Физика субъективного

А. В. Юров[i]

[i]Балтийский федеральный университет им. И. Канта
Калининград

Аннотация: Обсуждается проблема «субъективного» в физике. Показано, что некоторый класс квантовых автоматов (автоматы Альберта) способен демонстрировать нечто вроде «субъективного опыта». Делается предположение, что такие автоматы способны генерировать геделевские суждения. Дополнительно демонстрируется, что учет таких необычных свойств автоматов Альберта способен с новой стороны осветить проблему соотношения квантовых и классических свойств.

Ключевые слова: философия физики, философия сознания, Р. Пенроуз, гёделезация, квантовые автоматы Альберта.

Введение

В своей знаменитой статье «Пределы науки» Е. Вигнер написал следующее: «…и физика и психология претендуют на роль всеохватывающих, универсальных дисциплин: первая — потому, что она стремиться описать всю природу, вторая — потому, что рассматривает все явления, связанные с духовной деятельностью, а природу считает существующей для нас лишь постольку, поскольку мы познаем ее» (Вигнер, 1971, с. 177). И далее по тексту: «Однако чрезвычайно трудно, а может быть, даже и невозможно воспринимать эти две картины как различные аспекты одного и того же предмета» (там же).

Указанное противоречие в чрезвычайно острой форме конкретезировано Э. Шредингером: «Если вы спросите у физика, что, в его понимании, есть желтый свет, он вам ответит, что это поперечные электромагнитные волны, длина которых примерно равна 590 нм. Если вы спросите его: "а где тут желтый?", то он ответит: "в моей картине его нет совсем , но когда эти колебания попадают на сетчатку здорового глаза, у человека, которому принадлежит этот глаз, возникает ощущение желтого цвета."… Ощущение цвета невозможно объяснить в рамках объективной картины волн света, имеющейся у физиков… Все что я хочу сказать — мы можем быть уверены, что не существует такого нервного процесса, объективное описание которого содержит характеристику "желтый цвет" или "сладкий вкус", точно так же, как объективное описание электромагнитных волн не содержит никакой из этих характеристик» (Шредингер, 2000, с. 82).

Из (немногочисленных) современных физиков, высказывающих озабоченность по поводу такого разрыва в описании Природы, можно упомянуть Р. Пенроуза, А. Линде и Дона Пэйджа. Наиболее радикален Р. Пенроуз, который полагает, что для описания сознательной (а стало быть и чувственной) деятельности нам требуется новая физика, оперирующая невычислимыми процессами. Правда, для обоснования такого, прямо скажем, необычного утверждения, Пенроуз обращается не к загадке чувственных качеств (как характеризовал проблему Шредингер), а к загадке наличия у человека способности «геделезировать», т. е. составлять геделевские суждения. Другими словами, Пенроуз анализирует наиболее рациональную способность людей — умение понимать смысл математических утверждений. Вместе с тем, Пенроуз, недвусмысленно утверждает, что такое туманное явление, как «осознание» (а значит и чувственное восприятие, что бы под ним не понимали) есть следствие проявлений неведомых пока физических явлений (т. н. R-процедуры), носящих, как мы у же говорили, невычислимый, т. е. существенно неалгоритмический характер.

Насколько мы понимаем, точка зрения Пенроуза не пользуется популярностью у физиков. Более того, большинство вообще не считают противоречие между «объективным» и «субъективным» проблемой. Господствующую точку зрения можно наверно выразить так: разумеется, все что мы называет психическими и чувственными процессами суть проявления физических процессов в мозге, а значит могут быть поняты на основе наших знаний об этих процессах. Никакое другое знание нам не требуется. Что касается противоречий, которые так беспокоили Шредингера, то у нас с этим нет никаких проблем, поскольку мы сторонники функционального подхода. Это означает, что если кто-то построит действующую модель мозга, будь то нейронная сеть или хитроумная обучающаяся интерактивная программа, и эта модель будет вести себя так как настоящий мозг, в том смысле, что использование, скажем теста Тьюринга, не позволит отличить ее от «настоящего мозга», то мы должны признать, что программа чувствует, осознает и понимает. Именно против такой точки зрения и выступает Пенроуз, хотя и не особенно убедительно, судя по реакции большинства.

Несложно понять причины этого. Дело в том, что физика, по своей сути наука об объективных закономерностях. Слово «субъективный» (а именно к такому классу относятся понятия «желтый» или «сладкий») в физике всегда носит негативный характер, означая, в лучшем случае нечто не вполне достоверное и надежное, а в худшем нечто абсолютно ложное и лишь кажущееся истинным. Происходит это оттого, что, как учит нас опыт, очень многие интуитивные (сиречь — субъективные) представления действительно оказываются совершенно неверными. Ярким примером является субъективное понятие одновременности. Если наблюдатель A покоится, а B двигается относительно A с некоторой скоростью вдоль, скажем, оси OX, то события одновременные для A не будут таковыми для B. Есть лишь одно событие относительно которого у

А и В не возникнет разногласий, по поводу того, что это событие происходит «сейчас»: это точка в пространстве-времени где мировые линии А и В пересекаются. Относительно всех остальных событий происходящих вдоль прямой ОХ мнения А и В разойдутся: В считает, что «сейчас» происходит множество событий, половина которых для А уже произошла, а вторая половина еще произойдет! И это не иллюзия, а прямое следствие специальной теории относительности. Очевидно, что в данном конкретном случае наш прежний субъективный опыт оказывается просто неверным. Мы думаем, что не погрешим против истины предполагая, что, с точки зрения большинства физиков понятия типа «желтый» или «сладкий» являются примерами аналогичными понятию «одновременный» в дорелятивистской физике. Коль скоро, физика демонстрирует нам ущербность многих «очевидных» представлений (вроде одновременности), нам не следует вообще доверять субъективным ощущениям.

Однако можно сомнение в том, что это действительно правильная позиция. Нам кажется, что с такими понятиями-ощущениями, как «желтый-сладкий» действительно возникает проблема, в отличии от понятия «одновременный» и проблема эта относится именно к физике, а не к психологии.

Дело в том, что СТО не только изменила понятие одновременности, но и объяснила, откуда возникает иллюзия всеобщей одновременности, т. е. иллюзия того, что события происходящие «сейчас» для меня происходят «сейчас» и для всех остальных людей. Упомянутое выше расхождение точек зрения А и В тем больше, чем дальше в пространстве удалены события о которых идет речь. Более точно, если А и В говорят о том, что происходит сейчас в точке удаленной от них на расстояние 2 миллиона световых лет, и В движется относительно А со скоростью пешехода, то их мнения разойдутся на несколько суток. Это расхождение Δt прямо пропорционально пространственному расстоянию от наблюдателей до точки в пространстве о которой идет речь. Подчеркнем, что и А и В могут в этот момент находится в одном месте, поэтому речь не идет о какой то иллюзии.

С другой стороны, величина этого расхождения чрезвычайно мала на обычных земных расстояниях и именно поэтому мы так долго не знали об этом удивительном факте. Таким образом, хотя СТО и показала нам абсолютную ложность наших интуитивных (субъективных) представлений об одновременности, но она же и объяснила почему эти ложные представления казались нам такими надежными и правильными. Наш привычный мир получается при устремлении скорости света к бесконечности. Другими словами, СТО содержит в себе классическую механику[1], как предельный случай.

Ничего подобного нет в отношении таких понятий-ощущений, как «желтый-сладкий»[2]. У нас просто нет самостоятельной, пусть и неверной, теории в которой фигурировали бы эти качества, и которую можно было бы потом получить, как предельный случай более общей и правильной теории. Более того, ничто

[1]Конечно, только частично, так в СТО нет потенциальной энергии и сил. Чтобы получить всю классическую механику, как предельный случай нужна ОТО.

[2]Мы избегаем использовать популярный термин «квалиа» и говорим просто об «ощущениях».

не предвещает появление подобной теории в будущем. Физики интенсивно работают над единой теорией всех взаимодействий, которая, будучи построенной, должна венчать физическую картину мира. Однако поле приложения этой единой теории безмерно далеко от масштабов на которых происходит чувственное осознание человека. Отсюда следует, что даже единая теория, если она будет создана окажется перед той же проблемой, с которой мы сталкиваемся сейчас. Так что едва ли нам надо искать выход в этом направлении.

Возможно кто то скажет, что нам и не надо пытаться понять «ощущения» с точки зрения физики, по той причине, что это не входит в компетенцию физики. Однако вряд ли с этим можно согласиться. Если признать такую (на первый взгляд — разумную) позицию, то следует признать, что в нашей картине мира есть пробел. В самом деле, мозг состоит из атомов, динамика которых описывается уравнением Шредингера, а значит все что происходит с мозгом можно понять с помощью квантовой механики. Если же нет, то где гарантии, что аналогичные проблемы не возникают и «вне мозга»? Вся физика становится ущербной, поэтому мы не можем согласиться с такой точкой зрения и предполагаем, что ощущения типа «желтый-сладкий» на самом деле можно описать в рамках физики. Более того, это можно сделать в рамках уже известной физики (а именно — квантовой механики). Именно для доказательства этой точки зрения и была написана эта статья.

1 Стратегия

Однако как доказать такое утверждение? Наша цель показать, что это можно сделать в принципе. Мы не знаем, что значит «желтый», однако объяснение этого и не входит в наши задачи. А задача выглядит так: привести аргументы, показывающие, что используя известную нам физику, кто то в будущем сможет объяснить, что значит «желтый» или «сладкий». Другими словами нас интересует только принципиальный аспект. Детали же могут оказаться очень сложными и на настоящий момент мы не готовы обсуждать их.

Вот типичный пример: статистическая физика не является новой физикой по сравнению с механикой. Однако для описания газов мы используем именно статистическую физику. Ясно, что в принципе мы могли бы получить те же результаты (и, наверное даже большие) напрямую решая динамические уравнения описывающие движение атомов, но это нам не под силу. В этом смысле, механика (классическая и квантовая) в принципе решают проблему описания газов. На практике же потребовались колоссальные усилия Больцмана и Гиббса, для того чтобы получить это описание. Именно о таком принципиальном подходе мы и будем говорить ниже. Наша цель показать, что разрыв между «объективным» миром физики и «субъективным» миром ощущений в принципе можно устранить в рамках уже известных физических теорий.

Этот разрыв в описании похож на знаменитую проблему согласования квантовых и классических свойств[3]. В самом деле, мы хорошо знаем, что классиче-

[3]только в отличии от последней не является сколько нибудь точно определенной.

кую физику нельзя получить из квантовой просто устремив постоянную Планка к нулю. Часто встречаются утверждения, что переход к классике осуществляется с помощью декогеренции. Однако это сильное преувеличение. Декогеренция действительно демонстрирует каким образом чистое состояние превращается в смесь, но не отвечает на главный вопрос: как происходит отбор альтернатив. Наверняка многие физики считают, что ответа на этот вопрос просто не существует. Может быть и так, но в этом случае мы получаем принципиально неустранимый разрыв между классической и квантовой теориями. Это может означать или то, что (i) классическая теория вообще неверна, т. к. не получается из квантовой предельным переходом, или то (ii) что нам следует применять «принцип дополнительности» или то, что (iii) должна существовать более общая теория включающая в себя как два разных предельных случая квантовую теорию и классическую механику (или даже ОТО). Последняя точка зрения согласуется с позицией Пенроуза, но не физического большинства; вторая, понемногу потеряла популярность и, насколько мы понимаем, мало кем поддерживается всерьез[4]; первая же выглядит абсолютно нелепо. Мы убеждены, что мало кто из физиков всерьез согласится рассматривать вариант (i). Однако в разделе 6 данной работы мы приведем некоторые аргументы, которые, как нам кажется, можно использовать для защиты варианта (i).

Теперь заметьте, что какую бы точку зрения вы не избрали — вы в принципе решили задачу. Если вы приняли (i) или (iii) то вы при этом остаетесь в рамках физики (хотя в случае (iii) еще неизвестной). В случае же (ii) решение лежит вне рамок физики, но тем не менее это тоже решение!

Вот нечто подобное и надо сделать с загадкой «чувственных качеств».

2 Тактика

Для того чтобы сделать первый шаг, можно было бы попытаться свести задачу «желто-сладкого» к какой нибудь другой, пока не решенной, но физической задаче. Например, можно предположить, в порядке чистой спекуляции, что
(A) в нашем мозгу действует какой то квантовый чип (как его назвал Ю. Д. Манин), одним из проявлений работы которого и будет ощущение электромагнитных волн с длиной 590 нм, как «желтых».

Если бы и впрямь выяснилось, что мир наших субъективных ощущений порождается какими то квантовыми процессами в нервной системе, но эти ощущения относятся к классически воспринимаемому миру, то действительно можно было бы объявить проблему чувственных качеств физической и даже усилить это замечанием, что в действительности речь идет о старой пробле-

[4]Точка зрения (ii) никак не может быть удовлетворительной. Может быть природа и впрямь требует применять в одном случае классическое, а в другом квантовое описание, но откуда мы знаем когда какое правило надо применять? Интуиция и пр. не могут рассматриваться серьезно, т. к. если признать, что формирование интуиции физика лежит вне физики, то мы становимся мистиками. Если же признать, что есть некие физические законы, которые позволяют нам решать, когда какое описание применять, то эти законы явно должны лежать вне квантовой и классической физики. А значит мы получаем вариант (iii)

ме согласования квантовых и классических свойств! Однако начинать с таких предположений будет глубоко ошибочным, по крайней мере в данном конкретном случае.

В самом деле, ощущение «желтого» не может быть просто «чем то квантовым», потому что желтым мы называем совершенно классический свет. Однако оно не может быть и классическим, поскольку в классике нет никакого желтого, а есть только электромагнитные волны с длиной 590 нм. Таким образом возникает замкнутый круг.

Вместе с тем отбрасывать (A) тоже не стоит. Просто не надо использовать его как базовое допущение. Начинать надо с чего то другого, в надежде, что (A) получится как следствие. В этом случае у нас есть надежда, что в новом качестве (A) приобретет статус объяснения.

С чего же начинать? В данной работе мы будем использовать результаты полученные Альбертом (Albert, 1983) при описании им квантовых автоматов очень специального вида. Эти автоматы способны производить некоторые измерения над собой. В результате оказывается, что автомат приобретает некоторое знание, которое не может быть доступным никакому другому автомату, вне зависимости от его устройства! Таким образом, здесь возникает нечто похожее на «субъективное знание». Действительно, субъективными ощущениями мы называем как правило те ощущения, которые переживаем сами. Нет никакого способа передать именно это ощущение кому то другому. Собственно именно поэтому мы и называем такие вещи «субъективными». Разумеется, если я вижу «желтое», то я предполагаю, что мой сосед тоже видит «желтое», но точно знать я этого не могу. В этом то и проблема! Поэтому, если вдруг выясняется, что в квантовой теории возникает нечто похожее, то нам следует обратить на это самое пристальное внимание.

Мы вовсе не утверждаем, что мы «квантовые автоматы Альберта». Все что мы хотим сказать, это то что уже существующая физика на самом деле позволяет описывать нечто, очень похожее на субъективный опыт. Это весьма неожиданно и интересно само по себе, но главное здесь то, что если такое возможно, то у нас есть все основания считать, что мы решили задачу «чувственных качеств» в принципе. Другими словами описание субъективного мира, а значит и «желто-сладкого» осуществимо в рамках физики.

План остальной части статьи выглядит так: в следующем разделе мы изложим основные результаты Альберта, закончив тем самым наше принципиальное обоснование познаваемости субъективного в рамках существующей уже физики. Фактически, при изложении этих результатов, мы буквально воспроизведем самые главные места оригинальной статьи Альберта, которая, насколько нам известно, пока не переводилась на русский язык.

В четвертом разделе мы покажем, что концепция квантовых автоматов Альберта может привести нас к очень неожиданному (хотя и весьма спорному) решению проблемы как «геделизации». Более строго математическая часть выделена в пятом разделе, который неподготовленный читатель может опустить. В шестом разделе мы подводим итоги и, в частности, именно там приводим аргументы в пользу (i).

3 Автоматы Альберта

В своей работе Альберт рассмотрел автоматы которые работают по правилам квантовой механики и способны измерять и записывать некоторые физические не только собственного окружения, но и кое-какие свойства самих себя. Оказывается, когда такой автомат производит самоописание происходят нечто необычное, нечто непохожее на все что может произойти, при любых обстоятельствах, с классическим автоматом.

Предположим (следуя Альберту), мы конструируем автомат с некоторым механизмом для ввода и вывода информации, с разнообразными устройствами для измерения разнообразных физических наблюдаемых, с внутренней программой, которая включает набор правил для предсказывания поведения некоторых простых физических систем (включая самого себя) по данным начальным условиям и который сам действует по законам квантовой механики. Последнее утверждение означает, что каждое возможное состояние этого автомата является вектором в квантово-механическом гильбертовом пространстве и что каждый вектор этого пространства является возможным состоянием нашего автомата (состояние в котором автомат предсказывает некоторую частную вещь, будет одним вектором из этого пространства, а состояние в котором он предсказывает другую вещь будет ортогональным вектором) и что каждое измеримое свойство этого автомата (*например*, то что он предсказывает) будет соответствовать некоторому эрмитовому оператору в этом пространстве, и что, наконец, каждый такой оператор может быть, в принципе, измерен. Альберт разумно не предпринимает никаких попыток, чтобы решить являются ли обычные машины или компьютеры или мозги такими объектами, а задается вопросом: пусть мы построили такой объект (мы определенно можем вообразить это), как он будет вести себя?

Мы хотим знать какие ответы будут давать эти автоматы в различных обстоятельствах. Для начала, вообразим, что некоторая система (скажем S), находится в состоянии $|\alpha_1\rangle_S$ таком что

$$A|\alpha_1\rangle_S = \alpha_1|\alpha_1\rangle_S, \tag{1}$$

где (для простоты) A — интеграл движения; и что автомат проинструктирован измерить и записать величину A. Когда он заканчивает работу, мы спрашиваем его: «чему равна величина A»? Если это «хороший» автомат, то он конечно должен ответить, что это α_1. Более точно (поскольку наш автомат должен находиться в некотором квантово-механическом состоянии и поэтому задание вопроса такому автомату должно соответствовать измерению некоторой определенной физической наблюдаемой) мы требуем от «хорошего» автомата, чтобы по окончанию измерения A состояние составной системы S плюс автомат было бы

$$|\alpha_1^{(1)}\rangle \equiv |\alpha_1\rangle_A \bigotimes |\alpha_1\rangle_S, \tag{2}$$

где

$$P_A|\alpha_1\rangle_A = \alpha_1|\alpha_1\rangle_A.$$

Здесь наблюдаемая P_A измеряет значение *предсказания сделанного автоматом* для A ($|\alpha_1\rangle_A$, другими словами, это то состояние автомата в котором он предсказывает, что значение A есть α_1).

Если мы хотим качественно описать насколько хорошо автомат сделал свою работу, можно определить наблюдаемую E_A составной системы S+автомат, которая измеряет ошибку в предсказании сделанном автоматом для наблюдаемой A, а именно:

$$E_A = P_A - A.$$

Мы будем называть предсказание автомата для любой наблюдаемой A, точными, если $E_A = 0$ и займемся вопросами: что это за наблюдаемые и при каких условиях их комбинации могут быть точно предсказаны квантово-механическим автоматом. Например, точным является предсказание автомата в (2), т. е.

$$E_A|\alpha_1^{(1)})\rangle = 0.$$

Предположим, что S приготовлено в состоянии

$$|\beta_1\rangle_S \equiv 2^{-1/2}\left(|\alpha_1\rangle_S + |\alpha_2\rangle_S\right),$$

где

$$B|\beta_1\rangle_S = \beta_1|\beta_1\rangle_S,$$

и что мы даем этому автомату ту же инструкцию, что и раньше. Когда он закончит, итоговое состояние будет:

$$|\beta_1^{(1)}\rangle \equiv 2^{-1/2}\left(|\alpha_1^{(1)})\rangle + |\alpha_2^{(1)})\rangle\right) = 2^{-1/2}\left(|\alpha_1\rangle_A \bigotimes |\alpha_1\rangle_S + |\alpha_2\rangle_A \bigotimes |\alpha_2\rangle_S\right). \quad (3)$$

Таким образом $E_A|\beta_1^{(1)}\rangle = 0$, хотя $|\beta_1^{(1)}\rangle$ не является собственным состоянием ни P_A ни A; предсказания сделанные автоматом, несомненно, точны, несмотря на то, что это предсказание само по себе двусмысленно и непонятно. Но в этом нет ничего странного; как никак, говоря что предсказание автомата является точным мы имеем в виду, что если мы бы мы позже измерили A сами и сравнили результат с предсказанием P_A сделанным автоматом, то с определенностью обнаружили бы, что это было правильное предсказание и оно (для некоторой доли квантовых автоматов), вообще говоря и не должно быть хорошо определенным пока мы сделали измерение и не зафиксировали результат.

Что произойдет если мы скажем автомату измерить A тем самым, приведя его в положение, для предсказания A, а потом попросим его также предсказать результат последующего измерения B; или попросим его измерить B и предсказать результат последующего измерения A? Если автомат способен явно предсказать одновременно и A и B, то P_A и P_B должны коммутировать, но

$$[A,B] \neq 0,$$

поэтому

$$[E_A,E_B] = [A - P_A, B - P_B] = [A,B] \neq 0$$

(последнее уравнение следует из $[P_A,P_B]=0$), так что предсказания автомата о не коммутирующих наблюдаемых системы S не могут быть точными одновременно. Это находиться в полном согласии с принципом неопределенности.

Лучшее, что такой автомат может сделать при таких условиях — предсказать *вероятность* определенного результата, который был бы присущ большому ансамблю схожих экспериментов; любого полного измерения системы S, по крайней мере, достаточно чтобы ясно и недвусмысленно предсказать вероятности любых других таких измерений. Если, например, S приготовлена в начальном состоянии $|\beta_1\rangle_S$, а автомат проинструктирован измерить B и предсказать A, то после того, как все такие инструкции будут выполнены состояние окажется таким:

$$|\bar{\alpha}\rangle_A \bigotimes |\beta_1\rangle_S = 2^{-1/2}|\bar{\alpha}\rangle_A \bigotimes \left(|\alpha_1\rangle_S + |\alpha_2\rangle_S\right),$$

где

$$\bar{\alpha} \equiv \frac{1}{2}\left(\alpha_1+\alpha_2\right), \qquad \bar{P}_A|\bar{\alpha}\rangle_A = \bar{\alpha}|\bar{\alpha}\rangle_A,$$

где $\bar{P}_A$ измеряет *среднее значение* предсказанное автоматом для A. Если повторить тот же эксперимент для некоторого большого числа (N) систем, то состояние получившегося ансамбля будет приблизительно собственным для наблюдаемой,

$$\bar{E}_A(N) = N^{-1}\sum_{i=1}^{N}\bar{P}_{A_i} - A_i$$

когда $N\to\infty$ и отвечать нулевому собственному значению. Несложно убедиться, что этот же результат будет верен если B приготовлено не в состоянии $|\beta_1\rangle_S$, а в некоторой произвольной суперпозиции различных $|\beta_i\rangle_S$.

Все это привычная квантовая теория измерений, в которой автомат обладает теми же возможностями измерять и предсказывать, что и наблюдатель[5]. Но давайте представим себе (и тут мы начинаем воображать вещи и ситуации, которые обычно не воображают) что $B^{(1)}$ это наблюдаемая составной системы из S плюс автомат, такая что

$$B^{(1)}|\beta_1^{(1)}\rangle = \beta_1^{(1)}|\beta_1^{(1)}\rangle$$

(где $|\beta_1^{(1)}\rangle$ состояние определенное в (3)) , что S изначально приготовлено в $|\beta_1\rangle_S$ и что мы приказываем автомату измерить A и $B^{(1)}$ (в этом месте, помните, мы просим его измерить нечто *в самом себе*, нечто о его собственных предсказаниях для A[6]) и предсказать результат последующих измерений обеих этих величин. Когда он закончить всю эту работу, сложное состояние будет иметь вид

$$|\beta_1^{(2)}\rangle \equiv |\beta_1^{(1)}\rangle_{B^{(1)}} \bigotimes |\beta_1^{(1)}\rangle = 2^{-1/2}|\beta_1^{(1)}\rangle_{B^{(1)}} \bigotimes \Big(|\alpha_1\rangle_A \bigotimes |\alpha_1\rangle_S + |\alpha_2\rangle_A \bigotimes |\alpha_2\rangle_S\Big),$$

[5] Это, можно сказать, основа многомировой интерпретации квантовой механики, или «относительной» интерпретации описанной Путнамом в (Патнем, 1999).

[6] Нас не должно удивлять, что такое измерение может потенциально изменить память автомата об A. Автомат мы мыслим как квантовый объект, поэтому в его поведении нет

где

$$P_{_{B^{(1)}}}|\beta_1^{(1)}\rangle_{_{B^{(1)}}} = \beta_1^{(1)}|\beta_1^{(1)}\rangle_{_{B^{(1)}}}.$$

Состояние $|\beta_1^{(2)}\rangle$ очень необычно если вспомнить смысл всех символов, потому что

$$E_A|\beta_1^{(2)}\rangle = E_{_{B^{(1)}}}|\beta_1^{(2)}\rangle = 0,$$

несмотря на то, что

$$\left[A, B^{(1)}\right] \neq 0.$$

Если бы мы теперь поинтересовалась у автомата, в самом ли деле он может предсказать и $B^{(1)}$ и A, он должен ответить «да»; если мы попросим его еще раз измерить $B^{(1)}$ и затем A после чего сравнить их с исходным предсказанием он должен определенно подтвердить, что эти предсказания были правильны. Автоматдолжен предсказать обе эти величины, несмотря на то, что они не коммутируют!

Есть еще кое-что необычное. Предположим, у нас есть второй автомат, той же конструкции как и первый. Мы поинтересуемся предсказаниями второго автомата (о какой-нибудь частной наблюдаемой, скажем O) ${}^2P_O$ чтобы отличить от предсказаний первого, и также нас будет интересовать соответствующий оператор ошибки ${}^2E_O$. Получим

$$\left[{}^1E_A, {}^1E_{_{B^{(1)}}}\right]|\beta_1^{(2)}\rangle = 0,$$

но

$$\left[{}^2E_A, {}^2E_{_{B^{(1)}}}\right] \neq 0$$

потому что, коль скоро второй автомат занят, A и $B^{(1)}$ *обе* относятся к внешней системе. Если мы спросим второй автомат, в самом ли деле он находится или *мог когда либо находитьсяв* состоянии, чтобы предсказать и $B^{(1)}$ и A, он должен сказать «нет» (несмотря на то, что он может находиться в позиции позволяющей предсказать $B^{(1)}$ и выдать среднее значение для A, или предсказать A и выдать среднее для $B^{(1)}$, или предсказать и A и ${}^2B^{(1)}$, где ${}^2B^{(1)}$ связано со вторым автоматом также, как $B^{(1)}$ связано с первым).

Так что тут возникает нечто *субъективное*, подобно тому как это было со способностью первого автомата находиться в состоянии позволяющем предсказывать, одновременно, и A и $B^{(1)}$ (или со способностью второго предсказывать и A и ${}^2B^{(1)}$); ибо эта способность зависит не только от *структуры* автомата, но и от того о каком именно автомате идет речь, от их *индивидуальности*. Или мы можем представить это так: есть некоторые комбинации фактов, которые могут в принципе быть предсказаны автоматом *только о самом себе*. Конечно, первый автомат может сообщить второму, что он предсказывает для A и $B^{(1)}$, но это действие (обмен предсказаниями) должно сделать предсказание неверным, потому что случившаяся однажды корреляция между какими-нибудь переменными второго автомата и A приводит к тому, что смешанное состояние больше не может быть собственным для $B^{(1)}$.

ничего необратимого.

Впредь мы будем ссылаться на последовательность измерений типа той, что мы рассмотрели, как на «повторяющуюся» («итерактивную»), с каждым новым измерением такого типа, помимо других вещей связан результат предсказания этого измерения. Несложно понять, что такая повторяющаяся последовательность может быть продолжена до бесконечности. Например, предположим, что S приготовлена вначале в состоянии:

$$|\gamma_1\rangle_S \equiv |\beta_1\rangle_S + |\beta_2\rangle_S = |\alpha_1\rangle_S + |\alpha_2\rangle_S + |\alpha_3\rangle_S + |\alpha_3\rangle_S,$$

где

$$G|\gamma_1\rangle_S = \gamma_1|\gamma_1\rangle_S,$$

а автомат проинструктирован измерять и делать предсказания вначале об A, затем о $B^{(1)}$ и, наконец, о $G^{(2)}$, где

$$G^{(2)}|\gamma_1^{(2)}\rangle = \gamma_1^{(2)}|\gamma_1^{(2)}\rangle,$$

$$|\gamma_1^{(2)}\rangle = |\beta_1^{(2)}\rangle + |\beta_2^{(2)}\rangle.$$

По окончании работы автомата смешанное состояние будет

$$|\gamma_1^{(3)}\rangle \equiv |\gamma_1^{(2)}\rangle_{G^{(2)}} \bigotimes |\gamma_1^{(2)}\rangle;$$

где

$$E_A|\gamma_1^{(3)}\rangle = E_{B^{(1)}}|\gamma_1^{(3)}\rangle = E_{G^{(2)}}|\gamma_1^{(3)}\rangle = 0,$$

хотя

$$\left[A,B^{(1)}\right] \neq 0, \qquad \left[A,G^{(2)}\right] \neq 0, \qquad \left[B^{(1)},G^{(2)}\right] \neq 0.$$

А теперь третья неожиданность. Предположим S приготовлена в некотором произвольном начальном состоянии; мы просим автомат померить A, потом $B^{(1)}$ и затем просим его предсказать среднее значение для C (где C некоторая переменная системы S, такая что $[C,A] \neq 0$). Поразмышляем, каким образом он будет это делать. Несомненно, есть две альтернативы (и они будут в общем случае приводить к *разным* предсказаниям):

1. Автомат может принять во внимание собственное предсказание, α_i, для A и предсказать, что

$$\bar{P}_C = \langle\alpha_i|C|\alpha_i\rangle.$$

2. Он может учесть собственное предсказание,2 $\beta_j^{(1)}$, о $B^{(1)}$ и предсказать

$$\bar{P}_C = \langle\beta_j^{(1)}|C|\beta_j^{(1)}\rangle,$$

и какое бы из них не будет принято, это будет (статистически) точно! Что мы можем с этим поделать? Какое предсказание будет представлять для нас интерес? Нельзя сказать, что мол, нас интересуют «объективные» предсказания, предсказания связанные с тем, что находится «снаружи» наблюдателя, потому что *это* зависит от исходного состояния S, состояния о котором у автомата нет информации. Для примера, предположим автомат находит, что $A = \alpha_1$ и $B^{(1)} = \beta_1^{(1)}$. Возможно, что S была изначально приготовлена в $|\beta_1\rangle_S$ и в этом случае выбор (2) должен быть «объективным» выбором; однако может статься, что S изначально находилась в состоянии, скажем, $|\gamma_1\rangle_S$ и в этом случае, ни (1) ни (2) не будут «объективны». Альберт заключает, что понять такие вещи можно только в рамках многомировой, эвереттовской интерпретации квантовой механики. Мы согласны!

4 Геделизация

В предыдущем разделе мы изложили концепцию самоописывающих автоматов Альберта. Думаем, что вышесказанного достаточно для того, чтобы прийти к совершенно недвусмысленному мнению: в квантовой механике действительно содержатся некоторые вещи, которые трудно назвать иначе чем «субъективным опытом». А это, в свою очередь, означает, что мы можем быть оптимистами и считать физику достаточно могучей, чтобы описать внутренний мир человека, несмотря на то, что сейчас не знаем даже как подступиться к этой задаче.

Если внимательно изучить концепцию Альберта, то поневоле рождается мысль, что «измеряя некоммутирующие наблюдаемые» такие квантовые автоматы должны демонстрировать гораздо более богатый и неожиданный репертуар своего поведения, чем можно было бы ожидать у квантовых автоматов не занимающихся «самоописанием». В данном разделе, мы попытаемся показать, что у этих автоматов нет принципиальных запретов на порождение знаменитых «геделевских суждений», т. е. выполнения невычислимых операций, операций, которые в принципе не способна выполнить ни одна машина Тьюринга или формальная система. Разумеется, все нижесказанное не будет строгим доказательством (такого доказательства у нас, к сожалению нет). Мы просто попытаемся показать, что это не исключено.

В двух своих книгах (Пенроуз, 2003 и Пенроуз, 2005) Пенроуз предпринял попытку показать, что для понимания работы человеческого интеллекта нам необходима новая физика. В качестве ключевого аргумента Пенроуз использует теорему Геделя и способность человека осуществлять процедуру построения геделевских суждений, которую мы в дальнейшем будем называть *геделизацией*. Гедель фактически показал, что математик-человек может выйти за пределы любой формальной системы правил, а стало быть наш разум не является какой-то формальной системой[7]. С другой стороны, мы полагаем, что физика есть некая, очень сложная формальная система. Отсюда Пенроуз заключает, что наша физика неполна, т. к. она не способна описывать «разум» способный геделезировать. По мысли Пенроуза, необходимо включить в физику некоторые *невычислимые операции*. Тогда, существование «геделезирующего интеллекта» будет вполне совместимо с законами природы[8].

[7]В заключительном разделе мы представим свои соображения по этому поводу. Если коротко. по нашему мнению Пенроуз НЕ доказал, что человек — математик действительно способен выходить за эти формальные пределы. Тем не менее, в рамках данной работы, мы примем допущение Пенроуза.

[8]Следует отметить, что Пенроуз не считает познание «разума» главной целью такой перестройки здания физической науки. Он полагает, что невычислимость необходима для по-

В этом разделе мы попытаемся показать, что способность «геделезировать» есть некоторое свойство Альбертовских квантовых автоматов, описанных в предыдущем разделе. Более точно, автоматы Альберта способны к некоторым действиям, которые очень похожи на то, что мы обычно подразумеваем под геделизацией. Тем не менее, не ясно можно ли считать «геделизацию» альбертовских автоматов истиной геделизацией. Именно по этой причине, мы использую кавычки говоря, что альбертовские автоматы «геделезируют». С другой стороны, главная гипотеза этого раздела конечно звучит так: геделизация=«геделизация». Если это так, то нам не нужна новая физика для понимания «разума». Достаточно уже существующей квантовой механики!

Напомним неформальную суть теоремы Геделя. Пусть $p_n(w)$ суть суждение о целом числе w, где n номер этого суждения[9]. Цепочку суждений являющихся доказательством некоторой теоремы (например $p_n(w)$, если это утверждение истинно в нашей формальной системе) обозначим $U(x)$, где x номер доказательства. Саму же же (классическую) формальную систему обозначим W_c.

Фундаментальное геделевское суждение имеет вид

$$p_w(w) =\sim \exists\, x\ [U(x)\ \text{«доказывает»}\ p_w(w)]\,, \tag{4}$$

где $\sim$, $\exists$ — обычные логические кванторы : $\exists$ означает «существует... такое что», $\sim$ означает «неверно, что...».

Гедель показал, что суждение (4) действительно может быть записано в рамках достаточно мощной формальной системы W_c. Очевидно, само суждение $p_w(w)$ должно быть истинным (предполагается, что система W_c является «хорошей», в том смысле что она не порождает доказательств ложных суждений), а значит, *по самому смыслу суждения* (4), оно не может быть доказано (или, иначе, выведено из аксиом) в рамках системы W_c.

Геделевское суждение (4) может быть построены для любой достаточно мощной и непротиворечивой формальной системы, поэтому все такие системы, рассматриваемые как «машины» для порождения новых «теорем», будут ограничены в своих возможностях теоремой Геделя. Что касается нас (т. е.людей), то в этом случае теорема Геделя почему то не срабатывает[10]: мы умеем геделезировать, а стало быть, по какой то таинственной причине, превосходим любую формальную систему правил!

Именно отсюда Пенроуз (а до него Лукас) делает вывод, что работу сознания нельзя полностью смоделировать с помощью любой системы формальных правил, а значит, следует предположить, что работа сознания описывается некими невычислимыми функциями или неалгоритмическими процедурами. С другой стороны, наличие таких процедур в нашем мозгу подразумевает наличие ана-

---

нимания проблемы квантовых коллапсов и квантовой гравитации. Дополнительной же наградой, по мысли Пенроуза, будет долгожданное понимание того, что такое разум и как он функционирует!

[9]Важной частью работы Геделя было построение такой системы нумераций всех суждений генерируемых данной формальной системой.

[10]Подчеркнем — это точка зрения Пенроуза. Мы ее принимаем в данной работе за истинную, хотя истинна ли она — не факт. Лично я думаю, что она НЕВЕРНА.

логичных явлений и вне мозга. Чтобы не впасть в мистику, остается предположить, что современная физика неполна. В ней не хватает «невычислимых» ингредиентов, которые, по мысли Пенроуза, также необходимы для решения загадки квантового измерения.

Теперь ясно, что если нам удастся объяснить геделезацию в рамках уже известной физики, то тем самым мы лишим Пенроуза главного аргумента в пользу «неполноты» современной физики. По мнению автора этих строк, геделезация действительно может быть понята, как особое свойство квантовых автоматов Альберта.

Для того чтобы выразить эту идею в максимально понятном виде, мы рассмотрим парадоксальное суждение хорошо известное под названием парадокс Эпименида или парадокс лжеца. Некто утверждает: «то что я сейчас утверждаю — ложно». Очевидно, что это утверждение ложно тогда и только тогда, когда оно истинно и наоборот. Как поступать с такими высказываниями? То что от них нельзя просто отмахнуться было убедительно продемонстрировано Тарским, который сумел построить формальное (т. е. записанное в рамках некоторой формальной системы) суждение, утверждающее о себе, что оно ложно. Отличие суждения Тарского от парадокса лжеца, сформулированного на, скажем, русском языке, в том, что первое имеет *два уровня значения*. Первый совпадает с «русским» вариантом, тогда как второй представляет собой суждение о целых числах истинное тогда и только тогда, когда оно ложно! Другими словами, парадоксальность суждения Тарского в том, что оно *выглядит* полностью бессмысленным, но при этом является не просто суждением о себе (в этом случае парадоксальность привычна), а некоторым суждением о числах!

Итак, повторим вопрос: как следует понимать такие суждения? Интересно, что Пенроуз фактически полностью проигнорировал парадоксы такого типа. В своих книгах Пенроуз утверждает, что математические понятия существуют в некоем «платоновском» абсолютном виде и наш мозг каким то образом умеет «соединяться» с этим миром абсолютных математических истин. Мы знаем «истину» не потому, что наш мозг манипулирует символами, будучи некоей формальной системой, а потому, что умеем «наблюдать» эту истину занимаясь математикой.

Автор вынужден признаться, что он совершенно не понимает этой точки зрения. Дело даже не в том, что непонятно что такое «мир Платона». Пенроуз утвердает, что физика неполна и, надо полагать, как раз эту, недостающую часть физики он называет «платоновским миром абсолютных истин», истин, к которым мы таинственным образом «подключаемся» решая математические задачи. Однако, даже если согласиться с этим, все равно равно остаются вопросы. В частности, не ясно как Пенроуз объясняет парадокс Эпименида? Где здесь «истина»? Для защиты своей точки зрения Пенроуз аппелирует к геделевским суждениям, недоступным для формальных систем, но доступных нам (откуда и делает вывод, что «мы» не формальные системы). Но суждение Тарского, хотя и похоже на геделевское (в том смысле, что оно является самоописывающим) не является таковым. Геделевское суждение — истинно, а суждение Тарского не ложно и не истинно. Значит, мы тоже беспомощны перед парадоксами типа

Эпименида-Тарского. Ну и как это объяснить с точки зрения «платоновско-пенроузовского мира абсолютных математических идей», мира к которому мы можем «подключаться, чтобы узреть истину»?

Пенроуз не отвечает на этот вопрос, но его оппонент Дуглас Хофштадтер пытается высказать гипотезу, что причина появления парадоксов типа парадокса лжеца, носит *физический* характер. Согласно его в высшей степени оригинальной идее, нам не удается разрешить парадокс Эпименида потому, что при этом в мозгу должно произойти нечто не разрешаемое законами физики! Вот точная цитата: «Значит, необходимо искать нейронный субстрат парадокса Эпименида — низший уровень противоречащих друг другу физических событий, т. е. событий которые не могут произойти одновременно» (Хофштадтер, 2001, с. 546).

Хофштадтер не конкретезирует физический механизм, который не может сработать, когда мы пытаемся понять парадокс Эпименида, да это и невозможно на современном уровне знаний о функционировании мозга и связи физических процессов в мозге с формулированием суждений этим самым мозгом. По этой причине рассмотрим гипотетическую, но более определенную ситуацию: пусть мы построили «робота», предназначеного для вывода математических (в том числе логических) теорем и для проверки на истинность-ложность суждений, которые мы можем вводить в него извне[11]. Будем предпологать, что наш «робот» является квантово-механическим, в том смысле, что для описания его функционирования следует использовать квантовое, а не классическое описание. По этой причине, будем называть этого «робота» квантовым математиком роботом или КМР (в принципе, наш КМР можно представлять, как достаточно мощный квантовый компьютер).

Ключевым свойством является то, что работа КМР может быть описана на двух уровнях. Уровень I это уровень логического описания. Работая на этом уровне, мы заключаем, что КМР может генерировать определенные суждения, но не способен распознать, как истинное некоторое геделевское суждение и т. д. Уровень II — это уровень физического описания. Поскольку, по условию, КМР описывается на языке квантовой механики, то уровень II это уровень гамильтонианов и волновых функций. Более точно, для полного описания на уровне II используется две процедуры: унитарная эволюция заданная набором линейных операторов U (логических гейтов) и процедура квантовомеханического измерения, которую мы будем задавать набором проекторов. В принципе, оба описания должны давать один и тот же результат, хотя переход от уровня I к II и наоборот может оказаться невероятно сложной задачей. Нас однако интересует лишь принципиальная сторона дела.

Если КМР получает задание вывести семейство каких то теорем, то его работа будет представлять сложную унитарную эволюцию (т. е. вращение в гильбертовом пространстве состояний КМР), но рано или поздно КМР должен «напечатать» результат, что естественно предполагает проведение КМР кванто-

[11] Таких «роботов», как воплощение с ложных формальных систем с успехом использует Пенроуз для обоснования своих идей о «невычислимости» человеческого разума в (Пенроуз, 2003)

вомеханического измерения одной или нескольких наблюдаемых. Конечно, все ответы КМР будут носить вероятностный характер, но это не принципиально, поскольку для получения надежного ответа достаточно попросить КМР решить эту же задачу достаточно большое количество раз. Аналогично выглядит работа КМР, пытающегося определить истинность какого то суждения, которое мы «попросили» его проверить.

Что будет если мы попросим КМР вынести суждение об истинности суждения типа Эпеменида-Тарского? Очевидно, наш робот не справится с заданием так же, как и мы и это очевидно если рассматривать уровень I. А как выглядит дело с уровня II? Очевидно, должна найтись некоторая физическая причина по которой КМР не сможет завершить свою работу. Что же это за причина? Мне кажется вполне естественно предположить (кстати, полностью в духе Хофштадтера), что парадоксы типа эпименидовского возникают в том случае, если для осознания их истинности КМР приходится измерить несколько некоммутирующих наблюдаемых *одновременно*.

А теперь мы делаем еще более сильное
ПРЕДПОЛОЖЕНИЕ: *не только парадоксы типа Тарского, но и совершенно истинные, геделевские суждения не могут быть проверены КМР на истинность, потому что для этого требуется измерить несколько некоммутирующих наблюдаемых одновременно. Другими словами, в основе недоступности геделевских суждений лежат не логические, а физические причины.*

Если принять это предположение, то очевидно, что мы получаем лазейку для квантовых автоматов Альберта, описанных в предыдущем разделе. Действительно, если такой автомат сталкивается с задачей измерить несколько некоммутирующих наблюдаемых (скажем A и B), то, будучи не в состоянии решить эту задачу прямо он может подойти к ней окольным путем: перейти в состояние $|\beta_1^{(2)}\rangle$ (см. (3)); составить внутреннее описание этих наблюдаемых - A и $B^{(1)}$ — наблюдаемых которые автомат в состоянии (3) уже способен одновременно точно измерить, несмотря на то, что A и $B^{(1)}$ не коммутируют. Разумеется многие сочтут это жульничеством, поскольку автомат по прежнему не способен измерить A и B. Но, с другой стороны, можно возразить, что после проведения измерения A, автомат уже никак не может иметь дело с B. Теперь в его памяти записано не B, а $B^{(1)}$ и именно $B^{(1)}$ он должен «считать» реальным!

Но если это так, то вполне реально использовать подобный механизм для описания того, как квантовый автомат Альберта конструирует геделевское суждение.

Рассмотрим КМР предназначенного для «доказательства теорем» в рамках некоторой формальной системы W_c и проверки на истинность суждений, которые мы можем загружать в него через «интерфейс». Пусть S некоторая квантовая система с гильбертовым пространством состояний H_S. Введем наблюдаемую $P_k(w)$ такую, что для любого вектора $|w;k\rangle_S \in H_S$ выполняется уравнение

$$P_k(w)|w;k\rangle_S = k|w;k\rangle_S. \tag{5}$$

Т. е. собственные значения $P_k(w)$ будут номерами пропозиционных функций $p_k(w)$. В том же гильбертовом пространстве определим «логические гейты» U_j.

Полный список всех таких гейтов образует *квантовую формальную систему* W_q. Мы будем говорить, что *квантовая формальная система* может доказать (или вывести) суждение $p_k(w)$ (о числах w) если существует конечная последовательность гейтов

$$U = U(w,k) \equiv \prod_j U_j,$$

такая что

$$|\psi\rangle_S \to |w;k\rangle_S = U|\psi\rangle_S, \tag{6}$$

для некоторого начального состояния $|\psi\rangle_S \in H_S$. В этом случае, мы можем стартуя с $|\psi\rangle_S$ найти $|w;k\rangle_S$; затем измерить среднее значение $P_k(w)$ и получит номер суждения k:

$$k = {}_S\langle\psi|U^+P_k(w)U|\psi\rangle_S,$$

знание которого позволяет восстановить суждение $p_k(w)$. Детали такого процесса могут быть весьма сложны, но нас интересует только принципиальная сторона дела. Нет сомнений, что такая процедура может быть в принципе реализована, за исключением некоторых случаев, которые как раз и представляют для нас главный интерес.

Итак, при каких условиях вышеописанная процедура может быть осуществлена? Чтобы ответить на этот вопрос вообразим, что W_q является универсальной квантовой формальной системой, способной протестировать вектор $|w;k\rangle_S$ на предмет является ли он решением (5). Подставляя $|w;k\rangle_S$ на место $|\psi\rangle_S$ в (6) мы, разумеется должны получить тот же самый вектор[12].

$$U|w;k\rangle_S = u(w;k)|w;k\rangle_S,$$

т. е.

$$[P_k(w),U] = 0. \tag{7}$$

Уравнение (7) фактически выражает обычный принцип неопределенности. В самом деле, пусть квантовая формальная система W_q проинструктирована измерить (или вычислить, что означает тоже самое в нашем случае) величину $p_k(w)$. Наблюдаемой в W_q является U поэтому и $P_k(w)$ и U должны быть одновременно измеримы. Рассуждая таким образом, мы опять приходим к уравнению (7).

Ну а как обстоит дело с геделевским суждением (4)? Это пропозиционная функция от числа w с номером суждения w не может быть получено (доказано) в рамках систем W_c and W_q. Обозначим $P_w \equiv P_w(w)$. Очевидно, что должно выполняться неравенство

$$[P_w,U] \neq 0. \tag{8}$$

В противном случае, мы можем использовать U чтобы найти $|w\rangle_S \equiv |w;w\rangle_S$ и, в конце концов, вычислить геделево суждение $p_w(w)$. Но это невозможно, в силу теоремы Геделя!

[12] Мы можем получить, вообще говоря, другой вектор $|w;m\rangle_S$ с $m \neq k$. Если это действительно произошло, то следует переопределить U так чтобы исключить такую возможность. Можно показать, что это всегда осуществимо.

Итак мы дали грубое описания КМР, который не может создать геделевское суждение потому, что для этого ему надо измерить пару некоммутирующих наблюдаемых. Теперь ясно, как именно можно заставить автомат геделезировать: надо или заставить его работать по схеме Альберта или подключить к автомату Альберта. Для полноты картины мы описываем детали этой процедуры в приложении.

В заключение этого параграфа, остановимся на одном вопросе, способном вызвать недоразумение. Читатель не знакомый с теорий формальных систем или машин Тьюринга вправе усомниться, что суждение P_w действительно может быть построено. Разве теорема Геделя не утверждает, что P_w не достижимо для формальной системы? А раз так, то утверждение о том, что все дело в некоммутируемости (8) попросту притянуто, а вовсе не следует из каких то принципов, как утверждает автор этих строк.

Это не так. Для прояснения вопроса, полезно посмотреть на теорему Геделя с точки зрения работ Тьюринга. Хорошо известно, что Тьюринг доказал существование алгоритмически неразрешимых проблем. Например, нельзя построить универсальную процедуру предсказывающую остановится ли данная машина Тьюринга, вычисляющая некоторую функцию от данного натурального числа. Доказательство этого факта (как и доказательство теоремы Геделя) использует знаменитую процедуру диагонализации Кантора. В результате оказывается, что можно построить никогда не останавливающийся алгоритм α, такой, что универсальная машина Тьюринга не способна предсказать, что этот алгоритм никогда не остановится!

Алгоритм α родственен геделевским суждениям P_w. Важно понимать, что этот алгоритм действительно может быть легко построен за конечное число шагов. Необычность его, повторим, в том, что универсальная машина Тьюринга не способна предсказать, что будет с α дальше: остановится ли эта процедура или продлится вечно. Точно также, можно построить суждение P_w. Необычность P_w в том, что формальная система не способна установить его истинность, ибо установить истинность формальная система может лишь одним способом — вывести это суждение из исходных аксиом, что невозможно по самому смыслу суждения P_w.

5 Квантовые автоматы Алберта и самоописание

Рассмотрим автомат A[13] который проинструктирован (или запрограммирован) измерить и записать (запомнить) значения P_w и U. Имеют место соотношения

$$U|\psi\rangle_S = u|\psi\rangle_S, \qquad |\psi\rangle_S = \sum_w c_w|w\rangle_S, \qquad P_w|w\rangle_S = w|w\rangle_S,$$

где P_w, которые сравниваются с геделевскими суждениями p_w. Можно добавить к $|\psi\rangle_S$ дополнительные («нормальные», негеделевские) слагаемые $|w;k\rangle_S$ с $k \neq w$, но это не играет роли в дальнейших рассуждениях. Пусть система S

[13]Этот раздел можно опустить неспециалистам по квантовой механике

находится в состоянии $|w\rangle_S$. После того, как измерение P_w закончено, состояние комбинированной системы $S+A$ будет иметь вид [3]

$$|w^{(1)}\rangle = |w\rangle_S \bigotimes |w\rangle_{P_w}, \qquad (9)$$

где $|w\rangle_{P_w} \in H_A$, H_A — гильбертово пространство состояний автомата A и $|w\rangle_{P_w}$ — собственный вектор «альбертовской наблюдаемой» $G(P_w)$,

$$G(P_w)|w\rangle_{P_w} = \tilde{w}|w\rangle_{P_w}.$$

Это значение наблюдаемой предсказанное автоматом для P_w. Предсказание является точным, если

$$E(P_w)|w^{(1)}\rangle = (\tilde{w} - w)\,|w^{(1)}\rangle = 0,$$

где

$$E(P_w) \equiv G(P_w) - P_w.$$

Ниже мы будем рассматривать только точные предсказания ($\tilde{w} = w$).

Альбертовские наблюдаемые коммутируют с остальными наблюдаемыми[14],

$$[G(P_w),P_w] = [G(P_w),U] = [G(P_w),G(U)] = [G(U),P_w] = [G(U),U] = 0,$$

поэтому

$$[E(P_w),E(U)] = [P_w,U] \neq 0, \qquad (10)$$

что согласуется с принципом неопределенности и, в то же время, с теоремой Геделя. Действительно, геделевское суждение p_w не может быть доказано в рамках формальной системы W_q поэтому совместное предсказание автоматом величин U и P_w не может быть точным. Таким образом мы имеем ситуацию полностью аналогичную той, что была описана в предыдущем разделе. Если это так, то для чего мы ввели второй квантовый автомат A? Потому что, как показал Альберт, автомат A может преодолеть, в определенном смысле, ограничение (10).

Пусть исходное состояние системы S будет описано собственным вектором U: $|\psi\rangle_S$. Это означает, что выполняются процедуры формальной системы W_q, а значит геделевское суждение P_w недостижимо, в рамках этой системы. Несмотря на это A запрограммирован измерить P_w. По окончанию этого измерения состояние комбинированной системы $S+A$ примет вид

$$|\psi^{(1)}\rangle = \sum_w c_w |w^{(1)}\rangle, \qquad (11)$$

просто в силу линейности уравнений квантовой механики. $|w^{(1)}\rangle$ из (11) определен посредством (9). Теперь введем новую наблюдаемую $U^{(1)}$, такую что

$$U^{(1)}|\psi^{(1)}\rangle = u^{(1)}|\psi^{(1)}\rangle.$$

[14]Подчеркнем, что $G(P_w)$ не является функцией от наблюдаемой P_w также как и $G(U)$ не

Что представляет собой $U^{(1)}$? Мы можем назвать её «логическим гейтом» новой формальной системы $W_q^{(1)}$, частично записанной в памяти автомата A. Ясно, что новая формальная система $W_q^{(1)}$ не мощнее исходной (W_q). Но теперь, если автомат измерит наблюдаемую $U^{(1)}$, мы получим

$$|\psi^{(1)}\rangle \to |\psi^{(2)}\rangle = |\psi^{(1)}\rangle_{U^{(1)}} \bigotimes |\psi^{(1)}\rangle,$$

(см. (9)), где

$$G(U^{(1)})|\psi^{(1)}\rangle_{U^{(1)}} = u^{(1)}|\psi^{(1)}\rangle_{U^{(1)}},$$

поэтому

$$E(P_w)|\psi^{(2)}\rangle = E(U^{(1)})|\psi^{(2)}\rangle = 0,$$

хотя P_w и $U^{(1)}$ не коммутируют (см. (10))! Таким образом, автомат в состоянии $|\psi^{(2)}\rangle$ может точно предсказывать и P_w, и $U^{(1)}$. В некотором смысле мы можем назвать это «гёделизацией». Это не означает, разумеется, что автомат может вывести p_w в рамках формальной системы $W_q^{(1)}$. Правильная интерпретация этого результата звучит так: автомат A может «понимать» как гёделевское суждение p_w, так и формальную систему $W_q^{(1)}$. Классический автомат этого не может. Чтобы заставить классический автомат «понять» p_w и $W_c^{(1)}$, нам необходимо загрузить в него более мощную «формальную систему» $\tilde{W}_c$. С квантовым автоматом Альберта ситуация качественно другая.

И последнее. «Гёделизация», описанная выше, это «личный факт жизни автомата A». Наблюдаемая $U^{(1)}$ и наблюдаемая P_w являются внешними для другого автомата $\tilde{A}$, который поэтому не способен «понять» одновременно $U^{(1)}$ и P_w. Чтобы «гёделизировать» на таком уровне, ему надо попасть *собственное состояние* $|\tilde{\psi}^{(2)}\rangle$.

является функцией от U.

6 Заключение

Из этого следует два вывода. Первый свидетельствует «против Пенроуза», второй — за него.

Против. Нам не нужна новая физика, чтобы понять Гёделизацию. Квантовый автомат Альберта может «понять» одновременно формальную систему и гёделевское суждение, которое невозможно получить внутри этой системы.

За. Если наша «гёделизация» это на самом деле настоящая гёделизация, и если гёделизация это неалгоритмическая процедура, то нам приходится признать по крайней мере то, что квантовая механика содержит «нечто неалгоритмическое».

Заметим, что в рамки указанной парадигмы можно поместить новый взгляд на соотношение классичких и квантовых свойств (см. раздел 1, предположение (i)). Вообразим, фантастически сложный автомат Альберта, осуществляющий невероятно длинные цепочки вложенных измерений и формирующий наборы внутренне — внешних наблюдаемых, которые не коммутируют, но могут быть одновременно точно измеренны. Допустим, что этот автомат пользуется алгоритмами типа «машинное обучение» и ищет корреляции между этими наблюдаемыми. По самому смыслу понятно, что ничего подобного соотношению неопределенности в рамках наборов таких многократно измеренных наблюдаемых в этих корреляциях не появится. Другими словами, автомат Альберта придет к *классической* картине. Т. е. можно посмотреть на классическую динамику просто, как на *внутреннее описание* квантового мира, а не как проявление боровской дополнительности!

Боюсь, прочитав эту работу, читатель может сделать неверные выводы, поэтому следует резюмировать статью четким указанием на то, что мы НЕ утверждаем.

1. Мы не утверждаем, что для описания «субъективного» нужна квантовая механика. Более того, мы считаем это НЕВЕРНЫМ! Мозг слишком горячее место для того, чтобы в его функционировании были заметны квантовые эффекты.

2. Мы не утверждаем, что для описания работы «сознания» необходимо вводить невычислимые функции, как это делает Пенроуз. Вполне возможно, что кажущаяся «геделизация» человека — математика является либо следствием того, что математик совершает *ошибки* (точка зрения Тьюринга) либо того, что работа сознания по определению включает в себя самоописывающие суждения и значит приводит к невычислимости, в смысле непредсказуемости (точка зрения Ллойда).

Мы утверждаем только, что феномен «субъективности», служащий камнем преткновения для философов и когнитивистов, может быть вполне понят в рамках известной физики!

Список литературы

Вигнер Е. Пределы науки // Этюды о симметрии. — М., 1971. — С. 170—181.
Патнем Х. Философия сознания. — М. : Дом интеллектуальной книги, 1999.
Пенроуз Р. Новый ум короля: О компьютерах, мышлении и законах физики. — М. : Эдиториал УРСС, 2003.
Пенроуз Р. Тени разума: в поисках науки о сознании. — Москва-Ижевск, 2005.
Хофштадтер Д. Гёдель, Эшер, Бах: эта бесконечная гирлянда. — Самара : Издательский Дом «Бахрах-М», 2001.
Шредингер Э. Разум и материя. — Ижевск, 2000.
Albert D. On Quantum-Mechanical Automata // Physics Letters. 98A. — 1983. — No. 5, 6. — P. 249–252.

Об авторе

Артём Валерьянович ***Юров*** — доктор физико-математических наук, профессор, директор Института физико-математических наук и информационных технологий Балтийского федерального университета имени И. Канта, AIUrov@kantiana.ru.

Physics of Subjective

Artyem V. Yurov[i]

[i]Immanuel Kant Baltic Federal University
Kaliningrad

Abstract: The problem of "the subjective" in physics is discussed. It is shown that some class of quantum automata (Albert automata) is able to demonstrate something like "subjective experience". It is argued that such automata are capable of generating Gödel's propositions. In addition, it is demonstrated that taking into account such unusual properties of Albert automata is able to open a new perspective on the problem of the correlation of quantum and classical properties.

Keywords: philosophy of physics, philosophy of consciousness, R. Penrose, gödelerization, Albert quantum automata.

References

Albert, D. (1983). "On Quantum-Mechanical Automata". *Physics Letters. 98A*, no. 5, 6, pp. 249–252.

Khofshtadter, D (2001). *Gedel', Esher, Bakh: eta beskonechnaya girlyanda.* Samara: Izdatel'skii Dom "Bakhrakh-M".

Patnem, Kh. (1999). *Filosofiya soznaniya.* M.: Dom intellektual'noi knigi.

Penrouz, R. (2003). *Novyi um korolya: O komp'yuterakh, myshlenii i zakonakh fiziki.* M.: Editorial URSS.

— (2005). *Teni razuma: v poiskakh nauki o soznanii.* Moskva-Izhevsk.

Shredinger, E. (2000). *Razum i materiya.* Izhevsk.

Vigner, E. (1971). "Predely nauki". In: *Etyudy o simmetrii.* M., pp. 170–181.

About author

Prof. Dr. *Artyom V.* ***Yurov***, Director of Institute of Physics, Mathematics and Information Technology, Immanuel Kant Baltic Federal University, AIUrov@kantiana.ru.

RATIO.ru 18(1): 3–25