

Дмитрий Казеннов

Проблема «китайской комнаты» в контексте дискуссии о функционализме

В своей знаменитой работе «Сознание, мозг и программы»¹ Джон Серл продемонстрировал (на примере не менее знаменитой «игры в имитацию»² Алана Тьюринга) свои доводы в пользу соображений о том, что «компьютер с подходящей программой» не способен к пониманию и, следовательно, не «разумен» в той же степени, что и человек. Отличие между программой, оперирующей знаками на основании формальных правил, и мозгом, по логике Серла заключается в качестве, которое он называет «интенциональностью». При этом Серл утверждает, что некоторые машины, в числе которых он называет и человеческий мозг, способны к пониманию, то есть он выступает с натуралистических позиций³:

«Может ли машина мыслить?»

Ответ очевидным образом положителен. Мы как раз и есть такие машины. ... Если вы в состоянии точно продублировать причины, то вы в состоянии продублировать и следствия. И на самом деле, быть может, возможно продуцировать сознание, интенциональность и все такое прочее, используя какие-то другие химические принципы, чем те, что реализованы в людях.

Серл утверждает, что компьютер, выполняющий программу, или человек, действующий вслепую (говорящий на неизвестном ему языке) по соответствующим такой программе правилам, может успешно выполнить задание теста Тьюринга, то есть участвовать в диалоге и удовлетворительно для интервьюера отвечать на задаваемые вопросы, не понимая содержания диалога, а лишь формально соединяя символы (создавая синтаксически правильные фразы). Это должно

1 Серл Д. Сознание, мозг и программы // Аналитическая философия: становление и развитие. – М., 1998.

2 Turing A. Computing Machinery and Intelligence. // Mind, 59, 236, 1950.

3 Серл Д. Цит. соч., с. 393-394.

демонстрировать, во-первых, бесполезность теста Тьюринга для определения наличия интеллекта, а, во-вторых, невозможность реализации «сильного искусственного интеллекта». Аргумент Серла породил множество ответных реплик⁴. В целях настоящего исследования необходимо, во-первых, привести в качестве первого и второго тезиса краткие частные возражения аргументации Серла и, во-вторых, в качестве третьего тезиса коснуться содержания термина «понимание» у Серла в контексте дискуссии Патнема и Кальке о функционализме.

1. Первый тезис данной работы заключается в том, что «китайская комната» в описанных Серлем условиях мысленного эксперимента не может выполнить некоторые конкретные задачи, выполнения которых может требовать успешное прохождение теста («игры в имитацию») Тьюринга.

Первый тип таких задач – это ответ на оценочные вопросы, вопросы вида «как вы оцениваете элемент рассказа X». Не будет ничего удивительного, если подобные вопросы будут заданы в ходе теста Тьюринга в целях разоблачения компьютерного оппонента-имитатора. Программист, составляющий правила для соединения китайских символов «комнатой», должен понимать этот вопрос на китайском языке как обычный человеческий носитель китайского языка, отыскать в рассказе элемент X и иметь ценностное суждение об этом элементе рассказа, то есть сложившуюся реакцию на стимул, обозначаемый элементом X, или чувственно подобный ему (здесь и далее ценностные высказывания понимаются как иллюстрации эмоциональных реакций на соответствующие стимулы в духе эмотивизма). Если программист заложит вопрос «как вы оцениваете элемент рассказа X» в рукопись-3 (вопросы) и свой ответ «я считаю, что X – это хорошо; мне нравится X» на этот вопрос в рукопись-2 (ответы), которыми будет пользоваться человек, находящийся в «китайской комнате», то «китайская комната» будет в таком случае служить программисту, оценивающему элемент X китайского рассказа, лишь обыкновенным устройством вывода, вроде принтера. Сама по себе она задачу оценки не выполнит, и на вопросы об оценке элементов рассказа, не имеющиеся в инструкциях, составленных программистом, не ответит (если только не будут алгоритмизированы рефлексy, то есть то, что Серл бы назвал «интенциональным»).

4 Preston J. and Bishop M. (eds.) Views into the Chinese Room: New Essays on Searle and Artificial Intelligence. – N. Y., 2002.

Второй тип задач – это вопросы, требующие в качестве ответов наличия нескольких из наиболее общих эвристических способностей, составляющих содержание интеллектуальности, которые перечисляет Марвин Мински: *поиск, распознавание, обучение, планирование и обобщение*⁵. Задача *поиска*, например, успешно алгоритмизируется, но алгоритмы поисковых машин, на самом деле, не могут пересказать и, казалось бы, «не понимают» искомые тексты. Можно предположить, что правила, которыми пользуется человек в «китайской комнате», будут достаточно подробными и объемными, чтобы отыскать элемент X в рукописи-1, но что произойдет, если поисковый запрос будет дан по синониму находящегося в тексте термина? Если искомого знака в рукописи не будет (хотя будет означающий в точности то же самое термин), «китайская комната» сообщит об отсутствии термина, что приведет к разоблачению имитации. Если же инструкции «китайской комнаты» будут содержать и список синонимов, и некий алгоритм для определения одного нужного синонима (список равноценных синтаксически вариантов ответа на простой вопрос в духе современных поисковых машин будет провалом имитации) в зависимости от контекста (например, семантическую сеть), то не будет ли такая программа обладать некоторой семантикой (что означало бы наличие некоторого уровня «понимания» и подорвало бы аргумент Серла о том, что формальный алгоритм оперирует лишь с синтаксисом)? Другим хорошим примером может быть способность к *обучению*: всегда можно задать китайской комнате серию вопросов, требующих для нахождения ответов способности к обучению (использовать какую-нибудь словесную ролевую игру с головоломками). Если существует алгоритм, обеспечивающий такую способность, то «комната» сможет обучиться китайскому языку так же, как англичанин (то есть, «комната» научится соотносить символы-иероглифы если не с предметами, в отсутствии телесной чувственности, то с реакциями собеседника-человека: «боль – это то, в ответ на упоминание чего люди-собеседники реагируют так-то и так-то»; разумеется, если оборудовать такую «комнату» сенсорами, то она сможет соотносить символы с предметами). Если же такой алгоритм не существует, «комната» провалит тест.

Неспособность «китайской комнаты» Серла выполнить задания «игры в имитацию» Тьюринга не означает возможность или невозможность алгоритмизации разума или его составляющих

5 Minsky M. L. Steps toward artificial intelligence // Proceedings of the Institute of Radio Engineers, 49, 1961.

(эмоциональных реакций, эвристических способностей), а *изначальную некорректность* самого мысленного эксперимента Серла.

2. Второй тезис данной работы заключается в следующем: Джон Серл не раскрывает содержание понятия «понимание», а для целей своей статьи он считает достаточным два ясных случая, в которых можно буквально говорить о наличии и отсутствии понимания. Касаясь человеческой способности к пониманию, он лишь замечает⁶, что

Для способности людей понимать рассказы характерно, что люди способны отвечать на вопросы о данном рассказе даже в тех случаях, когда даваемая ими информация не выражена в рассказе явным образом.

В качестве примеров очевидно ложного атрибутирования способности понимания предметам Серл называет автоматическую дверь с фотоэлементом и термостат⁷:

Очень интересно, на каком основании мы делаем такие атрибуции, и это основание связано с тем, что мы распространяем на артефакты нашу собственную интенциональность; наши инструменты суть продолжения наших целей, и поэтому мы находим естественным приписывать им метафорическим образом интенциональность... Тот смысл, в каком автоматическая дверь «понимает инструкции» посредством своего фотоэлемента, – это вовсе не тот смысл, в каком я понимаю английский язык.

В целях данной статьи следует рассмотреть дискуссию о функционализме. В своей работе «Ментальная жизнь некоторых машин»⁸ Хилари Патнэм формулирует свой подход в рамках классической проблемы философии разума – проблемы соотношения тела и разума (mind-body problem). Патнэм дистанцирует свою позицию как от картезианского дуализма, так и от физикализма. По логике Патнэма, функциональные качества, выделяемые на определенном уровне абстрагирования, независимы от структуры реализующей их физической

6 Серл Д. Цит. соч., с. 377.

7 Там же, с. 381

8 Патнэм Х. Ментальная жизнь некоторых машин // Патнэм Х. Философия сознания. – М., 1999.

системы. Ментальность и поведенческий репертуар человека или машины здесь определяется как набор функциональных качеств в противовес различной физико-химической структуре этих машины и человека⁹:

...я постараюсь обосновать, что боль – это не состояние мозга, если под этим понимать физико-химическое состояние мозга (или даже состояние всей нервной системы), а состояние совершенно иного типа. Я выдвигаю гипотезу, что боль или состояние боли – это функциональное состояние всего организма.

Дискуссия, которой касается Патнэм, также затрагивает так называемую проблему *квалиа*. Развивая положения функционализма в статье «Роботы: машины или искусственная жизнь?», Патнэм настаивает на несводимости функциональных и структурных качеств¹⁰:

...сказать, что человек и робот имеют одинаковую «психологию»... значит сказать что поведение этих двух видов наиболее просто и ясно анализируется на психологическом уровне (в абстракции от деталей внутренней физической структуры)...

И также¹¹:

В самом деле, могло бы существовать сообщество роботов, которые не имели бы одинаковое физическое устройство, но все имели бы одинаковую психологию; и такие роботы могли бы говорить «У меня есть чувство красного», «у тебя есть чувство красного», «у него есть чувство красного», даже если бы три указанных робота не «осознавали бы физически» «чувство красного» одинаковым образом.

9 Там же, с. 59-60.

10 Putnam H. Robots: Machines or Artificially Created Life? // The Journal of Philosophy, 61, 21, 1964, p. 677.

11 Op cit., p. 675.

В работе «Что не так с функционализмом Фодора и Патнэма?»¹² Кальке возражает Патнэму, утверждая, что нельзя провести принципиальную демаркационную линию между структурой и функцией¹³:

... чтобы определить функциональную организацию физико-химической системы, вы должны определить не только предел системы, но и уровень абстракции для описания ее поведения. Для любой данной системы, положение, на котором уровень абстракции и предел наиболее полезным образом зафиксированы, ни в коем случае не определен уникально, — он будет различаться, в зависимости от некоторого числа прагматических соображений.

В самом деле, определение функциональных качеств произвольно, зависит от выбора пределов рассмотрения системы и уровня абстрагирования. При достаточном абстрагировании можно считать функционально изоморфными кошку и мышеловку (кошка и мышеловка, как замечает Кальке, выполняют одинаковую функцию ловли мышей), а задав жесткий предел тех структурных характеристик, которые останутся за пределами «черного ящика» (оставив в «черном ящике», то есть вне рассмотрения, только индивидуальные особенности физиологии) и низкий уровень абстрагирования, то есть крайне детальное рассмотрение поведения, можно определить две разные человеческие единицы как функционально сопоставимые (два разных человека по-разному реагируют на боль, по-разному ведут себя в определенных условиях и, следовательно, функционально не изоморфны).

Таким образом, подразделение на функциональные и структурные состояния произвольно. Ментальные состояния нельзя полностью абстрагировать от состояний структурных. Нельзя исключить возможность редукции. Ментальные состояния, вероятно с развитием нейробиологии, могут быть эмпирически сопоставлены определенным структурным характеристикам (поведение — содержанию памяти, а память — молекулярному состоянию клеток коры головного мозга). Функционализм Патнэма и Фодора уводит дискуссию в сторону от решения основного вопроса, поскольку не отвечает напрямую на вопрос о природе когнитивных способностей («интеллектуальности»).

12 Kalke W. What is wrong with Fodor and Putnam's Functionalism // Nous, 3, 1, 1969.

13 Ibid., p. 88.

Представление о структуре и представление о функции есть продукт чувственного наблюдения, исследования объекта. Мы говорим о том, что объекты имеют одинаковые функции, если эти предметы и их функционирование вызывают у нас, наблюдателей, одинаковую чувственную реакцию. То же справедливо в отношении отождествления структуры. Если же два объекта, наблюдаемые нами, подобны до степени смешения по своим функциональным свойствам, но мы знаем (в результате более подробного исследования), что структурно они различны, то мы не отождествляем эти объекты (мы можем, например, говорить о *конвергенции*). Осведомленность о структурной разнице уже достаточна для того, чтобы объект А вызывал у нас чувственную реакцию, отличную от чувственной реакции, вызываемой объектом В. В конце концов, *распознавание, узнавание и отождествление* – когнитивные поведенческие способности человека, позволяющие образовывать понятия, символы для множества индивидуальных фактов, вызывающих одинаковые чувственные реакции. Как точно подметил Мах, «каждое слово служит тогда для обозначения одного класса объектов (вещей или процессов) с определенной реакцией»¹⁴.

Однако, нельзя не признать, что некоторые *достаточно* схожие (где функциональной разницей, обусловленной структурным различием, можно пренебречь) функции действительно могут быть реализованы различными структурами. Можно представить себе ситуацию, в которой машина будет до некоторой степени *имитировать* в функциональном отношении человека. Степень достоверности имитации могла бы быть различна, но какой-то уровень имитации был бы достаточен для выполнения какой-либо сложной деятельности, требующей хотя бы таких когнитивных способностей, как распознавание. Другой уровень функциональной имитации был бы достаточен для эффективной социализации. В такой ситуации должен существовать структурный предел (boundary), за которым обнаруживалась бы разница («скрытая чужеродность») в структурных и функциональных качествах. Можно до определенной степени (до степени отождествления) уподобить разные структуры. Но две разные структуры можно и просто перепутать; при дальнейшем исследовании отождествление может показаться наблюдателям ошибочным.

Разумеется, чем более близок предел имитации к разрешающему пределу физико-химической структуры (в качестве допущения – к

14 Мах Э. Познание и заблуждение. Очерки по психологии исследования. – М., 2003, с. 145.

молекулярному уровню), тем ближе машина к клону. Однако уровень подобию, устраивающий человеческих наблюдателей, может быть достигнут ранее. Вероятно, довольно искусная механическая машина, управляемая микросхемами, будучи способна к имитации поведения, могла бы ввести человеческого наблюдателя в заблуждение на некоторое время.

Следуя данной логике, можно продемонстрировать отсутствие недвусмысленной демаркационной линии между человеческим «пониманием» рассказа и «пониманием», присущим автоматической двери с фотоэлементом, на существовании которой (демаркационной линии) настаивает Серл. В самом деле, «понимание» – это не структура мозга (или фотоэлемента), а функция работы структуры мозга, абстракция, которую мы определяем из прагматических соображений на произвольной основе. Развивая гипотезу редукции мышления ко множеству реакций на ряд стимулов, как настоящих, так и запомненных (вроде памяти боли от соприкосновения с горячим металлом), можно заметить, что на элементарном уровне работа мозга может быть грубо уподоблена *в функциональном отношении* работе фотоэлементов.

Если же изменить уровень предела черного ящика, оставив вне рассмотрения индивидуальные особенности физиологии и уровень абстракции функции «понимания», различив конкретные индивидуальные реакции на рассказ, можно будет утверждать что даже два человека не способны понять рассказ одинаково: Джон Серл рассказ понимает, а вот какой-нибудь Джон Смит – вовсе нет (потому как иначе на содержание рассказа реагирует, да и конституцией своей отличается).

Подобная логика лишает случай понимания человеком рассказа буквальности и ясности.

3. Третий и главный тезис данной работы заключается в том, что Серл допускает существование машин, которые могут быть способны к пониманию и, следовательно, обладанию интеллектом в полном смысле слова. Человек является такой машиной. Но остается неясным, какие именно машины способны обладать интеллектом, какие не способны, и в чем заключается структурная разница между ними. Почему можно моделировать, алгоритмизировать или формализовать устройство и работу одних машин, но не других? Серл лишь говорит об «интенциональности», но это не вносит ясности в вопрос о том, машины какого устройства такой «интенциональностью» могут обладать.

В своей работе «Что такое интенциональные состояния?»¹⁵ Серл пишет, что интенциональные состояния – это ментальные состояния, имеющие направленность. Интенциональные состояния относятся к объектам и состояниям дел (обстоятельствам) так же, как особый тип предложений, которые Серл называет иллокативными. Иллокативные предложения полностью соответствуют интенциональным состояниям и, что важнее, позволяют судить об этих ментальных состояниях: эти состояния «представляют объекты и состояния дел в точности в том же смысле, в котором акты речи представляют объекты и состояния дел»¹⁶. Серл указывает, что не интенциональность лингвистична, а сама лингвистика интенциональна.

Серл обращает внимание на то, что некоторые интенциональные состояния и соответствующие им иллокативные предложения направлены на несуществующий объект или ложное положение дел. Проблема разъяснения смысла таких предложений аналогична проблеме разъяснения смысла фиктивного (вымышленного) дискурса.

Серл вводит целую таксономию иллокативных речевых актов, то есть актов, выражающих интенции в противовес пропозиционным речевым актам, описывающим действительность (пропозиции, предложения, имеющие истинностное значение в терминологии Фреге и Витгенштейна). За английскими словами «утверждать» (state) и «верить» (believe) следует предлог «что» (that), эти сказуемые употребляются с одинаковыми инфинитивами глаголов, хотя очевидно, что утверждающее предложение и предложение уверенности не одинаковы ни в наличии истинностного значения, ни в плане интерсубъективности (вера не имеет истинностного значения и субъективна). Другие иллокативные речевые акты, такие как «предложение-приказ» или «предложение-желание» употребляются с теми инфинитивами глаголов, с которыми не употребляются «убежденности» и «утверждения». Исследуя синтаксис, направленность речевых актов, Серл выделяет утвердительный (assertive) тип предложений (утверждения, описания, выражения убежденности), которые имеют направленность «предложение-мир», а также директивный (приказы, желания) и комиссивный (обещания, клятвы) – типы предложений, которые имеют направленность «мир-предложение». Имеются также предложения, не имеющие такой направленности, такие как извинения и поздравления¹⁷.

15 Searle J. What Is an Intentional State? // Mind, 88, 349, 1979.

16 Ibid., p. 75

17 Ibid., pp. 75-77.

В целях данной статьи важно следующее: среди иллокативных актов, которые перечисляет Серл, есть выражения эмоций, реакции на сложившееся положение дел. Соответствующие им интенциональные состояния есть эмоциональные состояния (радость, грусть). Такие конкретные акты и такие соответствующие им состояния являются реакциями на стимулы, условными рефлексам. Целью данной работы является демонстрация того, что и другие иллокативные речевые акты и интенциональные состояния являются поведенческими реакциями, которые могут быть описаны в терминах стимула и рефлекса. В таком случае, во-первых, не будет необходимости говорить (как это делает Серл) об условиях удовлетворения (или истинности) иллокативных речевых актов: такие акты лишь поведенческий ответ, рефлекс.

Что еще более важно, ряд высказываний из тех, которые Серл называет иллокативными, являются ценностными высказываниями. И можно утверждать, что имеет смысл анализировать такие высказывания при помощи метаэтической теории эмотивизма, основания которой сформулировали Альфред Айер и Чарльз Стивенсон. Альфред Айер¹⁸ отделяет собственно этические суждения от анализа содержания этических терминов, анализа «морального опыта» и его причин и этических «увещаний». Предметом интереса Айера был вопрос о том, «могут ли утверждения этического значения быть переведены в утверждения эмпирического факта». Рассматривая этические суждения, Айер отказывается от субъективистского и утилитаристского подхода, как способов анализа существующих этических предложений. Субъективистский подход заключается в том, что «верное» – есть то, что поддерживается большинством. Но в утверждении о том, что «большинство иногда поддерживает неверные действия», нет логического противоречия. Следовательно, определение «верного» не связано с коллективным соглашением или индивидуальным одобрением. Утилитаристский подход также является неудовлетворительным в целях определения того, что считается «верным», так как нет логического противоречия в утверждении о том, что «не всегда действия, приносящие наибольшее удовлетворение в данной ситуации, являются верными». Айер пишет о том, что редукция этических утверждений к дескриптивному не совместима с конвенциями современного обыденного языка. Рассмотренные им утилитаризм и субъективизм отвергаются, таким образом, не как источники каких-либо конкретных этических

18 Ayer A. Language, Truth and Logic. – L.: Gollancz, 1946.

утверждений, а как способы анализа существующих этических понятий, как основания метаэтики. Также, Айер указывает, что следует различать нормативные этические символы и дескриптивные этические символы, где последние являются лишь описаниями предпочтений отдельно взятого общества, то есть являются обычными социологическими предложениями. Если же брать в качестве примера традиции обоснования этических суждений в некоей «интуиции» или метафизике, то легко обнаруживается, что из таких неясных оснований можно обосновывать произвольные утверждения. Для абсолютистских или интуитивистских метаэтических соображений нет релевантной эмпирической проверки. Главная же мысль Айера заключается в том, что этический символ в пропозиции не добавляет ничего к фактическому значению предложения: «ты поступил неправильно, украв деньги» = «ты украл деньги», где оценка есть лишь выражение эмоций, синтаксически тождественная восклицательному знаку. «Если в предложении вообще не содержится утверждения, очевидно, что нет смысла спрашивать, является ли сказанное истинным или ложным»¹⁹. Наконец, утверждения об эмоциях отличаются от выражений эмоций (то есть, этических суждений) тем, что выражение эмоции не обязательно сопровождается утверждением. Можно только добавить, что утверждение об эмоциональном переживании («в данный момент я скучаю», «в данный момент я напуган») не является суггестией и этически нейтрально, что более явно артикулировано Стивенсоном.

Чарльз Стивенсон в своей работе «Эмоциональное значение этических терминов»²⁰ подверг критике классические этические «теории интереса» Юма и Гоббса, так как последние определяли эмоциональные термины (такие как «благо»), используя чисто дескриптивные термины («то, что одобряется обществом»). Согласно Юму назвать нечто «благом» означает признать, что большинство одобряет это нечто, а согласно Гоббсу «благо» есть предмет желания. «Согласно определению Гоббса, человек может подтвердить окончательно свои этические суждения, показав, что он делает интроспективную ошибку о своих желаниях. Согласно определению Юма, кто-то может подтвердить (грубо говоря) свои этические суждения, приняв участие в голосовании. Подобное использование эмпирического метода, в любом случае, кажется далеким от того, что мы принимаем как доказательство, и подвергает сомнению полную релевантность определений, которые

19 Ibid., p. 112.

20 Stevenson C. The Emotive Meaning of Ethical Terms // Mind, 46, 181, 1937.

подразумевают его»²¹. Стивенсон считает необходимым исключить любую форму голосования по вопросам этики и отвергает теории интереса в их классической форме, поскольку даже из истинного утверждения о коллективном согласии в интересах никак не следует ни релевантное определение понятия «благо», ни индивидуальное принятие коллективного соглашения. Из эмпирических качеств самого предмета так же никаким образом не следует то, является ли такой предмет «благом». «Вместо этого, мы, должно быть, используем какое-то понимание «блага», которое не определяется релевантно в терминах чего-либо научно познаваемого»²². Стивенсон утверждает, что есть три ключевых требования, которым отвечает релевантное понимание «блага», но которые не выполняются классическими теориями интереса: во-первых, благо должно быть предметом рациональных разногласий, во-вторых, оно должно быть «притягательным» («magnetic»), и, в-третьих, оно не должно быть определяемым с помощью научного метода (эмпирически). Хотя *традиционные* теории интереса этим требованиям не отвечают, «это не подразумевает, что «благо» должно быть объяснено в терминах Платоновской Идеи, Категорического императива или быть уникальным неанализируемым качеством»²³.

Автор далее указывает, что в классических теориях интереса этические утверждения предоставляют информацию об индивидуальных интересах при данных условиях, на практике такие высказывания не описывают факты, а являются средством влияния на других людей, средством убеждения. В самом деле, нет смысла в констатации наличия у себя таких-то и таких-то интересов, подобные утверждения не станут предметом разногласия (столь характерного для этических вопросов). От цели употребления этических высказываний Стивенсон переходит к вопросу значения таких высказываний и выделяет два типа высказываний: дескриптивные, выражающие уверенность в факте, и динамические, не выражающие уверенности в каком-либо факте, но выражающие эмоции, направленные на создание настроения, содержащие суггестию. Стивенсон вводит понятие *эмоционального значения* слова, которое обозначает «тенденцию данного слова, возникающую в течение истории его употребления, порождать (или порождаться) аффективные реакции в людях»²⁴. В конце концов, автор

21 Ibid., p. 17.

22 Ibid., p. 18.

23 Ibid.

24 Ibid., p. 23.

модифицирует теорию интереса таким образом, чтобы она отвечала трем заданным ранее критериям. В таком случае, дискуссия об этических утверждениях является не разногласием о вере (в факт), а разногласием об интересах, и потому не может быть разрешена рационально. Проблема «притягательности» решается просто, поскольку рассуждения Стивенсона включают и заинтересованность говорящего. Наконец, третье и самое важное ограничение, введенное Стивенсоном, касается роли научного метода, эмпирических данных и согласия о вере (в факт) в разногласии об интересах. Эмпирический метод является релевантным этике лишь постольку, поскольку касается предмета или фактора, ставшего объектом нашего интереса. «Является ли эмпирический метод достаточным для достижения этического согласия? Очевидно, нет. Поскольку эмпирическое знание разрешает разногласие в интересах лишь до той степени, в которой это разногласие укоренено в разногласии о вере»²⁵. Этические утверждения не являются дескриптивными и не могут раскрыть некий особый тип априорной истины, который бы руководил *самими интересами* людей, «...спросить, что есть благо, означает, просить оказать влияние»²⁶. Таким образом, Чарльз Стивенсон вносит лингвистические коррективы в классические теории интереса и более подробно артикулирует базовые принципы метаэтической теории эмотивизма.

В целях же данной работы важно следующее: множество иллокативных высказываний, о которых пишет Серл, и множество ценностных высказываний, о которых пишет Стивенсон, пересекаются. Так, оценочные высказывания содержат суггестию, то есть команду, но команда – это содержание комиссивных иллокативных высказываний. Команда подразумевает некоторое желательное, но не имеющее места в настоящий момент положение дел, но суггестивное предложение вообще не рассматривается как дескриптивное. Таким образом, можно утверждать о тождестве интенции и эмоциональной потребности оказать влияние. Отметим, что Стивенсон вообще не подразделяет дескриптивные предложения на высказывания о вымысле и о достоверно известных фактах, это не представляется ему релевантным. Верным представляется следующий вывод: высказывания, которые Серл называет иллокативными, можно подразделить на несколько отдельных высказываний, одни из которых представляют собой оценку, эмоциональный ответ, суггестию, другие – представление о

25 Ibid., p. 29.

26 Ibid., p. 30.

последствиях, о будущем. О природе последних, как и о природе вымысла (вымышленный дискурс), Серл говорит не так много. Автор данной статьи предполагает, что вымышленные предложения можно назвать *мнемоническими*, то есть предложениями о содержании памяти, где «представление о будущем» является экстраполяцией прошлого опыта, а вымысел – нарушением памяти, соединением элементов памяти в таком порядке, в котором они не встречались в опыте.

Иллокативный речевой акт, по утверждению Серла, выражает, в частности, страх или радость, то есть эмоциональные реакции на стимул (и таким образом, «интенциональное состояние» является эмоциональной реакцией на стимул), где в качестве стимула выступает предмет интенции: утвердительная (подобная реакции предложения-о-факте!) реакция по отношению к предмету предложения-убежденности, отрицательная эмоциональная реакция по отношению к предмету высказывания-извинения, положительная эмоциональная реакция по отношению к предмету высказывания-поздравления. Но сам предмет таких высказываний, если он не является эмпирическим, является мнемоническим, то есть существующим в форме собирательной памяти. Иллокативная ложь (ложь в обычном смысле слова) есть продукт нескольких реакций (в числе которых – положительная реакция на предполагаемое поведение субъекта-адресата лжи на основании заблуждения), то есть сложный когнитивный акт: положительная реакция на воображаемый/предполагаемый результат поведения человека-адресата лжи, отрицательная реакция на предполагаемое поведение адресата лжи, если тот не будет введен в заблуждение, положительная реакция на правдоподобную и эффективную ложь, страх перед разоблачением. Хотя в представлении Серла иллокативные предложения имеющие направленность «высказывание-мир» статичны и не имеют причин во внешнем мире (и это большая логическая слабость рассуждений Серла), мы не можем не отметить, что у «иллокативных предложений» и, следовательно, у соответствующих им ментальных состояний, которые Серл называет интенциональными, есть причины-стимулы, как у рефлексов. Есть и условия удовлетворения, тождественные удовлетворению поведенческих потребностей. Следовательно, «интенциональные состояния» являются поведенческими потребностями, продуктами сочетания рефлексов. Полезность данного вывода заключается в том, что в отличие от интенциональных состояний, эмоциональные состояния являются нейрогуморальными (связанными с гормональным уровнем) и могут быть интерсубъективно определены.

В целом, для рассуждений Серла об интенциональных состояниях справедлива классическая бихевиористская критика Джоном Уотсоном²⁷ интроспективной психологии: такие рассуждения нельзя проверить опытным путем, они спекулятивны. Если же допустить интроспективные рассуждения о ментальном в качестве гипотезы, то человеческие «внимание», «желание», «волю» можно в духе рассуждений Эрнста Маха²⁸ редуцировать до сложного сочетания множества реакций на чувственные раздражители, как действующие на органы чувств в настоящий момент, так и запомненные, отразившиеся в содержании памяти. Причем реакции на запомненные раздражители и на раздражители настоящего момента равносильны, а память о чувственно подобных стимулах связана ассоциативно. Таким образом, директивный речевой акт, например, можно представить как поведенческий жест, имеющий цель. Понятие же «интенциональное состояние» представляется избыточным, не несущим дополнительной объяснительной функции, попросту говоря, подпадающим под действие бритвы Оккама.

Заключение

Три перечисленных тезиса, как представляется автору, лишают аргумент «китайской комнаты» Джона Серла эвристической ценности. Дело не в том, можно или нельзя алгоритмизировать работу головного мозга человека, а в том, что сам мысленный эксперимент Серла не выдерживает критики при подробном анализе. Но поставленная Серлем проблема интенциональных высказываний представляет интерес. Такие высказывания включают слова и высказывания, имеющие эмоциональные значения, а также высказывания о вымышленных предметах, которые мы можем назвать мнемоническими.

27 Уотсон Дж. Поведение как предмет психологии (бихевиоризм и необихевиоризм) // Гальперин П. Я., Ждан А. Н. (ред.) Хрестоматия по истории психологии. – М., 1980.

28 Мах Э. Цит. соч.