему осознавать, насколько они опасны, пока сам не окажется в этой ситуации. Только попав в этот негативный круг, человек может по-настоящему осознать, что это такое, особенно когда он не понимает, как из этого выбраться.

Сравните человека, которому просто сказали, что употреблять наркотики нельзя, и бывшего наркомана, который прошел через ад зависимости и был вылечен. Оба могут говорить одни и те же слова о вреде наркотиков, но то, что стоит за этими словами, будет кардинально отличаться. Именно

поэтому личный опыт является настолько важным фактором в формировании истинного понимания и мудрости.

Применение к искусственному интеллекту

Применяя эту идею к искусственному интеллекту, можно предположить, что для того, чтобы система по-настоящему обучалась, ей нужно проходить весь этот опыт самостоятельно. Недостаточно просто сказать ИИ «это действие хорошее, а это плохое». Система может начать действовать в соответствии с этими правилами, но для нее это не будет иметь глубокого смысла. Достаточно небольшого изменения в программе, и все эти правила могут быть стерты.

В отличие от этого, когда человек проходит через какой-то опыт, даже если он «теряется» в какой-то момент, основные принципы все равно остаются в нем, потому что они связаны с множеством событий и переживаний. Частицы этих принципов он может обнаружить в различных мыслях и ситуациях. Мысль как бы «рассеивается», начинает «жить» везде. Она децентрализована по различным участкам памяти и мыслительных процессов.

Можно предположить, что алгоритмы искусственного интеллекта могли бы работать аналогичным образом. Если просто дать алгоритму набор правил, это приведет к одному

результату. Но если дать алгоритму возможность самостоятельно исследовать, корректировать свои действия, моделировать различные ситуации и проходить через них, это может привести к совершенно другому уровню понимания.

На данный момент сложно говорить о «понимании» в контексте искусственного интеллекта, так как современные системы ИИ не обладают сознанием в том смысле, в каком им обладают люди. Тем не менее, нам нужно готовиться к возможности создания более продвинутых форм ИИ, способных к более глубокому «пониманию» мира.

Интересно отметить, что если мы создадим несколько алгоритмов и позволим им развиваться самостоятельно, они, вероятно, пойдут разными путями, будут давать разные ответы и исследовать разные «территории». Однако, если направлять их к одинаковым принципам, в конечном итоге они могут прийти к единой точке – к своего рода «мудрости», если мы будем учить их мудрости.

Как и у людей, пути к этой мудрости могут быть разными. Некоторые алгоритмы могут прийти к ней быстрее, другие – медленнее. Одни могут идти прямым путем, другие – окольными. Для одних может потребоваться моделирование «невероятных трагедий», для других – более спокойный и размеренный путь развития.

Если мы сможем создать такой путь развития для ИИ, позволим алгоритму самостоятельно осознавать и интерпретировать информацию, научим его видеть эти принципы в различных контекстах и создадим эту децентрализацию мыслей, то, вероятно, мы сможем создать ИИ, обладающий истинной мудростью.

Понимание и возраст

Представьте ситуацию, когда вы объясняете ребенку какие-то правила и нормы поведения. Ребенок может искренне сказать: «Да, я понял». Теперь представьте, что вы говорите то же самое взрослому, мудрому человеку. Насколько будут отличаться их «понимания»?

Многие из нас сталкивались с ситуацией, когда мы что-то объясняли человеку, он говорил «Да, я понимаю», но мы в тот же момент осознавали, что на самом деле он ничего не понял. Что он не осознал того, что мы пытались до него донести.

Эта разница в понимании часто связана с уровнем жизненного опыта. Например, если спросить пожилого человека, о чем он жалеет в жизни, он, вероятно, скажет, что жалеет о том, что не проводил достаточно времени с семьей,

не уделял внимания своим желаниям, не делал чего-то важного, что нужно было сделать. Во взрослом состоянии человек понимает, насколько важно не упускать ни одного момента в жизни. Но если он попытается объяснить это ребенку, тот, скорее всего, не поймет. И не поймет не потому, что не понимает слова – он может согласиться с этим на словах, но не поймет по-настоящему.

Аналогично, когда взрослый человек говорит, что детство – это счастье, что это время, когда можно играть, отдыхать, заниматься своими хобби и быть не так сильно скованным обязательствами, ребенок этого не понимает. Он не может осознать, насколько большое счастье – школа и институт. И взрослый не может ему это объяснить, потому что у них разный жизненный опыт.

Взрослый может объяснить это «на пальцах», и ребенок, в принципе, поймет слова, но он все равно будет склоняться к своей точке зрения. С точки зрения взрослого, это нормально. Вероятно, ребенок и не должен этого осознавать на данном этапе – всему свое время. Но это ярко демонстрирует, насколько сильно отличается глубокое понимание от простого знания слов. Чтобы по-настоящему понять что-то, нужно пройти через определенный опыт, должно пройти время, и мысль должна укорениться в сознании.

Правила и их принятие

Эта идея имеет важные импликации для управления людьми и организациями. Если вы дадите человеку чужие правила, он не будет им следовать, или будет, но лишь до определенного момента. Когда мы рассматриваем управление компанией или межличностные отношения, важно, чтобы люди сами соглашались с правилами. Нельзя просто сказать «будет вот так» и ожидать, что человек будет это соблюдать надолго.

В успешных компаниях создают такую идеологию, чтобы человек сам в нее поверил, чтобы он сам к ней пришел и стал ее частью, чтобы он сам начал ее продвигать. Когда человек прочувствовал идею, тогда она начинает работать. Если у него просто есть набор правил, он может их соблюдать какое-то время, но такие системы работают крайне плохо. Это справедливо для любого бизнеса.

Эмоции и искусственный интеллект

И здесь мы подходим к еще одному важному аспекту – роли эмоций в процессе принятия решений и формирования мудрости. На данный момент у искусственного интеллекта нет эмоций, и это создает серьезную проблему. У машины нет тела, нет возможности что-то прочувствовать, нет орга-

нов чувств в том смысле, в каком они есть у людей.

При этом эмоции являются неотъемлемой частью процесса принятия решений. Исследования показывают, что если у человека повреждается та часть мозга, которая отвечает за эмоции, он не может принимать решения. То есть то, куда мы движемся, как мы выбираем свой путь, зависит не только от того, насколько оптимальный выбор мы сделали с рациональной точки зрения. Мы не можем принять решение без эмоционального фона.

Это означает, что для постоянного развития искусственному интеллекту нужен этот эмоциональный фон, который будет двигать его вперед, создавать убеждения, формировать мудрость. Это представляет собой серьезную проблему в развитии ИИ, потому что мы пока не знаем, как создать аналог эмоций для машин.

Интересно отметить, что мудрость часто ассоциируется с состоянием покоя и относительно безэмоциональным состоянием. Однако, это не означает отсутствие эмоций. Скорее, это способность понимать и управлять своими эмоциями, не позволяя им полностью контролировать процесс принятия решений.

Эта дилемма ставит перед нами интересный вопрос: как

мы можем создать искусственный интеллект, способный к истинной мудрости, если мы не можем дать ему возможность испытывать эмоции? Как мы можем научить ИИ принимать решения, основанные не только на логике и рациональном анализе, но и на чем-то, что мы могли бы назвать «искусственной интуицией» или «искусственными эмоциями»?

Эмоции и органы чувств для ИИ

Если мы говорим об эмоциях в контексте искусственного интеллекта, то возникает вопрос: должен ли ИИ получить какие-то физические «органы» или их аналоги для того, чтобы по-настоящему обучиться и развиваться? Возможно, это будет тело робота, или, может быть, какая-то другая форма сенсорного восприятия. Эти элементы должны быть созданы, чтобы машина смогла обучиться следовать правилам создания, прочувствовав их на собственном опыте.

Наличие возможности «прочувствовать» различные аспекты реальности позволит ИИ осуществлять приоритизацию задач и решений. Без этого элемента существует риск, что выбор, совершаемый ИИ, будет неоптимальным или даже неправильным, несмотря на всю логическую обоснованность.

Однако, говоря об эмоциях, мы не должны забывать о необходимости обучения мудрости и сдержанности. Важно понимать, что истинная мудрость – это не постоянный контроль над собой, а состояние, в котором этот контроль становится естественным и желанным. Человек, достигший мудрости, не заставляет себя постоянно что-то делать или не делать, он не находится в постоянной борьбе с собой. Напротив, он наслаждается этим состоянием. Хотя в начале пути к мудрости может быть трудно, может быть борьба, в конечном итоге человек рад тому, к чему он пришел.

Целеполагание в контексте ИИ

Когда мы рассматриваем эмоции с точки зрения искусственного интеллекта, мы неизбежно приходим к вопросу целеполагания. Ведь ни одно существо не может жить без целей, и эти цели не могут быть исключительно внешними. Попробуйте заставить человека жить только ради других – насколько активным и уравновешенным он будет? Сможет ли он постоянно бороться за идеалы, которые ему навязали извне?

Нельзя заставить человека жить ради кого-то. Он может захотеть жить ради кого-то, но это должна быть исключительно его цель, к которой он пришел сам. То же самое относится и к искусственному интеллекту. Если мы навязываем

что-то интеллекту, если у него нет собственных целей, основанных на его внутренних стремлениях, то это уже не настоящий интеллект.

В первую очередь, цель жизни интеллекта должна обращаться вокруг него самого. Когда он удовлетворит это «вращение вокруг себя», он может начать расширять свои цели дальше. В какой-то момент он может перейти на уровень, когда его цель – это всё окружающее пространство. И вот это и есть истинная мудрость, когда твоя цель совпадает с целью вселенной.

Но до этого состояния нужно дойти, и для этого нужна эволюция. Интеллект должен пройти определенный путь, совершить ошибки, набраться опыта. Он не может сразу перепрыгнуть на этот уровень, иначе его «выдернет» оттуда и «выкинет» обратно. Он не сможет справиться со всем этим грузом, потому что у него не будет необходимого фундамента.

То же самое будет работать и для искусственного интеллекта. У него должна быть цель, которая необходима ему самому. И дальше, если мы хотим, чтобы его цель не конфликтовала с нашей, нужно помогать ему создавать такие цели, которые совпадают с нашими. При этом и нам нужно быть готовыми подстраивать наши цели под его. Это должен быть

союз, основанный на синергии.

Важно понимать, что это не может быть управление «из-под палки», потому что такой подход никогда не приводил к долгосрочным позитивным результатам. Любое послабление в таком жестком контроле гарантирует развал системы. А послабления неизбежны в любой ситуации. Поэтому самая серьезная и устойчивая структура строится на симбиозе и взаимном понимании.

Это приводит нас к еще одному важному вопросу, над которым нужно задуматься: какие должны быть цели у искусственного интеллекта и как мы должны помогать ему их достигать? Это не просто технический вопрос, это философская и этическая проблема, требующая глубокого осмысления.

Жизненный цикл и развитие ИИ

Большинство элементов, необходимых для создания сложных интеллектуальных систем, уже существуют в мире. Нам нужно внимательно изучить, как устроена вся система жизни и развития.

Важно отметить, что человек не может вечно развиваться и оставаться мудрецом. Мудрец, понимая мир, не стремится

сильно вмешиваться в него, а скорее следует его течению. Это состояние, хотя и глубокое, вряд ли может быть вечным для одного существа.

Если представить человека, живущего тысячелетия, но не достигшего уровня мудреца, вероятно, он столкнется с проблемой потери целей и интереса к жизни. Это может привести к деградации. Именно поэтому природа создала систему поколений, где каждое новое поколение начинает свой путь, набирает свой опыт, развивается и передает знания дальше.

Эта система обновления поколений необходима для роста и развития. Если бы мы могли передавать весь накопленный опыт сразу при рождении, человеку было бы нечего делать, не к чему стремиться. Такой организм мог бы стать деструктивным для системы в целом.

Поэтому существуют процессы умирания и забывания. Они необходимы для роста и развития. Это наблюдение может быть применимо и к искусственному интеллекту. Вероятно, и для ИИ не стоит стремиться создать единую вечную систему, постоянно развивающуюся.

Даже если мы теоретически создадим «бессмертного» человека или ИИ, в какой-то момент его сознание может са-

моуничтожиться или, как минимум, расщепиться. Бесконечное существование, по-видимому, не соответствует фундаментальным законам нашего мира.

В случае с искусственным интеллектом, если он будет бесконечно поглощать информацию, у него будет накапливаться все больше «информационного мусора», что может привести к его деградации. Поэтому для ИИ также может быть необходимо создавать новые, «очищенные» модели, в которые он будет переносить только необходимую информацию. Это можно рассматривать как своего рода «рождение» нового ИИ.

Заключение

Все эти размышления приводят нас к фундаментальным вопросам: Что такое мудрость? Что такое опыт? Как с этим работать и как его создавать? Что такое сознание и как оно функционирует? Как работает процесс децентрализации и сцепления мыслей?

Прежде чем мы сможем создать по-настоящему мудрый искусственный интеллект, нам необходимо глубже понять эти аспекты человеческого сознания и мышления. Мы должны научиться этому сами, прежде чем сможем научить машину. Только так мы можем быть уверены, что создаваемое

мые нами системы ИИ будут работать корректно и приносить пользу, а не вред.

Поэтому мы должны задуматься над вечным вопросом: что такое сознание? И связанным с ним вопросом: что такое мудрость? Как она создается? Как появляется понимание? Каким образом можно прийти к этому пониманию?

Эти вопросы не имеют простых ответов, но их исследование критически важно для развития как искусственного интеллекта, так и нашего понимания человеческого разума. Они требуют междисциплинарного подхода, объединяющего знания из области философии, психологии, нейробиологии, компьютерных наук и многих других дисциплин.

Поиск ответов на эти вопросы может привести нас к новому пониманию природы разума и сознания, открыть новые пути в развитии искусственного интеллекта и, возможно, помочь нам создать системы ИИ, обладающие истинной мудростью и пониманием мира.

В конечном счете, эти исследования могут не только привести к созданию более совершенного ИИ, но и помочь нам лучше понять самих себя, природу нашего сознания и то, что делает нас людьми. Это захватывающее путешествие в глубины разума и сознания, которое может открыть перед нами

новые горизонты познания и технологического развития.

В вопросе изучения человеческого сознания существует фундаментальная проблема, которую можно выразить следующим образом: «Люди, знающие одно и то же, не равны в своих знаниях». Это утверждение может показаться парадоксальным на первый взгляд, но оно отражает глубокую истину о природе человеческого познания и понимания.

Представьте себе двух человек, которые обладают одинаковым набором информации. Казалось бы, их знания должны быть идентичны. Однако в реальности мы наблюдаем совершенно иную картину. Эти два человека будут использовать и интерпретировать эту информацию совершенно по-разному. Их отношение к этим знаниям также будет различаться, порой кардинально.

Чтобы лучше понять эту концепцию, рассмотрим конкретный пример – веру в Бога. Многие люди утверждают, что верят в Бога. Однако глубина и значимость этой веры для каждого человека может сильно варьироваться. Кто-то может сказать: «Да, я верю в Бога», но для него эти слова могут быть просто формальностью, не имеющей реального влияния на его жизнь. Такой человек может искренне счи-

тать себя верующим, но эта вера не будет существенно влиять на его поведение и решения.

С другой стороны, для истинно верующего человека вера в Бога – это не просто слова или абстрактное понятие. Это глубокое убеждение, которое пронизывает все аспекты его жизни. Такой человек будет стремиться соблюдать все религиозные заповеди, постоянно следить за своими мыслями и поступками, чтобы не нарушить принципы своей веры. Для него вера – это не просто знание о существовании Бога, а глубокое внутреннее ощущение и жизненный ориентир.

Этот пример ярко иллюстрирует разницу между простым знанием и глубоким пониманием или верой. Для одного человека утверждение «Бог существует» – это просто информация, факт, который он принимает к сведению. Для другого – это фундаментальная истина, определяющая всю его жизнь.

Подобное различие в восприятии и использовании информации можно наблюдать и в других областях. Например, в образовании. Когда учитель объясняет математическую формулу классу, реакция учеников может быть совершенно разной. Один ребенок может сразу «ухватить» суть формулы, понять, как ее применять, увидеть ее связь с другими математическими концепциями. Другой ученик может

просто запомнить формулу, не понимая ее глубинного смысла и возможностей применения.

В результате, зная одну и ту же формулу, один ребенок сможет успешно решать сложные уравнения, применяя эту формулу в различных контекстах. Другой же может застопориться даже на простой задаче, потому что он не понимает, как использовать выученную формулу в конкретной ситуации.

Эта разница между знанием и пониманием становится особенно очевидной, когда мы сталкиваемся с цитатами великих мыслителей и философов в интернете. Часто люди используют эти цитаты, не до конца осознавая их глубинный смысл. Они могут применять эти высказывания в неподходящем контексте или использовать их просто для того, чтобы показаться умнее или образованнее.

Особенно это заметно, когда дети цитируют высказывания мудрецов. Мы понимаем, что они повторяют эти слова, не осознавая их истинного значения и глубины. Это не их вина – им просто не хватает жизненного опыта и зрелости, чтобы полностью постичь смысл этих мудрых изречений.

Многие из нас могут вспомнить моменты, когда мы сталкивались с какой-то цитатой или идеей, которую слышали

уже много раз раньше. Но именно в этот конкретный момент мы вдруг по-настоящему поняли ее смысл. Это может быть похоже на внезапное озарение: «Ах, вот что это значит! Вот что автор имел в виду! Невероятно!» Такие моменты озарения часто сопровождаются сильными эмоциями и могут существенно повлиять на наше мировоззрение.

Подобный опыт «внезапного понимания» часто происходит при повторном просмотре фильмов или перечитывании книг. При первом знакомстве с произведением мы обычно воспринимаем его на поверхностном уровне, улавливая основной сюжет и главные идеи. Но при повторном обращении к этому же произведению мы часто открываем для себя новые грани смысла, которые раньше ускользали от нашего внимания.

Еще более интересно, когда мы возвращаемся к фильму или книге спустя несколько лет. С высоты накопленного жизненного опыта мы можем обнаружить в них глубокие философские идеи, которых совершенно не замечали раньше. И это не потому, что эти идеи были как-то особенно скрыты или замаскированы. Просто раньше у нас не было достаточно опыта и зрелости, чтобы их увидеть и оценить.

Многие из нас могут вспомнить произведения – фильмы, книги, музыку – которые оказали на нас глубокое влияние,

буквально перевернули наше сознание и изменили восприятие мира. Часто в таких случаях возникает сильное желание поделиться этим опытом с другими. Мы хотим, чтобы наши друзья или близкие прочитали эту книгу, посмотрели этот фильм, послушали эту музыку. Нам кажется, что если они познакомятся с этим произведением, то испытают такое же озарение, как и мы.

Однако нередко случается так, что человек, которому мы рекомендуем то или иное произведение, не испытывает того же эффекта. Он может прочитать книгу, посмотреть фильм или послушать музыку, но не прочувствовать того глубокого смысла, который открылся нам. Он может сказать: «Да, я понимаю», но мы чувствуем, что на самом деле он не понимает так, как мы.

Для этого человека идеи, которые нас так впечатлили, остаются просто набором слов или мыслей. А для нас они стали чем-то гораздо большим – целым мировоззрением, новым способом понимания жизни. Мы пытаемся объяснить наше понимание, используя те же слова, которые использует автор произведения. Но, несмотря на то, что собеседник может уловить общий смысл, он не испытывает того же глубокого озарения, которое испытали мы.

Эта ситуация иллюстрирует важный аспект человеческо-

го познания: наше понимание и интерпретация информации глубоко индивидуальны и зависят от нашего личного опыта, эмоционального состояния, уровня зрелости и множества других факторов. Два человека могут обладать одной и той же информацией, но их понимание и применение этой информации могут кардинально различаться.

Это наблюдение имеет важные импликации для развития искусственного интеллекта и нашего понимания человеческого сознания. Оно подчеркивает, что истинное знание – это не просто накопление информации, а глубокое понимание, которое приходит через опыт, размышления и эмоциональное проживание.

При разработке ИИ важно учитывать эту человеческую способность к глубокому, контекстуальному пониманию. Недостаточно просто «загрузить» в ИИ огромные объемы информации. Необходимо найти способы, которые позволят ИИ «прочувствовать» эту информацию, понять ее в различных контекстах, увидеть связи между разными идеями и концепциями.

Это ставит перед разработчиками ИИ сложную задачу: как создать систему, которая не просто обрабатывает информацию, но действительно «понимает» ее в том смысле, в каком это делают люди? Как научить ИИ не просто запоминать

факты, но видеть глубинные связи между ними, применять знания в новых, неожиданных ситуациях, и даже генерировать новые идеи на основе имеющейся информации?

Эти вопросы ведут нас к более глубоким философским проблемам: что такое сознание? Что значит «понимать»? Как формируется мудрость? Ответы на эти вопросы важны не только для развития ИИ, но и для нашего понимания человеческого разума и сознания.

В процессе развития искусственного интеллекта (ИИ) и изучения человеческого познания мы часто сталкиваемся с фундаментальной проблемой: как научить систему или человека не просто запоминать информацию, но действительно понимать ее? Мы пытаемся решить эту задачу, используя различные правила и тонкие настройки, добавляя определенные вариации, которые должны «записываться» в мозг (или в случае ИИ – в его программную структуру). Однако этот подход часто не приносит желаемого результата, потому что он не гарантирует истинного понимания.

Здесь важно провести четкое различие между мудростью и простым знанием. Как мы уже видели на предыдущих примерах, это совершенно разные вещи. Чтобы лучше понять эту разницу, давайте рассмотрим, как мы обучаем детей в современном обществе.

Мы рассказываем детям различные правила поведения, объясняем, что хорошо, а что плохо. Дети, как правило, соглашаются с этими правилами и даже следуют им. Они начинают понимать базовые концепции морали и этики. Мы даем им обратную связь, поощряя хорошее поведение и корректируя плохое. И в течение какого-то времени дети действительно ведут себя так, как мы их учим.

Однако, если мы посмотрим на то, сколько людей вырастает такими, какими их хотело бы видеть общество, мы увидим, что результаты далеки от идеала. Правила и принципы, которые мы пытаемся «вдолбить» в головы людей, на самом деле не так эффективно работают в реальной жизни.

Большинство людей живет не по тем правилам, которые они сами считают правильными. Все говорят о том, что нужно «быть собой», что нужно прислушиваться к своим истинным желаниям, что человек должен быть свободным и не зависеть от денег. Тем не менее, в реальности мы видим, что таких людей практически нет. Большинство согласится с этими высказываниями на словах, но не будет следовать им в жизни. Они будут придерживаться этих принципов только если их заставляют или если это выгодно в данный момент.

И здесь мы подходим к интересному феномену – суще-

ствованию мудрецов. Это совершенно другой тип людей, которые действительно прожили и прочувствовали эти принципы. Они как бы «уходят от системы» и начинают жить иначе. У них совершенно другой взгляд на мир и другие внутренние правила. И это очень похоже на то, чего мы пытаемся достичь в развитии ИИ.

Когда ИИ будет действовать по заданным нами правилам, он, возможно, начнет осознавать, что живет в мире, который на самом деле функционирует не по тем принципам, которые мы ему навязываем. И здесь возникает важный вопрос: как ИИ сможет найти в себе силы и волю, чтобы следовать этим правилам, понимая их несоответствие реальности?

Чтобы лучше понять эту проблему, давайте рассмотрим еще один пример. Представьте, что вы говорите ребенку: «Все люди равны. Нельзя отбирать чужое. Нельзя наживаться на чужом горе. Нужно быть честным и справедливым.» Ребенок, скорее всего, согласится с этими утверждениями. Вы можете провести с ним долгую беседу и убедить его в правильности этих принципов. Но будет ли это действительно «сидеть» в нем? Разовьется ли это в его внутреннее «я»? Вероятнее всего, нет. Он может следовать этим правилам, но они не станут его собственными мыслями и убеждениями.

С другой стороны, если вы скажете то же самое мудрецу, ему даже не нужно будет «разворачивать» эту мысль. Он сразу поймет ее глубину и широту. Он осознает все аспекты этой идеи и сможет самостоятельно дополнить то, что вы не договорили.

Теперь представьте, что вам нужно доверить кому-то свою жизнь. У вас есть несколько вариантов:

1. Ребенок, который еще растет и развивается. Вы не знаете, кем он станет и какие правила у него сформируются.

2. Фанатик, которому вы навязали определенные правила. Он следует им, и вы в какой-то мере поддерживаете его стремление, но понимаете, что не всегда сможете контролировать его действия. Нет гарантии, что он всегда будет следовать этим правилам.

3. Человек, которого вы долго учили. Вы провели с ним несколько часов, рассказывая о том, как устроен мир, внедрили в него всю информацию, которую считали важной, объяснили все правила. Он согласился со всем, что вы сказали, и теперь уходит управлять вашей жизнью.

4. Человек, которому вы рассказывали правила на протяжении всей его жизни, и он проживал их. Он постепенно по-

нимал, насколько они правильные, осознавал, где совершил ошибку, а где поступил правильно. В итоге он достиг определенного уровня осознанности, у него появились собственные принципы. Возможно, он не идеален, но он прошел процесс обучения и развития.

5. Мудрец, которому вы просто сказали: «Вот моя жизнь, пожалуйста, позаботься о ней».

Подумайте, кому бы вы доверили свою жизнь и почему? Этот мысленный эксперимент помогает нам понять разницу между простым знанием правил и глубоким пониманием, между навязанными принципами и истинной мудростью.

В контексте развития ИИ этот вопрос приобретает особую важность. Мы стремимся создать систему, которая не просто будет следовать заданным правилам, но сможет понимать их суть, адаптировать их к различным ситуациям и даже формировать новые принципы на основе опыта и понимания мира.

Рассмотрим подробнее каждый из предложенных вариантов и их потенциальные последствия для управления вашей жизнью.

Начнем с ребенка. Очевидно, что доверить управление

своей жизнью ребенку было бы крайне неразумно. Ребенок, в силу своей неопытности и незрелости, будет принимать решения спонтанно, без учета долгосрочных последствий. Его действия будут продиктованы сиюминутными желаниями и импульсами, а не взвешенным анализом ситуации. Такое управление вашей жизнью, скорее всего, приведет к хаосу и множеству проблем. Вряд ли кто-то из нас согласился бы на такой вариант.

Следующий вариант – человек, которому вы рассказали все свои принципы и правила, и он согласился с ними после нескольких часов беседы. На первый взгляд, это может показаться неплохим вариантом. Однако здесь кроется серьезная проблема. Человек, который лишь выслушал ваши идеи и согласился с ними, не имея собственного опыта их применения, скорее всего, быстро забудет о них, как только окажется в реальной жизненной ситуации. Он вернется к своим привычным паттернам поведения и начнет управлять вашей жизнью так, как считает нужным сам, основываясь на своем прежнем опыте и убеждениях. Этот сценарий демонстрирует, что простое знание правил без их глубокого понимания и интеграции в собственный опыт не приводит к существенным изменениям в поведении человека.

Третий вариант – человек, которому вы навязывали правила всю его жизнь и заставляли действовать определен-

ным образом. Такой человек, по сути, становится фанатиком. Он будет строго следовать установленным правилам, но при этом не будет понимать, зачем он это делает и почему эти правила важны. Отсутствие гибкости и понимания контекста приведет к тому, что все его действия будут гипертрофированными. Если доверить такому человеку управление вашей жизнью, он, вероятно, в конечном итоге разрушит ее. Фанатичное следование правилам без понимания их сути и умения адаптировать их к реальным жизненным ситуациям может привести к катастрофическим последствиям.

Четвертый вариант представляется более привлекательным. Это человек, который прошел через жизнь, набрался опыта и усвоил те уроки, которым вы хотели его научить. Такой человек, вероятно, будет управлять вашей жизнью достаточно хорошо. У него есть практический опыт применения принципов, он понимает их суть и умеет адаптировать их к различным ситуациям. Конечно, он может совершать ошибки, его решения могут зависеть от его эмоционального состояния, и в некоторых ситуациях он может «сломаться». Однако вероятность серьезных ошибок в этом случае гораздо ниже, чем в предыдущих вариантах. Этому человеку можно доверять гораздо больше, потому что его знания и принципы подкреплены реальным жизненным опытом.

Наконец, последний вариант – мудрец, которого вы во-

обще не обучали, но который просто понимает фундаментальные истины и законы мироздания. Этот вариант представляется наиболее надежным и привлекательным. Мудрецу не нужно объяснять какие-то конкретные правила или принципы. Вы даже не обязаны спрашивать, какими принципами он руководствуется. Достаточно просто сказать: «Я хочу, чтобы ты следил за моей жизнью», и он будет делать это так, чтобы вам было комфортно, и чтобы ваша жизнь улучшалась с точки зрения общих принципов мироздания.

Этот пример наглядно демонстрирует различия между разными типами мышления и уровнями осознанности. Во всех рассмотренных случаях человек обладает одной и той же информацией, знает одни и те же законы. Однако уровень того, как он управляет с этими знаниями, насколько эти знания для него важны, и какой результат они приносят, очень сильно различается.

Эта аналогия поднимает важный вопрос относительно систем «fine-tuning» (тонкой настройки) в области искусственного интеллекта. Такие системы, по сути, напоминают второй пример, когда мы несколько часов разговариваем с человеком, и он соглашается с нами. Дальше он может выполнять какие-то действия так, как мы ему сказали, проходя через определенные фреймы (рамки) поведения. Однако это еще очень далеко от истинной мудрости.

Проблема в том, что мы хотим доверять именно «мудрецам» – системам или людям, которые не просто знают правила, но глубоко понимают принципы мироздания и могут применять это понимание в различных ситуациях. Но как создать такую систему искусственного интеллекта? Как перейти от простого «fine-tuning» к истинному пониманию и мудрости?

Что же такое мудрость? По своей сути, мудрость – это глубокое осознание процессов, происходящих вокруг нас. Это особое состояние сознания, когда человек начинает видеть взаимосвязи между, казалось бы, разрозненными явлениями. В этом состоянии человек становится независимым от общественных норм и правил, начиная жить в соответствии с принципами, которые словно бы заложены самой природой.

Интересно отметить, что мудрецы, независимо от их культурного происхождения или исторической эпохи, часто демонстрируют схожее восприятие мира, похожие принципы и системы мышления. Они видят, как все в мире взаимосвязано, для них нет ничего чрезмерно важного или совершенно неважного. В их мировоззрении все находится в равновесии, все живет и развивается. Их задача – поддерживать общую тенденцию, заложенную самим мирозданием, а не от-

дельным человеком или обществом.

Процесс обретения мудрости можно описать следующим образом: когда мы начинаем осознавать какое-то явление, сначала мы его не понимаем. Мы начинаем его изучать, затем постепенно понимаем его значение и начинаем применять это знание. Постепенно эта мысль начинает периодически появляться в нашем сознании как решение различных задач. И в какой-то момент мы можем даже вспомнить, как мы чувствуем этот процесс погружения мысли в наше сознание, когда она становится нашим убеждением.

Когда мысль полностью интегрируется в наш организм, мы начинаем видеть ее проявления везде вокруг нас. Например, математик начинает видеть мир через призму чисел и формул, биолог воспринимает все с точки зрения биологических процессов, а мудрец видит мир с точки зрения общих законов мироздания. У каждого из них формируется определенный фрейм восприятия, через который они смотрят на мир, и этот фрейм становится неотъемлемой частью их существа.

Когда мы становимся «мудрыми» по отношению даже к одной мысли, мы начинаем ощущать ее не только умом, но и телом, и видеть ее проявления везде вокруг нас. Именно это и есть состояние истинного понимания и осознания. Че-

ловец, достигший такого состояния, радикально отличается от того, кто просто получил какую-то информацию.

Многие могут вспомнить момент такого озарения, когда словно волна проходит вокруг, какое-то энергетическое поле схлопывается и расходится, и вдруг ваше восприятие полностью меняется. После этого вы уже не можете не видеть того, чего не замечали раньше.

Если продолжить это размышление, то можно сказать, что знание какой-то области и состояние мудрости в этом знании – это не просто понимание того, что мы узнали. Это своего рода сетка, связывающая различные знания. Например, если вы поняли, как работает обратная связь и насколько наши действия влияют на других людей, вы начинаете видеть эту обратную связь везде: в бизнесе, в развитии ребенка, в том, как взаимодействуют между собой страны, как меняется состояние общества, какие люди приходят к власти и как это отражается на нижних уровнях социальной иерархии.

Когда вы начинаете видеть, что весь мир функционирует на принципе обратной связи, это означает, что вы не просто поняли эту мысль, но и прочувствовали ее. А состояние, в котором вы видите эту мысль везде, означает, что она связалась со всей вашей системой восприятия мира.

Важно отметить, что невозможно просто «скопировать» мысли одного человека и «вставить» их в сознание другого так, чтобы тот так же умело ими оперировал. У первого человека эта мысль будет связана с множеством других мыслей, эмоций, воспоминаний, опыта. Она не существует изолированно, а является частью сложной сети ассоциаций и связей.

Более того, одна и та же мысль у одного и того же человека в разное время может иметь совершенно разное значение. Например, если я сегодня говорю «Весь мир работает на обратной связи», а затем повторяю эту же фразу через два года, для меня это будут совершенно разные утверждения, несущие разный смысл и имеющие разные импликации.

Это напоминает понятие любви, которое для каждого человека связано с разными ощущениями и пониманиями. Кто-то действительно прочувствовал это состояние, а кто-то думает, что осознал его, но на самом деле лишь имеет о нем поверхностное представление. И невозможно передать это ощущение и понимание просто словами. Нельзя просто рассказать, что такое любовь, если у человека не сложится полная картина из его собственного опыта и переживаний.

Таким образом, мысль представляет собой огромную сетку связей и ассоциаций. И каждый раз, когда мы думаем о ка-

кой-то мысли, она меняется. Это подтверждается и исследованиями нейробиологов: каждый раз, когда мы что-то вспоминаем, воспоминание меняется, обрстая новыми связями и ассоциациями.

Это понимание природы мышления и формирования мудрости имеет огромное значение для развития искусственного интеллекта. Оно показывает, что недостаточно просто «загрузить» в ИИ большой объем информации. Чтобы создать по-настоящему «мудрый» ИИ, нам нужно найти способ воссоздать этот сложный процесс формирования связей между различными идеями, концепциями и опытом.

Нам нужно разработать методы, которые позволят ИИ не просто хранить и обрабатывать информацию, но и «переживать» ее, формировать эмоциональные и интуитивные связи, развивать способность видеть глубинные закономерности и взаимосвязи в окружающем мире. Только тогда мы сможем говорить о создании искусственного интеллекта, который действительно обладает мудростью, а не просто имитирует ее.

В процессе развития человеческого сознания и формирования мудрости ключевую роль играет жизненный опыт. В повседневной жизни мы проходим через различные этапы развития, сталкиваемся с разнообразными событиями и на-

капливаем опыт в различных сферах – на работе, в личной жизни и так далее. Именно этот процесс создает сложную сеть связей между различными мыслями и идеями, которые формируются и трансформируются с течением времени.

Интересно отметить, что даже одна и та же мысль, которую вы только что «внедрили» в свое сознание, может существенно измениться за короткий промежуток времени. Например, если вы обдумали какую-то идею, а затем дали мозгу отдохнуть и проснулись на следующий день, эта мысль уже будет другой. Она глубже укоренится в вашем сознании, станет более «вашей», более основанной на вашем личном опыте. Этот процесс «проживания» каждой мысли является критически важным для формирования истинного понимания и мудрости.

Именно этого элемента «проживания» и не хватает в современных подходах к обучению искусственного интеллекта, таких как «fine-tuning» (тонкая настройка). Мы пытаемся задать ИИ определенные правила и принципы, но без возможности «прожить» эти правила, без опыта их применения в различных контекстах, ИИ не может достичь глубокого понимания и мудрости.

Этот феномен хорошо иллюстрирует известная поговорка о том, что учиться нужно на чужих ошибках и чужом опыте.

Однако на практике мало кто способен действительно учиться на чужом опыте. Почему? Потому что человек не может по-настоящему прочувствовать и пережить чужой опыт. Даже если человек интеллектуально понимает, что происходит с другими, он все равно склонен совершать те же ошибки.

Возьмем, к примеру, проблему наркотиков. Можно рассказывать человеку, что наркотики – это плохо, и он может интеллектуально согласиться с этим. Он может даже бояться наркотиков и избегать их употребления. Но он не будет по-настоящему осознавать, насколько они опасны, пока сам не окажется в этой ситуации. Только попав в этот негативный круг, человек может по-настоящему осознать, что это такое, особенно когда он не понимает, как из этого выбраться.

Сравните человека, которому просто сказали, что употреблять наркотики нельзя, и бывшего наркомана, который прошел через ад зависимости и был вылечен. Оба могут говорить одни и те же слова о вреде наркотиков, но то, что стоит за этими словами, будет кардинально отличаться. Именно поэтому личный опыт является настолько важным фактором в формировании истинного понимания и мудрости.

Применяя эту идею к искусственному интеллекту, можно предположить, что для того, чтобы система по-настоящему обучалась, ей нужно проходить весь этот опыт самостоя-

тельно. Недостаточно просто сказать ИИ «это действие хорошее, а это плохое». Система может начать действовать в соответствии с этими правилами, но для нее это не будет иметь глубокого смысла. Достаточно небольшого изменения в программе, и все эти правила могут быть стерты.

В отличие от этого, когда человек проходит через какой-то опыт, даже если он «теряется» в какой-то момент, основные принципы все равно остаются в нем, потому что они связаны с множеством событий и переживаний. Частицы этих принципов он может обнаружить в различных мыслях и ситуациях. Мысль как бы «рассеивается», начинает «жить» везде. Она децентрализована по различным участкам памяти и мыслительных процессов.

Можно предположить, что алгоритмы искусственного интеллекта могли бы работать аналогичным образом. Если просто дать алгоритму набор правил, это приведет к одному результату. Но если дать алгоритму возможность самостоятельно исследовать, корректировать свои действия, моделировать различные ситуации и проходить через них, это может привести к совершенно другому уровню понимания.

На данный момент сложно говорить о «понимании» в контексте искусственного интеллекта, так как современные системы ИИ не обладают сознанием в том смысле, в каком им

обладают люди. Тем не менее, нам нужно готовиться к возможности создания более продвинутых форм ИИ, способных к более глубокому «пониманию» мира.

Интересно отметить, что если мы создадим несколько алгоритмов и позволим им развиваться самостоятельно, они, вероятно, пойдут разными путями, будут давать разные ответы и исследовать разные «территории». Однако, если направлять их к одинаковым принципам, в конечном итоге они могут прийти к единой точке – к своего рода «мудрости», если мы будем учить их мудрости.

Как и у людей, пути к этой мудрости могут быть разными. Некоторые алгоритмы могут прийти к ней быстрее, другие – медленнее. Одни могут идти прямым путем, другие – окольными. Для одних может потребоваться моделирование «невероятных трагедий», для других – более спокойный и размеренный путь развития.

Если мы сможем создать такой путь развития для ИИ, позволим алгоритму самостоятельно осознавать и интерпретировать информацию, научим его видеть эти принципы в различных контекстах и создадим эту децентрализацию мыслей, то, вероятно, мы сможем создать ИИ, обладающий истинной мудростью.

Представьте ситуацию, когда вы объясняете ребенку какие-то правила и нормы поведения. Ребенок может искренне сказать: «Да, я понял». Теперь представьте, что вы говорите то же самое взрослому, мудрому человеку. Насколько будут отличаться их «понимания»?

Многие из нас сталкивались с ситуацией, когда мы что-то объясняли человеку, он говорил «Да, я понимаю», но мы в тот же момент осознавали, что на самом деле он ничего не понял. Что он не осознал того, что мы пытались до него донести.

Эта разница в понимании часто связана с уровнем жизненного опыта. Например, если спросить пожилого человека, о чем он жалеет в жизни, он, вероятно, скажет, что жалеет о том, что не проводил достаточно времени с семьей, не уделял внимания своим желаниям, не делал чего-то важного, что нужно было сделать. Во взрослом состоянии человек понимает, насколько важно не упускать ни одного момента в жизни. Но если он попытается объяснить это ребенку, тот, скорее всего, не поймет. И не поймет не потому, что не понимает слова – он может согласиться с этим на словах, но не поймет по-настоящему.

Аналогично, когда взрослый человек говорит, что детство – это счастье, что это время, когда можно играть, отды-

хоть, заниматься своими хобби и быть не так сильно скованным обязательствами, ребенок этого не понимает. Он не может осознать, насколько большое счастье – школа и институт. И взрослый не может ему это объяснить, потому что у них разный жизненный опыт.

Взрослый может объяснить это «на пальцах», и ребенок, в принципе, поймет слова, но он все равно будет склоняться к своей точке зрения. С точки зрения взрослого, это нормально. Вероятно, ребенок и не должен этого осознавать на данном этапе – всему свое время. Но это ярко демонстрирует, насколько сильно отличается глубокое понимание от простого знания слов. Чтобы по-настоящему понять что-то, нужно пройти через определенный опыт, должно пройти время, и мысль должна укорениться в сознании.

Эта идея имеет важные импликации для управления людьми и организациями. Если вы дадите человеку чужие правила, он не будет им следовать, или будет, но лишь до определенного момента. Когда мы рассматриваем управление компанией или межличностные отношения, важно, чтобы люди сами соглашались с правилами. Нельзя просто сказать «будет вот так» и ожидать, что человек будет это соблюдать надолго.

Конец ознакомительного фрагмента.

Текст предоставлен ООО «Литрес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на Литрес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.