

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В.ЛОМОНОСОВА
ФИЛОСОФСКИЙ ФАКУЛЬТЕТ

На правах рукописи

Ковалев Максим Александрович

**Роль символических методов
в задачах общего искусственного интеллекта**

5.7.6. Философия науки и техники

ДИССЕРТАЦИЯ

на соискание ученой степени

кандидата философских наук

Научный руководитель:

к.филос.н., доцент

Чусов Анатолий Витальевич

Москва – 2024

Оглавление

Введение	4
<i>Актуальность темы исследования.....</i>	<i>4</i>
<i>Степень разработанности темы исследования.....</i>	<i>5</i>
<i>Основные понятия и термины.....</i>	<i>18</i>
<i>Цели и задачи исследования.....</i>	<i>28</i>
<i>Объект и предмет исследования.....</i>	<i>32</i>
<i>Методология исследования.....</i>	<i>32</i>
<i>Положения, выносимые на защиту.....</i>	<i>33</i>
<i>Научная новизна исследования.....</i>	<i>33</i>
<i>Теоретическая и практическая значимость результатов исследования... </i>	<i>35</i>
<i>Степень достоверности и апробация результатов исследования</i>	<i>36</i>
<i>Структура диссертации</i>	<i>37</i>
Глава 1 Исторический анализ процессов возникновения и эволюции понимания роли символьных методов в задачах проектирования ИИ-систем.....	38
§ «Дофизический» период	39
§ Первые реализации на физическом субстрате. «Физический».....	51
§ «Компьютерный период»	54
§ Символьные методы и исследовательские программы ИИ	56
§ Проблемы современного дискурса	73
Глава 2 Анализ онтологического статуса ИИ	80
§ Сущность техники, пределы и граничные условия функционирования технических объектов.....	81
§ Место техники в социальном.....	103
§ Критика основных технологических подходов	112
Символьные методы	112
Коннекционизм	113
§ Критика возможности выхода ИИ за границы техники	120
§ Обоснование необходимости гибридного подхода как следствия вывода о «подручности» ИИ	138
§ Определение граничных условий существования ИИ-систем.....	147

Глава 3 Основные аспекты применимости символьных методов при проектировании ИИ-систем.....	149
§ <i>Эпистемологический и гносеологический аспекты</i>	153
§ <i>Онтологический аспект</i>	207
§ <i>Информационный аспект</i>	213
§ <i>Коммуникативный аспект</i>	219
Заключение.....	225
Приложение 1 Правила генезиса новых понятий, правил и знаков.	230
Список литературы	233

Введение

Актуальность темы исследования

В настоящее время существует практически необозримое количество попыток инженерной реализации систем искусственного интеллекта (ИИ-систем). Задачи общего ИИ в целом можно охарактеризовать как моделирование процессов интеллектуальной деятельности человека. Диссертационное исследование посвящено исследованию роли символических методов как технологического подхода к задачам общего ИИ.

Философскую проблематику ИИ-систем можно рассматривать на разных уровнях общности. Мы ограничиваем исследование проблемами философии науки и техники. В частности, не теряет актуальности проблема онтологического статуса ИИ-систем. Многие современные философы рассматривают ИИ-системы, абсолютизируя их самостоятельность, в отрыве от реальных условий их производства и воспроизводства. Они, по сути, линейно экстраполируют некоторые современные тенденции развития техногенной среды. В то же время многие философы техники, такие как П.К. Энгельмайер, Освальд Шпенглер, Карл Ясперс, Мартин Хайдеггер, Хосе Ортега-и-Гассет, Фридрих Дессауэр, Бруно Латур, Жильбер Симондон, Бернар Стиглер и т.д., рассматривают общие вопросы существования технических устройств в неразрывной связи с их включённостью в человеческое общество, т. е. как техники, произведённой людьми и получающей существование в составе искусственной природы. Такое положение дел демонстрирует большую актуальность прояснения действительных условий и предпосылок развития технических систем и техногенной среды.

Существенная и специфическая особенность ИИ-систем заключается в том, что они непосредственно основаны на получении и переработке информации. При этом явно или неявно проводится аналогия с человеком в

том аспекте, что последний обычно имеет знания, получает знания и использует знания. Здесь присутствует проблемный гносеологический и эпистемологический и технико-технологический узел. В гносеологическом аспекте классической является проблема возникновения знаний у субъекта. В эпистемологическом аспекте классической является проблема возникновения у субъекта системы знаний или же проблема возникновения системного знания. В интересующем нас случае технико-технологической проблематики реализации идеальных конструкций проблема типично выражается как проблема обучения и самообучения ИИ-системы. Актуальность этой проблематики также вполне очевидна (библиографии работ, посвящённых обучению и самообучению ИИ-систем, насчитывают многие тысячи позиций).

Инженерные подходы к реализации ИИ-систем основаны на тех или иных общих теоретических предположениях о строении и структуре интеллекта. При этом исторически основные модели искусственного получения результатов преобразования знаний были основаны на аналогии с математическими вычислениями как с символьными преобразованиями. С развитием техники эта аналогия всё упрочивалась практически до 60-х годов XX века, когда стали всё более широко разрабатываться коннекционистские модели. В настоящее время они почти полностью превалируют среди исследований в области ИИ. Однако такое смещение исследовательского интереса обусловлено не теоретическими причинами. На него повлияли как неудачи прямых реализаций символьных методов, так и появление более дешёвых и производительных вычислительных мощностей. Тем не менее, остаётся под вопросом применимость символьных методов в задачах общего ИИ.

Степень разработанности темы исследования

Тема нашего исследования непосредственно сочетает проблематику общего ИИ с проблематикой технологических подходов к реализации ИИ-систем. Но философский ракурс рассмотрения расширяет эту область, в

частности, посредством обращения к её действительным условиям и предпосылкам. Среди них в первую очередь следует указать на человеческое общество и человеческую деятельность (в том числе и мыслительную), без чего в известной нам реальности не возникает и вопроса о возникновении ИИ-систем как таковых. При этом естественно наличны разные философские уровни рассмотрения проблемы реализации ИИ-систем: онтологический, гносеологический, эпистемологический и инженерно-технический.

На наш взгляд, основная проблематика определения статуса ИИ как искусственного объекта заключается не в общем определении того, «что есть ИИ?», и не в сформулированной в общем виде практической постановке задачи: «можно ли построить систему общего ИИ?». Гораздо более значим и продуктивен поиск ответа на вопрос:

Как можно построить систему общего ИИ, при условии её функционирования в человеческом обществе?

Поставленный таким образом вопрос подчеркивает необходимость философского осмысления системы взаимодействий искусственных объектов с миром и человеком. Процессы практической реализации и дальнейшего самостоятельного функционирования системы общего ИИ не могут основываться исключительно на (основанных на прошлом опыте) идеализированных фрагментах мировых ситуаций, как на источниках новых знаний ИИ-системы. Инженерная реализация процессов обучения и самообучения, как процессов производства знания, нового для ИИ-системы, является необходимым условием, обеспечивающим достижение требуемого уровня автономности функционирования систем общего ИИ.

Ещё одна характеристика современного положения дел в области ИИ заключается в фактическом исключении философии из современного научного и технического дискурса, что само по себе не ново. Линия на элиминацию философии из инженерной и естественнонаучной проблематики

исторически связана с позитивистскими попытками исключения метафизики. Ее другую сторону демонстрирует история ИИ, начиная с 60-х годов XX века. Об этом свидетельствуют воспоминания Хьюберта Дрейфуса о царившей тогда в Массачусетском технологическом институте атмосфере, о дискуссиях и проблемах того периода и о последовавшей за этим «зиме ИИ». Он указывает на упрощенное, «нефилософское» понимание проблемы ИИ, заключавшееся в том, что в своих работах, «фактически не осознавая этого», исследователи в области ИИ опирались на идеи логицизма и бихевиоризма [Dreyfus, 2007].

Но и Мартин Хайдеггер ещё в 1964 году в работе «Конец философии и задачи мышления» диагностирует разложение философии на отдельные науки. При этом кибернетика охарактеризована им как основа «техничности» современного научного дискурса. По нашему мнению, тут обнаруживается смещение общего вектора научных исследований в области ИИ от (в целом логически ориентированной) естественнонаучной позитивистской парадигмы в гораздо более узкую инженерную область информационных технологий. Последующее доминирование компьютерной метафоры способствовало дальнейшему исключению философии из основного дискурса разработчиков ИИ-систем. Одной из главных причин такого положения дел является кажущаяся очевидная простота бихевиористского подхода в моделировании психических процессов с точки зрения ее технической реализации, не требующая своего философского осмысления.

Сегодня, более полувека спустя, ситуация в области ИИ находится в аналогичном положении. Даже беглый анализ списков публикаций и докладов основных мировых конференций в области ИИ говорит о недооценке роли философии в области проектирования ИИ-систем.

Но как включить философию в основной дискурс в области ИИ? Поиск ответа на этот вопрос является сегодня одной из наиболее актуальных задач философии науки и техники. На наш взгляд, без возвращения философии в этот дискурс невозможно понять всю глубину и сложность таких проблем, как: определение возможной степени автономности ИИ-систем; определение места

и значимости символических, коннекционистских и др. методов в области ИИ; установление теоретических и практических пределов возможностей самообучения в задачах общего ИИ.

Обсуждение роли символических методов в задачах общего искусственного интеллекта невозможно без понимания как сложившегося положения дел, так существующих ограничений на условия и перспективы решения задач общего ИИ. Прежде всего, речь идет о возникновении представлений о безграничности возможностей ИИ. Свидетельством этому является целый ряд утверждений о неизбежности достижения «сингулярности» и возникновения «сверхразума» в работах ряда авторов, таких как Ник Бостром и Рэй Курцвейл. Эта уверенность в безграничности возможностей ИИ является универсализацией надежд на получение нужных результатов. В этом смысле она похожа на проекты создания универсального языка науки в начале XX века. Одним из важных результатов распространения и укоренения таких надежд стало убеждение в возможности реализации творческих способностей в рамках какого-либо одного технологического подхода, прежде всего - одних только коннекционистских методов.

Результатом доминирования со стороны коннекционистских подходов стал уход далеко на второй план таких проблем как определение роли тех или иных технологических подходов, перспективы их интеграции; определение границ и поиск пределов автономности ИИ, а также ряда других «предельных» проблем.

Поиск философского ответа на эти вопросы, на наш взгляд, не потерял своей актуальности, и может быть построен прежде всего не столько на критике самих коннекционистских подходов как таковых, сколько на осмыслении роли и места разработанных сегодня технологических подходов.

Как указывает В.С. Степин, любая деятельность всегда регулируется ценностями и целями. Ценность отвечает на вопрос: «для чего нужна та или иная деятельность?» Цель – «что должно быть получено в деятельности?». Таким образом – цель понимается им как идеальный образ продукта [Степин,

2001, с. 38]. Описывая структурную характеристику элементарного акта действия, он выделяет субъектную и предметную (объектную) структуру деятельности. Именно субъектная структура задает ценности и цели. Средства и действия могут быть отнесены одновременно и к объектной, и к субъектной структурам. Двойственность такого представления определяет необходимость рассматривать ИИ-системы как с точки зрения «искусственных органов человеческой деятельности» [Степин, 2000, с.37], так и в качестве «естественных объектов» [там же], взаимодействующих с другими объектами.

Анализируя цели и ценность ИИ-систем с первой точки зрения, как некоего искусственного органа человека, с необходимостью встает вопрос об существовании граничных условий, задающих степень автономности ИИ-систем. На наш взгляд, даже при автономных действиях ИИ-системы её ценность определяется полезностью результатов ее функционирования для человека, понятностью с точки зрения коммуникации и обмена информацией, предсказуемости ее действий.

С другой стороны, ИИ-система, понятая как некоторый объект, самостоятельно действующий среди других объектов, должна быть осмыслена с точки зрения процессов обучения и самообучения, без чего невозможно говорить о сколь-нибудь самостоятельном существовании ИИ-системы в качестве автономного субъекта, способного к самостоятельному принятию решений. Исследование этой проблемы непосредственно связано с исследованием принципиальной возможности обнаружения со стороны ИИ-системы как существующей «для-себя». Говоря о существовании «для себя», мы имеем в виду озабоченное присутствие в мире, имеющее структуру присутствия, решающее задачу познания «именно себя – как познание мира» [Хайдеггер, 2015, с. 59]. Данный вопрос можно сформулировать так:

Может ли изделие (технический объект) обнаружить себя существующим «для-себя», в качестве «всегда-моего»

присутствия? Возможно ли такое существование вне связи с человеком и культурой?

Говоря о связи ИИ-системы с человеком и культурой, мы имеем в виду как человечески смысловую, культурную обусловленность первоначального, иницирующего момента обучения, так и объективное (или по меньшей мере интерсубъективное) существование целей и ценностей, задающих последующие траектории процесса самообучения в ходе функционирования ИИ-системы.

На наш взгляд, ответы на оба вопроса на сегодняшний день являются отрицательными. Существующий сегодня уровень понимания техники, также как и уровень философской рефлексии относительно понимания не только природы сознания, процессов обучения, но и места технических объектов в мире, не дают оснований для возможности выхода ИИ-систем за пределы «подручности», обретения ими самостоятельности в культурном и познавательном аспекте. Напротив, в настоящее время скорее следует говорить о пластичности человека, о его добровольном, но далеко не всегда осознанном, подчинении законам функционирования технических объектов, им же самим и созданным.

Что касается значимости символьных методов для области ИИ – то они являются «первыми» методами, которые легли в основу разработок ИИ. К настоящему времени возникло большое количество различных инженерных решений, реализующих те или иные подходы, основанные на символьных методах. Примерами тому служат как ранние попытки, такие как General Problem Solver Ньюэлла и Саймона, так и современные исследования, такие как ДСМ-метод В.К. Финна. Кроме того, примеры таких систем можно встретить в работах «Искусственный интеллект...» [Люгера, 2003], «Чего не могут вычислительные машины...» [Дрейфус, 2010], «Верховный алгоритм...» [Домингос, 2016], «Искусственный интеллект...» [Финн, 2011] и целом ряде других работ. Среди конкретных технологических подходов и

построенных на их основе решений, можно выделить такие методы как: семантические сети [Люгер, 2003, с. 212], деревья решений [Домингос, 2106], метод «классная доска» [Люгер, 2003, с. 212], разного рода производственные системы, ДСМ-метод [Финн, 2011] и т.д..

Особняком стоят немногочисленные попытки построения гибридных символично-коннекционистских систем [Домингос, 2016; Гридин, 2017].

По нашему мнению, основная трудность подхода к практической реализации систем, моделирующих интеллектуальную деятельность, заключена не в выборе рационального решения из множества доступных альтернатив, а в реализации процессов обучения и самообучения. Можно констатировать, что в настоящее время практически все работы по обучению ИИ-систем ограничиваются исключительно инженерной проблематикой. Примером этого является предлагаемая Жан-Луи Лорьером следующая иерархия уровней обучения [Лорьер, 1991, с. 516-517]:

1 уровень. Программируемое обучение, когда «специальный фиксированный код управляет тем, что сделано». Ему соответствуют все существующие сегодня промышленные роботы.

2 уровень. Запоминание наизусть. Все ситуации запоминаются вместе с действиями, которые нужно предпринять в каждой из них.

3 уровень. Расширенный уровень 2. В системе хранится лишь классификация ситуаций, полученная на основании анализа достаточно большого числа случаев. При этом запоминаются только компоненты, соответствующие вполне определенному действию.

4 уровень. Обучение с помощью учителя. Такая система способна к обобщению.

5 уровень. Обучение без учителя. Система способна индуктивным образом порождать новые ситуации, выдвигать гипотезы и новые концепции¹.

¹ В целях исключения искажения исходного смысла предлагаемого автором (кем именно) мы считаем целесообразным привести данную им классификацию дословно не взирая на объем соответствующего фрагмента. Необходимая атрибуция, отсылающая к исходному тексту предпослана ему.

Схема Лорьера является классическим примером того, что понимается под обучением в сфере решения прикладных инженерных задач. Из недостатков такого понимания можно указать на отсутствие научной рефлексии и на отсутствие связанного с этим методологического осмысления различий, на что обращали внимание В.В. Краевский, И.Я. Лернер и ряд других исследователей. Речь о различиях между опытным освоением предметной области и овладением опытом, учитывая существование разных типов опыта [Серикив, 106]:

- «опыт готовых знаний», где важен объяснительно-иллюстративный процесс,
- «опыт способов деятельности» с его репродуктивной (тренинговой) направленностью,
- «опыт творческий», имеющий проблемную направленность.

Не менее важной характеристикой процесса обучения, согласно В.В. Краевскому, является его «со-бытийность», притом что в этом процессе обязательно наличие двух оппозиций: «ученик-учитель», «ученик-учебный материал». Возвращаясь к концепции уровней обучения у Лорьера, следует заметить, что уровни 1 и 2 не относятся к собственно обучению, так как не подразумевают двусторонней коммуникации между обучаемым и учителем.

Следует различать обучение и самостоятельное опытное освоение мира, что также в значительной степени или не замечается или игнорируется разработчиками в сфере машинного обучения.

Символьные методы, реализующие обучение на основе принципов индукции, разные исследователи реализуют и называют по-разному. Так, Дж. Люгер называет целый ряд таких методов: эвристический поиск, обобщение в пространстве понятий, алгоритм исключения кандидата, индуктивный алгоритм построения дерева решений ID3 [Люгер, 2003, с. 376, 380]. В.Н. Финн в качестве такого механизма предлагает свой ДСМ-метод [Финн, 2011].

Ж.-Л. Лорьер называет свой метод «программой обучения планам» [Лорьер, 1991] и т.д.

Основным недостатком всех этих подходов, по мнению не только самого Люгера, но и целого ряда других исследователей, является требование наличия качественных обучающих выборок. По нашему мнению, источник проблемы лежит более глубоко – основная трудность заключается в сложности генезиса нового знания в условиях недостатка информации, которая не позволяет построить модели представлений знаний при отсутствии в данный момент всех необходимых примеров свойств, понятий и типичных правил.

Кроме того, одной из принципиально важных является проблема получения первичных знаний, необходимых для начального, иницирующего момента в процессе обучения ИИ-систем. Это с необходимостью ставит вопрос о смещении акцентов в сторону гибридных, человеко-машинных систем. Роль человека в этом случае, по мнению Лорьера, состоит по меньшей мере в предоставлении первичных знаний [Лорьер, 1991, с. 517]:

«...каким бы ни был уровень, вопрос о том, чтобы обучаться, начиная «с нуля», не стоит и обсуждать. Именно минимум знаний для процесса обучения в той среде, где мы находимся, представляет основную трудность для программирования системы, действительно способной к приобретению знаний».

Еще одной проблемой, демонстрирующей сложность процесса машинного обучения, является проблема формализации условий, при которых должно быть принято решение о целесообразности и непосредственном механизме генезиса нового понятия знака или правила, особенно в условиях отсутствия большого объема эмпирически полученных фактов. Основная ценность науки заключается в ее предсказательной силе. И здесь реализация способности системы машинного интеллекта строить теоретические модели,

обладающие предсказательной силой в условиях недостатка информации, является одной из наиболее актуальных проблем.

Таким образом, проблемы философского осмысления процесса обучения, анализ существующих и потенциально возможных дидактических систем и стратегий являются одними из ключевых в задачах общего ИИ.

Если рассматривать в качестве таковых уровни предлагаемые Лорьером, то относительно уровней 1-3 успешность частных инженерных реализаций, ограниченных конкретными прикладными задачами, не вызывают сомнения.

Относительно 4-го уровня уже возникает целый ряд проблем, прежде всего связанных с трудностями обучения, возникающими при наличии малого количества обучающих примеров, с высокими требованиями к их качеству, а также с неспособностью таких систем к самообучению.

На наш взгляд, сегодня не существует сколь-нибудь успешных инженерных реализаций систем уровня 5.

Для дальнейшего понимания степени разработанности проблемы обучения речь будет идти об уровнях 4 и 5. Кроме того, по нашему мнению, задачу необходимо разделить на три подзадачи:

1. порождение/изменение новых понятий;
2. порождение новых знаков (символов) (процесс символизации);
3. порождение/изменение новых правил.

Одной из причин такого положения дел является та же проблема, что и проблема генезиса теорий в связи с опытом. «Теория, хотя и обусловлена опытом, но непосредственно не может быть получена из индуктивного обобщения данных» [Степин, 2015, 15]. К сожалению, философское осмысление теоретических идей, обеспечивающих процессы генезиса новых понятий, знаков и правил, является сегодня одним из наименее разработанных моментов. Начиная с 1940-х годов позитивизм (и следующий за ним неопозитивизм) отказался от поиска подходов, которые могли бы с логико-методологической точки зрения обеспечить генезис новых теоретических

идей [Степин, 2015, с. 17]. Но современный дискурс продолжает вестись в русле позитивистских и, по выражению В.С. Степина, «созерцательно материалистических» концепций. По его мнению, для решения проблемы необходимо осуществить поиск из совершенно иных гносеологических представлений. Необходимость такого поиска является одним из тех ключевых условий, которые должны обеспечивать процессы обучения систем машинного интеллекта.

На наш взгляд, проблематика обучения и самообучения ИИ-систем центрирует и в наибольшей степени позволяет понять как принципиальную роль человеческого субъекта в процессе реализации интегрированных в человеческое общество систем общего ИИ, так и место различных технологических подходов в процессе реализации этих систем.

К числу наиболее значимых причин, препятствующих разрешению указанных нами проблем, сегодня следует отнести, помимо исключения философской проблематики из современного дискурса в сфере проектирования ИИ-систем, также недооценку роли символьных методов в задачах общего ИИ.

Недооценка роли символьных методов в задачах общего ИИ

Д.А. Поспелов в статье «Десять горячих точек...» пишет о борьбе и смене в 1980-х приоритетов между символьными (продукционными) и коннекционистскими (сетевыми) подходами. Результатом этого стало практически полное размежевание исследовательских программ и возникновении де-факто двух различных наук. Его слова можно считать во многом пророческими [Поспелов, 2019, с.9]:

«Этот разрыв нарастает, что, по-видимому, приведет к становлению двух разных наук, связанных с построением интеллектуальных систем. Одна из них будет по-прежнему опираться на уровень ментальных (информационных) представлений, а другая — на уровень структурной организации (по типу нервных тканей), порождающий нужные решения. Во всяком случае, в начале нового тысячелетия вряд ли можно ожидать спада интереса к сетевым моделям и многочисленным нерешенным проблемам, связанным с их построением и функционированием.»

Существующие сегодня очевидные успехи коннекционистских методов в области распознавания образов, поиска паттернов путем поиска и выявления статистически значимых связей внутри больших массивов предварительно размеченных данных. Все это только усиливает указанную тенденцию. Подтверждением этого является анализ структуры большинства конференций в области ИИ, о чем уже упоминалось выше. В перечне докладов последних 10–15 лет (начиная с 2000-х годов) практически полностью отсутствуют доклады на тему символьных методов. Существующие сегодня немногочисленные попытки анализа гибридных систем, такие как исследования группы Гридина [Гридин, 2017] носят ограниченный характер. Основная сложность здесь заключается в проблематике обучения ИИ-систем в условиях отсутствия исчерпывающей информации, когда обобщение опыта требует генезиса новых понятий, правил и знаков.

Аналогичная ситуация складывается и в области перевода, где на первый взгляд достигнуты значительные успехи. Но, рассматривая имеющиеся решения в области машинного перевода, нужно констатировать, что они, в основном, используют частотный анализ парных вхождений слов (словоформ, текстовых фрагментов). Суть же проблемы заключается в контекстной зависимости выражений и их значений, что усугубляется наличием разных прагматических контекстов не только для носителей разных

языков, но и для коммуникантов, принадлежащих к разным культурным и профессиональным слоям.

Проблема обучения не менее актуальна и для символьных методов, о чем, в частности, пишет Джордж Люгер [Люгер,2003]. Основной акцент он делает на обучении как процессе обобщения опыта и связанной с этим проблемой неосуществимости полной индукции. Последняя понимается им как проблема выбора инвариантов в рамках ограниченного пространства вариантов, где сама по себе индукция, не является достаточным условием производства нового знания. Одной из главных причин такого положения дел является, невозможность решения проблемы обучения без технической реализации процессов генезиса новых базовых понятий и продукционных правил, без процесса генезиса новых знаков системой общего ИИ, функционирование которой предполагается в широком спектре предметных областей.

На сегодняшний день сложилась ситуация, при которой все попытки решения задачи создания общего ИИ предпринимаются главным образом в рамках одних только коннекционистских методов. Такое положение дел позволяет фиксировать проблему установления места и роли символьных методов в задачах общего ИИ.

Отметим также, что потребность в экспертных системах, способных не просто к выявлению каких-либо паттернов, а к полноценному логически и предметно обоснованному выводу остается актуальной для целого ряда областей человеческой деятельности. Примерами таких областей являются планирование и управление производством, медицина, автономное функционирование технических средств (принятие решений в сложных и динамично изменяющихся обстоятельствах, когда возможности присутствия человека ограничены) и т.д.

Сказанное позволяет сформулировать общую **проблему** изучения онтологического, гносеологического, эпистемологического и инженерного аспектов *роли символьных методов в задачах общего ИИ.*

Основные понятия и термины

Сильный ИИ

В рамках настоящего исследования под термином «**сильный ИИ**» будем понимать определение, данное Джоном Сёрлем как ИИ-системы как

демонстрирующей «точно такое же поведение, которое человек демонстрирует в отношении некоторого когнитивного феномена» [Parsing the Turing Test, 2008, p. 140] и в буквальном смысле слова обладающей когнитивными состояниями [Searl, 1980].

Общий ИИ

К сожалению, на сегодняшний день не существует согласия относительно того, что понимать под термином «**общий ИИ**» и «**сильный ИИ**». На наш взгляд, существующее сегодня понимание «**сильного ИИ**» является в значительной степени размытым и требующим уточнения. Прежде всего, это связано со специфичностью самого теста и невозможностью однозначной оценки того, смогла ли в действительности та или иная ИИ-система пройти тест. Среди таких неоднозначностей можно выделить, во-первых, отсутствие конкретных требований к составу экзаменационной комиссии. Во-вторых, тот факт, что способность к поддержанию диалога только косвенно связана со способностью к рациональному мышлению и самообучению. Более детальная критика теста Тьюринга выходит за рамки настоящего исследования.

В рамках настоящего исследования предлагается ограничить понимание **сильного ИИ**, путем разделения этих понятий и определения термина **«общий ИИ»** в качестве

системы (инженерного решения), моделирующей интеллектуальную деятельность человека, в которой реализованы процессы обучения и самообучения (производство нового знания) в целях принятия самостоятельных рациональных решений понимаемых как выбор одного из вариантов действий среди множества возможных (а также смоделированных в качестве прогнозных самой ИИ-системой) альтернатив в условиях неполноты исходной и текущей информации, обусловленной динамично изменяющимися условиями внешней среды.

В качестве примеров **задач общего ИИ** можно привести такие задачи как: построение диалоговых систем; создание автономных роботов для управления транспортными средствами на дорогах общего пользования, а также для работы в условиях агрессивной окружающей среды; космических аппаратов; системы слежения в медицине и т.д.

Гибридный ИИ.

В настоящее время существует два понимания термина «гибридный ИИ». Во-первых, как ансамбль различных технических подходов (в работах Педро Домингоса, Гэри Маркуса и т.д). Во-вторых, как человеко-машинная система так, как понимал это Венда В.Ф.

В рамках настоящего исследования термин «гибридный ИИ» понимается прежде всего как человеко-машинная система в соответствии с идеями В.Ф. Венда.

Символьные методы

По мнению Дональда Э. Кнута, уже начиная с 1960-х годов компьютеры все чаще начинают применяться в задачах, в которых непосредственно вычисления начинают играть второстепенную роль. На первый план начинают выходить задачи, направленные на принятие решений [Кнут, 2001, с. 13].

Признавая глубокую связь между математикой и информацией, исследователь считает, что предмет их имеет существенные отличия. Если математика занимается *теоремами, бесконечными процессами, доказательствами существования объектов*, то информатика как прикладная наука, занимается *конечными процессами, и непосредственно конструктивными объектами как уже данными* [Кнут, 1977]. Алгоритмы, с его точки зрения, представляются уже не столько предметом исследования, сколько (вместе с языком) непосредственными инструментами и механизмами мышления.

Важным этапом на пути такого разделения математического и кибернетического понимания концепта «символические методы» является гипотеза, выдвинутая Аланом Ньюэллом и Гербертом Саймоном в ходе реализации проекта «Универсальный решатель задач» [Newell, 1976]. Они полагают, что *существует принципиальная возможность создания физической системы, способной к осуществлению основных интеллектуальных операций на любом физическом носителе.*

Необходимыми условиями для существования физической символической системы они считают [там же]:

1. Требование существования набора элементов-символов, имеющих какую-либо физическую реализацию. Символ может использоваться для обозначения любого выражения без ограничений. Этот произвол относится только к символам; лексемы символов и их взаимные отношения определяют, какой объект обозначается сложным выражением.
2. Системы символов представляют собой наборы шаблонов и процессов, причем последние способны создавать, разрушать и модифицировать первые.
3. Существуют выражения, которые обозначают каждый процесс, на который способна машина.
4. Существуют процессы для создания любого выражения и для изменения любого выражения произвольными способами.

5. Выражения стабильны; после создания они будут продолжать существовать до тех пор, пока явно не будут изменены или удалены.
6. Не ограничено количество выражений, которые может содержать система.

В рамках предлагаемого ими подхода алгоритмы начинают пониматься уже во втором «кибернетическом» смысле. При этом алгоритмы рассматриваются как явления, связанные с вычислительными машинами [там же]. Тем самым алгоритмы становятся инструментами, которые реализуют те или иные состояния и сценарии поведения системы, где необходимость эмпирических исследований становится значимой не менее теоретических. Принципиально важным аспектом при таком подходе является необходимость исследования связи между теоретической моделью и реальной ситуацией.

Кроме того, центральными для понимания символов и выражений являются такие понятия как «обозначение» и «интерпретация». Так, выражение обозначает объект, если система может влиять на объект или вести себя способом, аналогичным объекту. Толкование понимается как способность вызвать и выполнить какой-либо процесс в зависимости от выражений, которые их определяют. Тем самым достигается, в частности, корреляция между теоретическими моделями и реальной действительностью. Оценка поведения должна осуществляться исключительно по последствиям, не зависимо от того выполняется она людьми или нет. Таким образом:

На абстрактном теоретическом уровне под «символьными методами» будем понимать процедуры манипулирования символами как многоуровневыми представлениями.

В аспекте же инженерных реализаций, под «символьными методами» будем понимать совокупность инженерных решений, реализующих процедуры создания и преобразования символьных выражений.

Примерами алгоритмов, основанных на продукционных правилах, являются экспертные системы Фредерика Хейес-Рота, Дональда Уотермана и Дугласа Лената, ДСМ-метод В.К. Финна и т. д.

Коннекционистские методы

Под коннекционистскими методами будем понимать совокупность технологических подходов, включающих в себя алгоритмы, основанные на сетевых моделях как моделях архитектуры мозга.

В настоящее время существует большое количество разнообразных моделей, таких как цепи А.А. Маркова, нейронные сети Джона Хопфилда, Ричарда Хэмминга, «машина Больцмана» Джеффри Хинтона и Терри Сайновски, сверточные нейронные сети Яна Лекуна и т. д.

Роль символьных методов

«Роль» будем понимать как определенный набор действий, которые может выполнять тот или иной элемент-актор внутри той или иной системы.

«Роль» также часто понимают как степень *значимости* тех или иных элементов-актеров относительно друг друга и того вклада, который каждый из них вносит в общий результат. В этом смысле говорят о «ведущей», «подчиненной» или «второстепенной» роли. Сегодня в задачах реализации ИИ-систем доминируют коннекционистские подходы, притом что отсутствуют теоретические основания их выбора в качестве основных. Но при этом остаются недостаточно исследованными и эксплицированными возможности использования символических методов как элементов технологических подходов к задачам общего ИИ.

Технологический подход

Под «технологическим подходом» будем понимать множество инженерных решений, реализующих символные, коннекционистские и т.д. методы.

«Подручность»

Давая определение общего ИИ мы декларировали тот факт, что система общего ИИ должна быть способной к самостоятельному принятию решений в широком спектре задач в динамично меняющихся условиях. Отличительной чертой такой системы является способность к обучению и самообучению в ходе своего функционирования. Такого рода постановка вопроса с необходимостью ставит вопрос о степени автономности ИИ-системы. Для понимания этой проблематики становится актуальным поиск ответа на вопрос о том, является ли создаваемая нами ИИ-система зависимой от человека, и если да, то в каком именно образом.

Начиная с Эрнста Каппа, Хосе Ортега-и-Гассета техника понималась исключительно как продолжение человека (органопроекция у Каппа) или как «реакция» человека на внешние обстоятельства у Ортега-и-Гассета.

Позднее, Мартин Хайдеггер, рассматривая сходную проблематику, вводит понятие подручности. Подручность у него понята через обращение Dasein (Человека) с внутримирным сущим. Средство – как вещь – не просто «бывает» в смысле моментального присутствия определённого нечто в человеческом окружении, а существует в несамостоятельном, связанном с присутствием человека модусе «для того чтобы». «Сработанное изделие», по Хайдеггеру, отсылает, в том числе, к своему пользователю. Именно по «телу» пользователя происходит «выкраивание» изделия. [Хайдеггер, 2015, с.70].

Однако, для понимания ИИ, такое понимание подручности не является достаточным. Для философа подручность связана с «сакральным» и «неподлинным», где существование техники рассматривается из перспективы индивидуальной экзистенции человека, из базовой включенности человека как

в мир людей, так и в мир вообще. В рамках такого понимания невозможно понять место ИИ-системы как некоторого субъекта, более активно (нежели «просто» машины и инструменты), включенного в различного рода взаимодействия.

Более поздние исследования таких философов как Жильбер Симондон, Бруно Латур привели к новому пониманию места технических объектов как полноценных акторов, включенных в целый ряд интерсубъективных взаимодействий и, прежде всего, социальных коммуникаций. Однако, техника в их понимании все же остается скорее средством медиации между человеком и природой; актором, ограниченным рамками трансиндивидуальной коммуникации [Симондон, 2011].

Система ИИ, являющаяся предметом настоящего исследования, обладает в значительно большей степени признаками субъектности, чем тем технические объекты, о которых идет речь у вышеперечисленных философов. Отсюда в полной мере встает вопрос о необходимости исследования степени автономности и степени зависимости ИИ-систем от человека в вопросах обучения, самообучения и самостоятельного принятия решений в процессе их функционирования. В целях такого исследования необходимо определить особенный смысл, в котором далее будет использоваться концепт «подручность» (в отличие от того смысла, в каком он понимается у М.Хайдеггера):

под «подручностью» в настоящем исследовании будем понимать способ существования ИИ-системы в человеческом мире как системы, автономной в широком диапазоне условий и задач, причём действия этой системы зависят от человека и человеческого мира в вопросах обучения, самообучения и принятия решений в ходе своего функционирования.

В трактовке подручности мы исходим из перспективы гибридных взаимодействий естественных и искусственных объектов и субъектов.

Знание.

В случае системы машинного интеллекта речь идет о формировании некоторого экземпляра концептуальной системы, реализуемого средствами программно-аппаратных комплексов. В рамках настоящего исследования примем следующие положения:

Под «знаниями» в рамках настоящего исследования будем понимать произведенные в результате процессов генезиса и поэзиса новые/измененные понятия знаки (символы) и правила вывода, организованные в единую концептуальную систему, понимаемую в дальнейшем как систему знаний того или иного экземпляра системы общего ИИ.

«Познавательная деятельность» понимается как процесс производства знания некоторым субъектом (человеком, ИИ-системой).

Под «опытом» будем понимать знания, полученные в результате познавательной деятельности. Производство знания предполагается на основании поступающей извне или уже имеющейся у системы информации.

Под «обучением» будем понимать процесс передачи от учителя к ученику (ИИ-системе) опыта в форме знаний, в котором участвуют не менее двух субъектов.

Под «самообучением» будем понимать процесс производства нового знания, производимый одним субъектом.

Обучение и самообучение.

Рассматривая ИИ-систему не просто как изделие, но и с точки зрения закладываемой в ней способности к обучению, с необходимостью встают вопросы обучения и самообучения. На наш взгляд, одной из первых среди этих

проблем, является проблема первоначального обучения, проблема первичного, иницирующего акта в этом процессе. В связи с чем с необходимостью встает вопрос о способе «обнаружения» себя со стороны ИИ-системы:

во-первых, как одного из акторов двустороннего процесса обучения, в котором человек выступает в роли учителя;

во-вторых, как ограниченного рамками человеческой культуры как той «сверхприроды» (Ортега-и-Гассет), которая формирует цели и ценности, без которых невозможны процессы генезиса новых смыслов.

По мнению В.В. Краевского следует проводить четкое различие между наукой, как деятельностью «в которой происходит выработка и теоретическая систематизация объективных знаний о действительности» [Краевский, 2006, с.7] и педагогической деятельностью, направленной на *передачу опыта*, в которой наличие пар «ученик-учитель» и «ученик-учебный материал» является обязательной. Важно отметить, что, по его мнению, процесс обучения имеет в значительной степени нормативный характер.

Таким образом, процесс *самообучения* носит принципиально иной характер относительно процесса обучения. Соглашаясь с пониманием ИИ В.К. Финном как веберовского типа интеллекта, деятельность ИИ-системы в области самообучения может быть понята прежде всего как рациональная, соответствующая принципам научного познания, деятельность.

Понятия. Знаки. Продукционные правила.

Под «понятием» в рамках настоящего исследования будем понимать теоретические объекты как элементы концептуальных структур, порождаемые путем идеализации и схематизации в результате процесса теоретической «реконструкции фрагментов действительности, взятых с определенных сторон и определенным образом схематизированных» [Смирнов, 1964].

Фрагменты концептуальных структур технически реализуются аппаратно-программными средствами как структурные элементы Базы знаний и Базы фактов в понимании их В.К. Финном.

Под «знаками» будем понимать символные выражения и имена, присвоенные полученным теоретическим объектам, с учетом правил грамматики того языка, на котором осуществляется именование.

Под «правилами» в рамках настоящего исследования будем понимать пропозиционные высказывания, реализуемые теми или иными программно-аппаратными средствами.

Примерами таких формальных систем могут являться модель Крипке для интуиционистской логики высказываний, λ -исчисления или другие системы неклассической, прежде всего, логики. К этому следует добавить, что речь идет главным образом об их реализации в качестве частично рекурсивных функций, в соответствии с тезисом Черча-Тьюринга [Клини, 1973, с. 430].

Говоря о понятиях необходимо сказать о проблеме соотношения «системы понятий - системы терминов» [Никитина, 2014, с. 44]. Система понятий здесь выражает структуру знания – онтологию и гносеологию научной области. В то же время система терминов является инструментом «унификации, стандартизации и нормирования» [Никитина, 2014, с. 43], образуя разного рода подъязки, например язык науки. Важно отметить, что не все понятия обретают соответствующие им термины одновременно со своим возникновением. В частности, понятия могут номинироваться с помощью «дидактических» или «педагогических» метафор, представляющих собой не отдельные имена, а текстовые пояснения [Никитина, 2014, с. 37].

Кроме того, С.Е. Никитина считает важным различать значение (концепт, понятие) и дефиницию или словесно выраженный интенционал. Таким образом, проблема соотношения понятия и термина, выражения их средствами

языка представляют собой комплексную проблему, являющуюся предметом терминоведения. Как указывает Ф.В. Риггз, терминоведение (терминология) тесно связана с такими дисциплинами как семантика, лингвистика и т.д. Философия, в свою очередь, связана с терминоведением через логику и концептоведение [Никитина, 2016, с.43].

Переходя непосредственно к понятиям, необходимо отметить, что понятия можно определить как феномены, возникающие в процессе перехода от процесса живого созерцания к абстрактному мышлению. Так, результаты наблюдений всегда уже являются результатом *схематизации*, что делает акт наблюдения категориальным, тем самым вводя его область мышления [Смирнов, 1964, с.24]. Выражая структуру знания, понятия не могут быть эксплицированы исключительно в терминах непосредственного наблюдения. Более того, как пишет В.А. Смирнов, неправильно было бы думать, что предложения чистого наблюдения являются «базисными» предложениями как это считали позитивисты [там же]. Схематизация, таким образом, является неизбежным этапом при переходе от непосредственной наблюдаемости к описанию явления в терминах того или иного языка.

В.С. Степин также отмечает, что теоретические схемы не являются результатом чисто индуктивного опыта. Построение теоретических моделей путем непосредственной схематизации осуществляется только на ранних этапах построения теоретических моделей [Степин, 2000, с.314]. Позже, по мере теоретического освоения, непосредственная схематизация эмпирического опыта используется лишь в качестве функции для построения *абстрактных объектов* теоретических моделей. Понятия необходимо рассматривать в качестве прямого результата схематизации как в процессе первичного выдвижения гипотез, так дальнейшего их обоснования.

Цели и задачи исследования

Целью настоящего исследования является:

Определение роли символьных методов в задачах общего искусственного интеллекта.

На наш взгляд, одним из главных направлений выхода из сложившейся затруднительной ситуации является возникновение полноценной исследовательской программы, в которой символьные методы, так же, как и коннекционистские, должны стать обязательными элементами. Если рассматривать символические методы как аналитические, а коннекционистские как эмпирические, то можно согласиться с высказыванием Джорджа Люгера [Люгер, 2003, с. 780]:

«...если мы пытаемся довести исследования искусственного интеллекта до уровня науки <...> то в процессе конструирования, использования и анализа артефактов должны применять смесь из аналитических и эмпирических методов.»

В понимании автора конструирование ИИ-системы есть уже сам по себе эксперимент, вопрос, «задаваемый разработчиками природе» [там же]. Результатом этого является конкретный экземпляр ИИ-системы, осуществляющий те или иные действия. Каждая программа в его понимании должна рассматриваться как эксперимент, который ставит вопросы перед природой, а ответами на них являются результаты выполнения программы.

В этом смысле основной задачей настоящего исследования является определение роли символьных методов сквозь призму понимания способа их существования в рамках модели, представляющей собой единую систему: человек - коннекционистские методы – символьные методы.

В рамках такой системы каждый элемент является необходимым и выполняющим свои специфические роли. Включение символьных методов является необходимым в ходе решения как минимум следующих задач:

- Обмена информацией в системах «машина-человек» и «машина-машина», с учётом его специфических особенностей: ограниченности скорости и объемов передачи информации; неизбежности наличия рассогласованности концептуальных структур каждого из участников.
- Создания эффективных механизмов самообучения интеллектуальной системы, прежде всего - в условиях динамически меняющихся условий внешней среды и значительной автономности процессов производства новых знаний.
- Реализации механизма символических операций как преобразований на трёх уровнях (знаков, значений и понятий).

Прежде всего речь идет об обосновании необходимости включения символьных методов в процессы самообучения ИИ-систем, в том числе с использованием обучающих выборок, аналогично тому, как это происходит при обучении нейронных сетей.

Основными задачами настоящего исследования являются:

1. Установление основных этапов процесса исторического развития исследовательских программ области ИИ, а также анализ процессов возникновения и дальнейшей эволюции понимания роли символьных методов при проектировании ИИ-систем и в том числе:
 - 1.1. установление и анализ исторического периода, характеризующегося поиском основополагающих принципов, позволяющих моделировать внутренние процессы мышления, присущие человеку вплоть до момента возникновения первых интеллектуальных физических систем;
 - 1.2. установление и анализ исторического периода, характеризующегося возникновением первых физических реализаций интеллектуальных систем вплоть до периода создания первых электронных вычислительных систем;

- 1.3. установление и анализ периода от момента возникновения первых электронных интеллектуальных вычислительных систем до наших дней;
 - 1.4. исследование исторических процессов возникновения непосредственно символьных методов, а также эволюцию в понимании роли символьных методов в ходе проектирования ИИ-систем;
 - 1.5. исследование современного философского дискурса в области проектирования ИИ-систем.
2. Определение онтологического статуса существования ИИ-систем, в том числе:
 - 2.1. определение сущности техники и граничных условий инженерной реализации ИИ-систем, моделирующих интеллектуальную деятельность человека;
 - 2.2. определение места техники и человека в социальном как созданном самим человеком искусственным мире;
 - 2.3. проведение критического анализа сильных и слабых сторон основных технологических подходов, таких как коннекционистские и символьные методы;
 - 2.4. проведение критического анализа возможности выхода ИИ-систем за границы «подручности» человеку;
 - 2.5. обоснование необходимости создания гибридных систем общего ИИ как человеко-машинных систем;
 - 2.6. определение граничных условий существования ИИ-систем с точки зрения возможности автономии ИИ-систем в процессах самообучения и самостоятельного принятия решения в ходе их функционирования.
 3. Определение основных ролей символьных методов в задачах общего ИИ, в том числе:
 - 3.1. обоснование тезиса о необходимости символьных методов в задачах генезиса новых понятий, знаков и продукционных правил;

- 3.2. обоснование тезиса о возможности воспроизводства процессов генезиса новых понятий, знаков и правил в рамках машинного интеллекта как языковых моделей;
- 3.3. обоснование тезиса о ведущей роли символьных методов в процессах генезиса новых понятий, знаков и правил в ходе процесса обмена информацией при осуществлении актов коммуникации в рамках моделей «машина-машина» и «человек-машина»;
- 3.4. обоснование тезиса о необходимости символьных методов в задачах обмена информацией при осуществлении процессов коммуникации в рамках моделей «машина-машина» и «человек-машина».

Объект и предмет исследования

Объект исследования – многообразие исследований, подходов и разработок в области концептуализации и реализации искусственного интеллекта.

Предмет исследования – символьные методы в общем контексте задач концептуализации и реализации общего искусственного интеллекта.

Методология исследования

В ходе исторического анализа исследований в области общего искусственного интеллекта и символьных методов, а также в ходе построения теоретической модели концептуальной структуры процедур генезиса новых понятий, знаков и правил, применялись элементы системного, герменевтического, структурного и конструктивного подходов. Использовались некоторые методы теории информатики.

Положения, выносимые на защиту

1. ИИ является техническим объектом, готовым к использованию и «подручным» человеку. Первоначальное обучение ИИ-системы невозможно без человека. Анализ того, как должен быть устроен общий ИИ должен проводиться в рамках прежде всего гибридных человеко-машинных концепций.
2. Граничные условия, которым должен удовлетворять процесс самообучения, нарушение которых приводит к невозможности существования ИИ-системы или потере ценности полученных результатов однозначно задаются концептуальными системами человека.
3. Эффективная система общего ИИ должна с необходимостью представлять собой систему гибридного интеллекта, включающую в себя человека, коннекционистские и символьные методы.
4. Роль символьных методов определяется тем, что символьные методы являются эффективными инструментами и механизмами, обеспечивающими:
 - a. процессы (само)обучения ИИ-систем путем генезиса или трансформации существующих понятий, знаков и правил в том числе с использованием обучающих выборок;
 - b. процессы коммуникации в рамках моделей «человек-машина» и «машина-машина».

Научная новизна исследования

Все вышесказанное определяет необходимость поиска решения одной из наиболее критически важных проблем - проблеме машинного самообучения в аспекте рассмотренных выше проблем 1-3. В соответствии с целью настоящего исследования - определения роли символьных методов в задачах общего ИИ - осуществлен поиск философских методологических оснований для проектирования такого рода творческих механизмов в рамках систем

машинного интеллекта за счет: 1) переосмысления роли символьных методов; 2) обоснования принципиальной возможности средствами символьных методов выхода за пределы существующего сегодня замкнутого круга аналогия – операция обобщения. С этой целью:

1. *Разработана методология, обеспечивающая процедуры генезиса новых представлений, знаков и правил.* В частности, разработаны базовые правила, позволяющие в значительной степени расширить возможности символьных систем в области производства нового знания. Символьные методы могут выступать в роли учителя в задачах обучения коннекционистских (сетевых) моделей.
2. *Определена роль символьных методов как необходимых в обеспечении коммуникации как между машиной и человеком, так и между машинами как таковыми.*
3. *Определены граничные условия, которым должен удовлетворять процесс самообучения, нарушение которых приводит к невозможности эффективного функционирования ИИ-системы или потере ценности полученных результатов.*
4. *Дается обоснование требования о невозможности и нецелесообразности построения полностью автономных систем общего ИИ.*
5. *В качестве наиболее приоритетного пути развития ИИ обосновывается необходимость гибридных, человеко-машинных систем, исключая, по крайней мере на современном этапе, возможность создания полностью автономного машинного интеллекта, осуществляющему полностью автономные процессы самообучения, без присутствия человека.*

Теоретическая и практическая значимость результатов исследования

Основной результат представленной работы – теоретическая модель, которая может быть положена в качестве основания для разработки методологии, обеспечивающей процессы обучения как с учителем, так и без в рамках символьных методов на базе обучающих выборок аналогично тому, как это сегодня реализуется в рамках коннекционистских подходов.

Определена роль символьных методов как источника новых понятий, знаков и правил.

Кроме того, значимость в теоретическом плане определяются тем, что:

1. осуществлен анализ пределов творческих возможностей систем машинного интеллекта;
2. определена роль человека как необходимого и ограничивающего ценность результатов обучения ИИ-системы условия;
3. осуществлен анализ причин, почему внутреннее устройство систем машинного интеллекта должно быть реализовано в парадигме гибридного ИИ.

В практическом плане значимость определяется тем, что разработанная теоретическая модель может быть реализована как значимая в плане практической реализации методология, позволяющая значительно расширить творческие возможности систем машинного интеллекта. Прежде всего речь идет о процессах обучения в ситуациях недостатка и или отсутствия достаточного объема эмпирических данных.

Степень достоверности и апробация результатов исследования

Достоверность полученных результатов достигается автором при помощи применения научной методологии и принципа объективности. Диссертационное исследование опирается на широкий круг источников и исследовательскую литературу на русском, английском и немецком языках.

Диссертация была обсуждена и рекомендована к защите на заседании кафедры философии и методологии науки философского факультета Московского государственного университета имени М.В. Ломоносова.

Основные положения и выводы диссертационного исследования были изложены в 4-х научных статьях, опубликованных в изданиях, отвечающих требованиям п. 2.3 Положения о присуждении ученых степеней в Московском государственном университете имени М.В. Ломоносова.

Основные результаты диссертационного исследования представлены в рамках научно-практических конференций и семинаров:

1. Международная конференция «Суперкомпьютерные дни в России» (Москва, 24.09.2019) – доклад «Суперкомпьютерные системы в задачах распознавания смысла текста»;
2. Международная конференция «Суперкомпьютерные дни в России» (Москва, 22.09.2020) – доклад «Роль символического подхода в задачах общего искусственного интеллекта»;
3. Международная междисциплинарная конференция «Искусственный интеллект в новой коммуникативной реальности» (Москва, 06.11.2020-07.11.2020) – доклад «От гибридных интеллектуальных систем к гибриднему интеллекту»;
4. Научный семинар «Методологические проблемы современной науки: эпистемология научных объяснений» (Москва, МГУ имени М.В.

Ломоносова, 20.10.2021) – доклад «Является ли ИИ чем-то большим, чем техника?»;

5. Всероссийская научно-практическая конференция с международным участием «Философские контексты современности: искусственный интеллект и интеллектуальная интуиция» (Ижевск, 25.02.2022-26.02.2022) – доклад «Онтологические проблемы в задачах общего искусственного интеллекта»;
6. XVIII конгресс «Нейронаука для медицины и психологии» (Судак, 30.05.2022-10.06.2022) – доклад «Гносеологические проблемы в задачах общего искусственного интеллекта».

Структура диссертации

Диссертационная работа состоит из введения, трех глав, заключения, одного приложения и списка литературы.

Глава 1 Исторический анализ процессов возникновения и эволюции понимания роли символьных методов в задачах проектирования ИИ-систем

Историю возникновения искусственного интеллекта, на наш взгляд, можно условно разбить на три периода.

К первому, «дофизическому» периоду, можно отнести промежуток истории начиная с Пифагора и до возникновения первых прототипов интеллектуальных физических систем в виде машин Беббиджа [Петренко, 1979] и Корсакова [Алексеев, 2019]. Таким образом этот период охватывает порядка 25 столетий начиная с VI в. до н. э. и продолжается до XIX века н.э..

Второй, «физический», продолжавшийся чуть больше одного столетия вплоть до появления в середине 1940-х первых электронных машин, таких как «Bomb» [Wright, 2017] Тьюринга, ENIAC от IBM [McCarthy S., 1999] и т.д.

Третий, «компьютерный», самый короткий, с середины 1940-х продолжающийся до сегодня.

Первый период, характеризующийся поисками самих принципов, позволяющих моделировать внутренние процессы мышления, присущие человеку. Именно в этот период формируются идеалы научного метода и классической рациональности без которых было бы невозможно создание современных вычислительных систем.

Второй период ознаменован первыми относительно успешными попытками создания «интеллектуальных» машин, иными словами, физических систем, способных воспроизводить определенные интеллектуальные операции самостоятельно, хотя и после предварительной настройки оператором. Создание таких машин если и не изменило хода

истории в тот период, но в известной степени повлияло на ход философской мысли так как подтверждало идеи Декарта и Гоббса о том, что мыслительные процессы не отличаются от механических. Характерной особенностью второго периода необходимо считать возникновение нового феномена – первых механических вычислительных устройств, значительно усиливших картезианские тенденции в философии науки.

Третий, «компьютерный» период ознаменовывается появлением электронных устройств, что привело к научно-технической революции середины XX века и смене технологического уклада, свидетелями которых мы являемся в настоящее время.

§ «Дофизический» период

По мнению Бертрانا Рассела дедуктивное обоснование вывода, по крайней мере в математике, начинается именно с Пифагора в VI веке д.н.э. [Рассел, 2018]. Правда, замечает Рассел, главным образом это было связано скорее с особым рода мистицизмом. Если большинство наук в то время были связаны прежде всего с «некоторыми формами ложных верований» [Рассел, 2018, с. 63], астрономия - с астрологией, химия - с алхимией, то начиная с Пифагора математика становится уже не просто практическим знанием, но чем-то большим, что дает «идеал знания», полученного путем чистого размышления. На основании этого вывода было сделано предположение, что мысль выше чувства, и если что-то в мире не укладывается в идеальные рамки математики, то тем хуже для самого мира. Ввиду того, что чувственные предметы не совершенны, а совершенны только идеи, то становится очевидной мысль о том, что «мысль благороднее чувства, а объекты мысли более реальны чем объекты чувственного восприятия» [Рассел, 2018, с. 64].

Начавшееся с Пифагора сочетание математики и теологии привело к дальнейшей рационализации западноевропейской мысли. Именно эта исходная рационализация, на наш взгляд, и послужила в дальнейшем отправной точкой на пути возникновения самой идеи о возможности

реализации физических интеллектуальных систем на субстратах отличных от биологического. Прежде всего, имеются в виду исследовательские программы по созданию сильного искусственного интеллекта в понимании Сёрля [Searle, 1980].

Важной вехой в дальнейшем развитии идей Пифагора, уже столетием позже, являются идеи Парменида. По утверждению [Рассел, 2018, с. 80] Парменид «в действительности изобрел метафизику, основанную на логике», а не логику как таковую. Он был первым, кто отделил «путь истины» от «пути знания», а также первым [Рассел, 2018, с. 81], кто дал пример аргументации от «мысли и языка к миру в целом» [там же]. Кроме того, не употребляя этого термина, Парменид дал понятие «субстанция» [там же], которое стало наиважнейшим понятием в дальнейшем развитии таких наук как философия, психология, физика. С точки зрения нашего исследования, понятие «субстанция» можно считать основой, благодаря которой существует принципиальная возможность создания разного рода баз знаний, разного рода тезаурусов и т.д., так необходимых для создания физических интеллектуальных систем.

К этому стоит добавить Эмпедокла с его идеей о том, что все в мире происходит не с какой-либо целью, а определяется только Случайностью и Необходимостью [Рассел, 2018, с. 88]. Правда, мысль европейской науки еще долго не анализировала серьезно теории хаоса, следуя по дальнейшему пути, проложенному прежде всего Аристотелем. Позже, уже в начале XIX века стохастические процессы были сведены к идее о Лапласовской детерминации. По нашему мнению, идеи Эмпедокла занимают не менее важное место в процессе производства нового знания, чем идеи Аристотеля.

Возвращаясь к рассмотрению нашего вопроса, а именно к IV веку д.н.э., прежде всего необходимо рассмотреть влияние Аристотеля на дальнейшее развитие европейской науки. По мнению Д.А. Поспелова созданная Аристотелем теория силлогизмов, впервые в истории стала целой наукой, созданной одним человеком [Поспелов, 1982]. Конечно же, теория

силлогизмов прежде всего была создана с совершенно конкретной целью: «безупречное» ведение диалогов в духе Сократа. На протяжении столетий логика Аристотеля, изложенная в «Органоне», является практически неизменной. Несомненно, его заслугой следует считать само понятие «силлогизма». Кроме того, Аристотель «сумел четко сформулировать условия, при которых истинность вывода в дедуктивном рассуждении всегда обеспечивается, если большая и малая посылки верны» [Поспелов, 1982, с. 18].

В дальнейшем его ученики, Теофраст и Эвдем, отделили логику от метафизики и теории познания, привели к ее окончательной форме, сформированной Порфирием [там же]. Большинство авторов связывают теорию силлогизмов Аристотеля прежде всего с правилами дедуктивного вывода, но без индуктивных выводов не удалось обойтись даже ему. Дело в том, что знаменитая большая посылка Аристотеля «Все люди смертны» по сути является индуктивной, декларируя наличие множества фактов о смертности отдельных людей. Нигде не дается доказательства всеобщности данной посылки. Следовательно, вывод, формируемый на ее основе, является обобщением. Впрочем, честь четкого пояснения сути индуктивных рассуждений все же принадлежит Френсису Бэкону [Поспелов, 1982, с. 22]. Здесь важно отметить, что несмотря на то, что в ряде случаев даются посылки, полученные на основании индуктивного вывода, именно дедуктивная логика и теория силлогизма, являются одними из величайшей заслуг Аристотеля и его последователей.

С точки зрения понимания процессов мышления и онтологических построений огромной заслугой Аристотеля можно считать идеи категорий как прообраз классов. Много позже универсалии получают свое развитие в логике исчислений классов Буля, а также при формировании онтологий непосредственно в задачах создания физических интеллектуальных систем.

Но прежде идеи родов и универсалий получили свое дальнейшее развитие в работах Авиценны, который сделал первый шаг на пути примирения взглядов реалистов, последователей Платона и Аристотеля, с

одной стороны, и с другой – номиналистов, через тройственное различие сущего как существующего в уме Бога, природе и ее проявлениях, а также в уме человека [Асмус, 2019, с. 40]. Именно идеи Авиценны послужили основой для последующих работ Альберта Великого и Фомы Аквинского.

В своих работах Авиценна развил и уточнил понятия категорий. Прежде всего речь идет о категориях качества и количества. По словам В.Ф. Асмуса, Авиценне была свойственна тенденция к противопоставлению своего понимания тому, что давала платоновско-пифагорейская школа. Если первые считали их началом, то Авиценна свел их к акциденциям. Не менее важным является развитие Авиценовского учения о четырех видах качества суждения, а также подробное развитие мысли о связи Первой и Второй «Аналитик» Аристотеля. По мысли философа «Первая аналитика» была посвящена теории умозаключений, а вторая – теории доказательства. У Авиценны имеются обстоятельные комментарии к первой и второй аналитике Аристотеля, в которых он не только излагает само учение, но и излагает собственные мысли, усиливая то, что было сделано самим Аристотелем.

В это же время появляется первое философское осознание проблемы символа. Так в «Федоне» Сократ говорит о необходимости отвлеченных понятий, в них предлагает он рассматривать истину бытия несмотря на то, что такое уподобление может быть «ущербным» [Платон, 1970, с.70]

Фредерик Коплстон пишет о возникновении в XIII веке разделение средневековых философских школ на школы старого («*via antiqua*») и нового пути («*via moderna*») [Коплстон, 2003, с. 325]. Новое направление Коплстону видится как номиналистическое и терминистическое.

Понятие «терминизм», употребляемый как особое направление в логике, по его мнению, необходимо понимать как синоним номинализма: оно связано с теорией *suppositio*, как изучением самостоятельной функции языковых терминов.

Зарождавшийся в XIII веке кризис, был обусловлен целым рядом причин, среди которых можно назвать рост интереса к анализу конкретных

проблем, в противовес попыткам поиска всеохватывающих систем. Крепла тенденция к анализу вероятных аргументов, истинность которых ранее подвергалась сомнению. Происходит «открытие» новой, аристотелевской логики. Начинает преобладать аналитический подход, критический подход, переходя к эмпирическому [Коплстон,2003, с. 326]. Все это привело к началу размежевания теологии и философии. Середина XIII века ознаменуется целым рядом работ нового толка. Уильям Шервуд публикует «Введение в логику» в 6 частях [Коплстон,2003, с. 330]. В пятую часть он вводит *suppositio*. Выходит трактат Петра Испанского «Краткий обзор логики». В ней он пишет о том, что логика есть наука размышления, управляемая языком как инструментом [там же].

Одним из наиболее ярких представителей школы новой школы является Уильям Оккам. Философом написан целый ряд работ в сфере логики такие как толкования Аристотеля о категориях, толкования на аристотелевские «Об истолковании» и «О софистических опровержениях». Им же написана «Сумма всей логики» [Коплстон,2003, с. 334] и ряд других работ. Одной из главных идей Оккама является убеждение о существовании только одного способа постижения сути вещей – смотреть и понимать. В результате- выводимые нами истинные суждения, являются условными истинами. Философ вводит различие между «реальными» и «умозрительными» науками [там же]. Вводит понятие термина как многозначного понятия, которое 1) может использоваться в широком смысле в качестве подлежащего или сказуемого; 2) термин может соотносится с целым высказыванием; 3) термин может рассматриваться как обозначающее все то, что может быть как субъектом, либо предсказателем в суждении [Коплстон,2003, с. 336]. Вводит понятие естественного символа или знака как условного обозначения какого-либо мысленного понятия.

Следующим важным шагом на этом эволюционном пути являются идеи философов Нового времени. Предшествующий ему Коперникианский переворот первой половины XVI века, утвердил мысль, что непосредственный чувственный опыт может обманывать. И только конструирование мира с

помощью математики может дать ясное представление о природе. Это была революция «с ее постоянным и упорным стремлением математизировать природу, сопровождавшимся не менее упорной и не менее постоянной переоценкой роли опыта и эксперимента» [Койре, 2001, с. IX]. Она заложила не только новый взгляд на мир, но и дала начало новому переосмыслению работы нашего сознания, заново переосмыслило логику Аристотеля.

Первым среди таких философов является Френсис Бэкон. Во многом именно его критика схоластики и прямое указание на недостаточность дедуктивного метода Аристотеля, дали начало дальнейшему переосмыслению методов познания. Кроме того, по утверждению Рассела, он был первым, «кто подчеркивал важность индукции в противоположность дедукции» [Рассел, 2018, с. 680], развитые впоследствии Джоном Стюартом Миллем, хотя последнему до конца так и не удалось в полной мере решить проблему простого перечисления. Стоит отметить, что Асмус не соглашается с такой точкой зрения и возводит индукцию Милля к «агностицизму Юма, а вовсе не к материалистической теории Бэкона» [Асмус, 2019, с.75]. К сожалению, несмотря на уверенность Бэкона в том, что ему известен некий универсальный метод научного познания, в реальности такого эффективного метода предложено не было. Действительно, предлагаемый им в работах «Великое восстановление наук» и «Новый Органон» метод индукции, основанный на сравнительной классификации, является не более чем исключаящим перебором сущностей, обладающих некоторой природой и тех, которые ей не обладают для получения итоговой совокупности только тех «форм», которые методом исключения «должны» ей обладать. И все же не стоит недооценивать метод индукции Бэкона - именно он положил начало дальнейшим исследованиям в этом направлении. Пожалуй, одним из главных недостатков Бэконовской индукции можно считать то, что он полагал свой метод индукции способом получения именно достоверного, а не вероятностного знания [Рассел, 2018, с. 682].

Безусловно, эмпиризм Бекона и рационализм Декарта положили начало современному пониманию методов научного познания и повлияли, в частности, на идеи, развитые позднее Томасом Гоббсом. С нашей точки зрения, именно Томас Гоббс дал в наибольшей степени толчок тем исследованиям, которые в последствии привели к возникновению символических методов; Гоббс был одним из первых философов нового времени кто наиболее рационально подошел к рассмотрению вопроса Разума и основных принципов его работы. Прежде всего, речь идет о его механистическом подходе к пониманию жизни души как к совокупности движений и усилий. Гоббс, в отличие от Декарта, считал, что наша душа также имеет физическую, т.е. механическую природу. Гоббс пишет: «дух все же есть тело, но тонкое» [1989, с.566]. Все действия человека, согласно Гоббсу, определяются внешними воздействиями. Кроме того, наши рассуждения управляются законами и не являются произвольными и зависят от целей нашего мышления.

«Когда человек рассуждает, он лишь образует в уме итоговую сумму путем сложения частей...ибо рассуждение...есть ни что иное как подсчитывание» [Гоббс, 1989, с.30].

Тем самым, как указывает Асмус, именно Гоббс был первым, кто попытался ввести понятие исчисления, но ограничился только выдвиганием такого рода положения [Асмус, 2019, с.116]. Далее, именно Лейбниц был первым, кто впоследствии попытался разработать логическое исчисление, в то время как Гоббс является родоначальником ассоциативной психологии, рассматривая «деятельность нашего ума как основанную на причинных связях или ассоциациях» [Асмус, 2019, с.94]. Обработка первоначальных идей происходит путем их сравнения, сочетания и разъединения. Все науки учат нас, по мысли Гоббса, построению отношений, будь то математика геометрия или логика. Рассуждение является процессом подсчитывания «связей общих имен с целью отметить и обозначить наши мысли» [Гоббс, 1989, с.31]. Отсюда

по необходимости возникает идея номинализма как развития идеи единичного опыта. Метод рассуждений должен быть только синтетическим, при этом доказательства должны строиться на правильной последовательности и быть введены по средствам демонстрации. Доказывающий и слушающий должны следовать одному методу.

Оперируя понятием «благоразумия» как современного эквивалента понятия «вероятность», Гоббс пишет, что мы можем знать только о единичном и связях между единичными фактами. Мы можем лишь благоразумно предполагать, что наши выводы являются верными. Таким образом, научное знание может строиться, во-первых, на правильном употреблении имен, а во-вторых, путем использования правильного метода оперирования этими именами [Гоббс, 1989, с.31].

Не менее значимыми являются мысли Гоббса относительно развития и трансформации идей номинализма. Развивая свои идеи исключительно опытного приобретения знаний, отрицая наличие врожденных идей, философ продолжал идеи номинализма. Прежде всего речь идет о понимании языка как человеческого опыта. В рамках такого понимания слова в его работах часто встречается обозначение слов именами, которые несут только условный смысл по отношению к вещам [Гоббс, 1989, с.17-18].

Дж. Локк в «Опыте человеческого разума» впервые упоминает термин семиотика как один из трех ветвей человеческого познания. Согласно Локку, задача семиотики [Фомин, 2017]:

«рассмотреть природу знаков, которыми ум пользуется для уразумения вещей или для передачи своего знания другим».

Позднее, уже в XIX-XX веках, идей Гоббса и Дж. Локка, Юма, Беркли и ряда других мыслителей Нового времени сформируются в современное понимание семиотики, например, в работах Чарльза Пирса, Ю.М. Лотмана, Джона Дили и т.д..

Таким образом, идеи Гоббса можно рассматривать как первые предпосылки к возникновению как аналитической философии, так и гипотезы Ньюэлла-Саймона [Newell, Simon, 1975].

Касательно влияния Декарта на историю возникновения символических методов, можно сказать, что главным его вкладом являются как минимум две идеи. Во-первых, это идея переосмысления роли математики. Мысль Декарта о всеобщности математики, о том, что математика есть прежде всего наука не о величинах и их вычислениях как таковых, а наука о соотношениях вообще. Во-вторых, сама по себе как таковая постановка вопроса о поисках границы человеческого ума в качестве объекта полноценной исследовательской программы. Что касается непосредственно вклада с точки зрения логики, то следует согласиться с Асмус: «точная характеристика взглядов Декарта на логику-дело сложное» [Асмус, 2019, с.75].

Несмотря на то, что сама по себе проблема индукции возникла гораздо раньше, чем об этом писал Дэвид Юм, Карл Поппер в своей работе «Объективное знание» [2002] начинает анализ проблемы индукции именно с Юма. Как пишет Поппер со ссылкой на самого Юма, прежде всего последнего интересовал вопрос о возможности оправдания человеческих верований какими-либо достаточными основаниями. Юмом были сформированы две проблемы, логическая и психологическая» [Поппер, 2002, с.15].

« H_L : Оправдан ли в наших рассуждениях переход от случаев, [повторно] встречающихся в нашем опыте, к другим случаям [заключениям], с которыми мы раньше не встречались?»

Ответ Юма на H_L : нет, как бы велико ни было число повторений.

H_P : Почему, не смотря на это, все разумные люди ожидают и верят, что случаи, не встречающиеся раньше в их опыте будут соответствовать случаям из их опыта? Иначе говоря, почему мы так уверены в некоторых своих ожиданиях?

Ответ Юма на H_P : это происходит по «обычаю или привычке».

По мнению Поппера, в результате своих рассуждений Юм превратился из одного из самых рационально мыслящих людей в «верующего в иррациональную эпистемологию». Таким образом, первоначально Бэконом и Гоббсом, позднее – Юмом - была сформулирована одна из фундаментальных проблем гносеологии, а именно: проблема того, как наши знания могут расти и продвигаться вперед [Поппер, 2002, с.15].

Заканчивая обсуждение индукции, и продолжая анализ Поппера следует упомянуть о проблеме источника знания, сформированную в том числе Кантом, на что обращает внимание Поппер [Поппер, 2002]. В частности, он указывает на то, что Кант защищает идею о том, что мы не обладаем способностью интеллектуальной интуиции, а наши понятия остаются пустыми или чисто аналитическими, если мы не обращаемся к нашим чувствам или чувственной интуиции. Только пространство и время дают нам возможность упорядочить информацию. В частности, согласно Канту, аксиомы математики основываются на чистой интуиции. Для того, «чтобы следить за доказательством, нам требуется глядеть на (нарисованный) чертеж [Поппер, 2002, с.131].

Следующим не менее важным этапом в истории возникновения прежде всего символических методов является Готфрид Лейбниц. С точки зрения настоящего исследования интерес представляют сразу несколько его идей. Но прежде, чем приступить к их изложению, следует отметить, что Лейбниц твердо верил в то, что логика имеет первостепенное значение не только в математике, но и в вопросах метафизики. Итак, к таким идеям относятся:

Во-первых, «*lingua characteristic*» как формальный язык, на котором могут быть выражены любые знания и «*calculus ratiocinator*» - метод исчисления умозаключений. Предложенная им Универсальная Характеристика представляет интерес как одна из первых систематических попыток создать логику исчисления понятий. Существует как минимум три варианта универсальной Характеристики [Журавлева, 2012]. Многие

исследователи рассматривают их как отправную точку в возникновении разного рода формальных языков включая современные нам языки программирования. Подобного представления придерживаются Г. Фреге, Дж. Буль, Э. Кассирер, У. Эко и ряд других исследователей [там же].

Во-вторых, разработанный им арифмометр, речь о котором пойдет далее.

К сожалению, Лейбниц не успел довести задуманное до конца. Поэтому первым, кому удалось значительно продвинуться вперед в этом вопросе, был, по мнению Д.А. Поспелова, Джордж Буль. Именно Буль сумел представить дедуктивную логику в форме математической системы [Поспелов, 1982, с. 27] и ввел в обиход понятие исчисления. Математик создал три вида исчислений: классов, высказываний и отношений. Впоследствии, исчисления классов и отношений превратились в логику исчисления предикатов. Д.А. Поспелов так охарактеризовал вклад Буля [Поспелов, 1982, с. 27]:

«В чем же заслуга Буля, Ответ на этот вопрос парадоксален. Буль сумел исключить из логики семантику. Во всяком случае, свести ее к точной и однозначно понимаемой математической семантике».

Таким образом, именно идеи Буля послужили тому, что в первой трети ХХ была окончательно сформирована идея формальной системы [там же], что будет чрезвычайно важно для формирования в дальнейшем символических методов.

Последним, но безусловно не единственным, на ком стоит остановить свое внимание, является Джон Стюарт Милль. Идеи, изложенные в работе «Система логики» привели к значительной формализации подходов к процессу познания. Необходимо отметить, что одной из главных предпосылок для этого является позитивизм Огюста Конта, исключивший требование об обязательности объяснений причин наблюдаемых явлений [Степин, 2015, с.15]:

«Законы определялись только по признаку быть устойчиво повторяющимися отношениями явлений».

Милль разделял такую точку зрения, полагая что сознание и материя должны трактоваться не как субстанции, а как понятия, обозначающие особые ассоциативные сочетания ощущений [там же]. Считая индукцию главным методом, получения новых знаний, он отдавал должное и дедуктивным методам. Что более важно, в отличие от Конта, Милль допускал объяснения, выдвижение гипотез, но не как ссылку на сущности, а как апелляцию к казуальным законам [Степин, 2015, с.551]. Соглашаясь с Контом, Милль все же считал гипотезы и дедуктивные методы вспомогательными средствами, и именно в русле этих идей Милль разрабатывал свою индуктивную логику. Согласно математику, именно индуктивные методы переводят первичные гипотезы в ранг казуальных законов.

В рамках настоящего исследования первостепенное значение имеет то, что Милль [Финн, 2010]:

«...пытался не просто обосновать общие принципы индукции, но сформировать правила индуктивных рассуждений как логические средства, подобные доказательству утверждений, представляющих знание, извлеченное из фактов».

Сформулированные им пять правил индуктивного рассуждения легли в основу ДСМ-метода Финна В.К [Финн, 2011].

Таким образом позитивизм, работы Милля, окончательно сформировали условия для возникновения систем машинного интеллекта и символьных методов, в частности.

§ Первые реализации на физическом субстрате.

«Физический»

Если начало теоретических поисков, связанных с разработкой методов научного познания, дедуктивной и индуктивной логики можно отнести ко временам древнегреческой философии, то первые попытки создания «думающих» машин, судя по всему, начались не ранее XIV века. Первым представителем из ряда таких исследователей, пожалуй, можно назвать мистика и фантаста Раймунда Луллия (1235–1315) выдвинувшего идею, о том, что с помощью механического артефакта (механического устройства) можно производить рассуждения [Лапшин,1999, с.312]. Машина Луллия представляла собой устройство для механического комбинирования мыслей и состояла из шести концентрических кругов. На двух обозначались субъекты, на трех - атрибуты, а шестой -неподвижный, содержал вопросы. Как пишет И.И. Лапшин, у Луллия были многочисленные последователи, среди которых можно встретить Джордано Бруно и Лейбница [Лапшин,1999, с.313].

Позднее, в примерно в 1500 году Леонардо да Винчи спроектировал свою суммирующую машину. Описание машины встречается в его дневниках известных сегодня как двухтомник «Мадридский Кодекс». Недавно проведенная реконструкция показала, что его проект является вполне работоспособным.

Период Нового времени XVII-XVIII, характеризующийся бурным развитием науки и техники, очевидным образом породил интерес к созданию разного рода вычислительных машин. Одним из наиболее известных примеров является «Паскалина», построенная в 1642 году Блезом Паскалем. Паскаль считал, что «арифметическая машина производит эффект, который кажется более близким к мышлению по сравнению с любыми действиями животных» [Рассел, Норвиг, 2016, с.41]. Одним важнейших значений «Паскалины» в истории является то, что она укрепила позиции Декарта и

Гоббса, утвердив в будущем мысль о том, что мыслительные процессы не отличаются от механических.

Паскаль был не единственным, кто пробовал создавать такого рода машины. За 20 лет до Паскаля, в 1623 году подобные попытки были предприняты немецким математиком Вильгельмом Шиккардом [там же]. Продолжателем идей Луллия и Бруно об алфавите мыслей является Лейбниц, который в 1673 году он создает «Ступенчатый вычислитель». В его машине использовались специальные шестерни, получившие название «шестерни Лейбница». Принцип ее работы был аналогичен «Паскалине».

В 1805 году Жозеф Мария Жаккард разрабатывает первое программируемое устройство – ткацкий станок, в котором инструкции по плетению узлов записывались на специальных перфокартах.

В 1822 году Чарльз Беббидж публикует статью и в том же году представляет Королевскому обществу модель разностной вычислительной машины. Беббидж выдвинул идею создания универсальной вычислительной машины, которую он назвал «аналитической». В единую логическую схему Беббидж увязал арифметическое устройство (названное им «мельницей»), регистры памяти, объединенные в единое целое («склад»), и устройство ввода/вывода, реализованное с помощью перфокарт трех типов. По мысли Дотцлера [Дотцлер, 2010] целью математика было описать сложные механизмы с помощью «буквальных и семиотических аббревиатур», которые можно будет легко конструировать. Практическая реализация машины привела его к идее создания специального кода для аналитической машины. Ну и конечно же, в этой связи нельзя забывать об Аде Лавлейс, сделавшей первое описание ЦВМ и правил ее программирования.

Чуть позднее, в 1832 году российским изобретателем Корсаковым С.Н. была предложена идея использования перфокарт для выполнения рутинных операций при статистической обработке данных. Предложенная машина для работы со «сложными идеями» предназначалась, в частности, для быстрого поиска гомеопатических лекарств в зависимости от множества симптомов.

Машина представляла собой пять механизмов: три гомеоскопа – устройств исследования подобий в идеях, идеоскоп – прибор анализа сложной идеи и построения общей идеи на основании более простых, компаратор – прибор оперативного сравнения идей по признакам [Алексеев, 2019].

Но история этого периода не ограничивается только историей построения тех или иных машин. В 1872 Иоганн Готлоб Фреге расширил булеву логику включив в нее новые объекты и отношения. В работе «Смысл и значение» им вводится семантическая модель, связывающая такие понятия как «знак», «значение» и «смысл», что является принципиально важным для понимания связи какой-либо языковой или формально-логической системы и реальности. Кроме того, математик ввел такие понятия функция и класс, а также два абстрактных понятия – «истина» и «ложь» как аргументов логических функций.

В 1895 году, после создания под руководством Эрнста Маха кафедры философии индуктивных наук, возникает первый «Венский кружок». Позднее, в 1922 году, уже под руководством Морица Шлика возникает второй, ставший уже ядром логического позитивизма.

Согласно основной доктрине логического позитивизма все знания могут быть охарактеризованы с помощью логических теорий. В основе этих идей легли работы «Логико-философский трактат» Людвиг Витгенштейна и «Principia mathematica» Бертрана Рассела. Среди участников кружка были такие известные личности как Рудольф Карнап, Герберт Фейгель, Альфред Уайтхед, Курт Гёдель. Позднее к участникам кружка примыкают Джон фон Нейман и Френк Рамсей.

Под впечатлением дискуссий Венского кружка Карнап пишет свою первую крупную работу «Логическая структура мира» как попытку объединить логику и эмпиризм. По выражению Силен и Мейсмана [Силен, 2017] эта работа стала первой теорией мышления как вычислительного процесса. Необходимо отметить, что поскольку в рамках логического позитивизма основная часть метафизики осталась «за бортом» (в чем,

собственно, и состояло одно из намерений создателей данного направления) логический позитивизм в некоторых кругах встретил неодобрительное отношение.

В течение 30-х и 40-х годов возникает целый ряд ярких работ, так и или иначе приведших к созданию физических интеллектуальных систем, представленных сегодня компьютерными системами в привычном сегодня понимании. Впрочем, по нашему мнению, все началось, немного ранее - с 1900 года - когда Давид Гильберт на Парижском Конгрессе Математиков в Париже поставил проблему отыскания универсальной алгоритмической процедуры. В некотором смысле именно этот момент можно считать началом истории символьных методов.

7 сентября 1930 года Курт Гёдель на проходившем в Кенигсберге научном конгрессе обнародовал две теоремы о неполноте. Автор доклада показал, что существует эффективная процедура доказательства любого истинного высказывания в логике Рассела и Фреге, но при этом она не позволяет выразить принцип математической индукции в отношении мира натуральных чисел [Рассел, Норвиг, 2016, с.44]. Иными словами, Гёдель показал, что существуют истинные высказывания, которые являются недоказуемыми.

Таким образом, к середине 1930-х годов все было готово к возникновению первых электронных вычислительных систем.

§ «Компьютерный период»

В 1934 выходит работа Карла Поппера «Логика научного исследования» [Popper, 1980], в которой обосновывает идею фальсифицируемости научного знания.

В 1936 теорема Гёделя была переформулирована Аланом Тьюрингом [Turing, 1938]. Тьюринг заменяет универсальный формальный язык Гёделя на гипотетическое устройство, впоследствии названное «Машиной Тьюринга».

Пожалуй, именно эту статью можно было бы считать началом компьютерной эпохи.

Чуть позже, в 1936 году, Алонзо Чёрч создает λ -исчисление, давшее теоретическую основу процедурно-функциональной технике программирования. В основе λ -исчисления лежит понятие функции, а основная идея заключалась в том, чтобы исследовать возможность дать ответ о том верно или нет некоторое логическое высказывание. Исследование должно проводиться с помощью какого-либо алгоритма, построенного на основе заданных аксиом. Несмотря на то, что задача не имеет решения в общем случае, именно λ -исчисление легло в основу последующих работ Алана Тьюринга.

В 1936 в статье «On Computable Numbers, with an Application to the Entscheidungsproblem» А. Тьюринг впервые описал автомат, получивший название «машина Тьюринга».

В 1938 году К. Цузе создает механический вычислитель с электрическим приводом и ограниченной возможностью программирования. Этот вычислитель Цузе по праву можно считать первой электронной вычислительной машиной.

В 1943 году Уоррен Мак-Каллок и Уолтер Питс [McCulloch, 1943] предлагают идею искусственной нейронной сети.

Альфред Тарский создает теорию ссылок, впервые показав, как связать логические объекты с объектами реального мира [Рассел, Норвиг, 2016, с.43].

Таким образом к началу третьего, «компьютерного» периода, уже сформированы все основные вопросы [там же] необходимые для создания интеллектуальных физических систем:

1. Каковы формальные правила формирования правильных заключений?
2. Как определить границы вычислимости?
3. Как проверить рассуждения с использованием недостоверной информации?

Даны если и не исчерпывающие, то достаточно полные ответы на эти вопросы, имеющие первостепенное значение для создания Искусственного интеллекта и реализации символьного подхода, в частности.

§ Символьные методы и исследовательские программы

ИИ

В.К. Финн так классифицирует состояния исследований в области искусственного интеллекта [Финн, 2011, сс. 250–251]:

1. *процедурный*, когда создавались некоторые процедуры, делались, скажем, шахматные или шашечные программы, интеллектуальная деятельность отождествлялась с этими процедурами;
2. *декларативный*, когда стали пытаться путем систем утверждений описывать некоторую предметную область, т.е. создавать представления наших знаний о мире, стали формироваться декларативно-процедурные средства обработки знаний.

Само по себе понятие символьного подхода базируется на двух концепциях:

Во-первых, понятия символьных методов и систем компьютерной алгебры [Семенов, 2004, с. 128]. В отличие от численных методов, занимающихся непосредственным образом обработкой вычислений (как получение результата в виде числа или набора чисел), символьные методы оперируют равенствами и формулами как последовательностями символов. Примером могут служить, например, правила для упрощения выражений, подстановки одних выражений в другие.

Во-вторых, гипотезе Ньюэлла-Саймона [Newell, Simon, 1976]. Исходя из тезисов гипотезы физическая символическая система имеет все необходимое

для решения задач, аналогично тому, как делает человеческий мозг. В 1985 году, с легкой руки Джона Хогланда [Haugeland,1985], вся совокупность решений, основанных на этой гипотезе, получает название «старый добрый искусственный интеллект» или GOF AI («Good Old-Fashioned Artificial Intelligence»).

В соответствии с тезисом Черча-Тьюринга, всякая вычислимая функция вычислима с помощью машины Тьюринга. В терминах Клини [1973, с.430] это звучит как «все частичные (то есть не обязательно определённые для всех значений аргументов) функции, вычисляемые посредством алгоритмов, являются частично рекурсивными». В терминах Тьюринга это означает, что функция вычислима с помощью машины Тьюринга, путем частично рекурсивной последовательности элементарных операций. Тезис Черча – Тьюринга является всего лишь математическим определением, *гипотезой*.

Можно спорить о том какую из дат считать началом того феномена, который сегодня объединяют под термином Искусственный интеллект. Такой датой можно считать, например, лето 1956 года, когда его ввел в повсеместное использование Джон МакКартни на исторической конференции в Дартмутском университете. На наш взгляд, более правильным было бы считать отправной точкой всех исследовательских программ в области ИИ статью Алана Тьюринга «Вычислительные машины и разум», опубликованную в 1950 году в журнале «Mind» [Turing, 1950].

Дискуссии вокруг данной статьи и сегодня составляют огромный пласт разного рода статей и целых монографий. С точки зрения истории возникновения и развития символьных методов данная статья примечательна сразу несколькими фактами:

Во-первых, факт того, что машина Тьюринга представлена как первое символьное устройство, система, основанная на правилах вывода.

Во-вторых, факт того, что он исключил из участников теста машины, пусть и механические, но работающие на основе метода проб и ошибок по

причине того, что для правильной интерпретации результатов необходима объяснимость результатов с точки зрения рационального мышления.

В-третьих, выделение важности Тьюрингом проблемы творчества со стороны искусственной интеллектуальной системы, о чем он пишет в главе VII «Обучающиеся машины». В частности, Тьюринг делает вывод о том, что утверждение Лавлейс о невозможности творчества со стороны машины является спорным. Машина, по его мнению, теоретически сможет удивить нас новыми идеями в том случае если машина является обучающейся. Причиной этому, по его мнению, будет являться тот факт, что система определений и суждений будет эволюционировать если не случайным, то сильно зависящим от контекста способом. Форд, Глимур и Хейз [Parsing the Turing Test, 2008, p.11] в своих комментариях к статье Тьюринга говорят о проглядывающейся за этим аналогией с экспертными системами.

В-четвертых, приводимый Тьюрингом пример обучения «ребенка-машины» является реализацией концепции именно символического подхода, в котором - говоря в терминах ДСМ-метода Финна В.К. [Финн, 2011] - база знаний, т.е. правил вывода строится в значительной степени в соответствии с эволюционистскими принципами. Руководящие принципы выбора наилучшего правила должны здесь задаваться учителем. На что еще стоит обратить особое внимание — это обязательность объяснимости как процесса обучения, так и получаемых машиной выводов в ходе ее работы. По мысли Тьюринга «интеллектуальное поведение» вполне допускает отклонение от абсолютной детерминированности поведения, но только в незначительной степени. В противном случае поведение машины может стать беспорядочным.

Комментируя статью Тьюринга, Поппер К. замечает [Popper, 1950], что Тьюринг, по сути, ставит интеллектуальную ловушку. Описывая форму поведения, мы, тем самым, уже проектируем новую вычислительную программу.

С этого момента в мире возникает бум разного рода исследовательских программ. Например, как вспоминает А. Эттингер, автор первого машинного словаря (1954):

«представление о полностью автоматизированном машинном переводе...взлелеянное теми, кто в этом заинтересован...разрослось как буйный сорняк» [Дрейфус, 2010, с.33].

В том же году выходит статья Клода Шеннона о машинах, играющих в шахматы, в которой обсуждалась принципиальная возможность создания такой программы. В ней он описывает программу, которая «будет играть довольно сильно и по скорости сравнимо с человеком» [Shannon, 1950].

По мнению Дрейфуса основной толчок к большинству последующих исследовательских программ дали работы в корпорации RAND Алана Ньюэлла, Джорджа Шоу, и Герберта Саймона [Дрейфус, 2010, с.35]. В 1958 году появляется разработанная ими программа для игры в шахматы. Спустя несколько лет авторы называли свой подход «моделированием процесса познания». Дрейфус в несколько ироничной форме отзывается по поводу энтузиазма Саймона в связи с шахматной программой и восторга Д. Митчи в октябрьском Science Journal 1968 [Дрейфус,2010, с. 27].

В это же время возникает сразу несколько программ-проектов, которые как представлялось тогда, должны были если не решить, то по крайней мере практически приблизится к решению проблемы Искусственного интеллекта. В большинстве случаев в основе всех этих программ лежали методы символьного подхода. Дрейфус пишет о четырех основных направлениях в исследованиях того периода [Дрейфус, 2010, с. 28]: программирование игр, автоматический перевод, решение задач и распознавание образов.

В 1954 Энтони Эттингер создает первый машинный словарь. Несмотря на то, что на первых порах работа над словарем шла успешно, Эттингер обращает внимание на «чрезвычайно загадочные семантические процессы, благодаря которым большинство людей в своих рассуждениях

интерпретирует большинство осмысленных предложений в большинстве случаев однозначно». [Дрейфус, 2010, с. 52].

В это же время начинается разработка специализированных языков программирования. Примером такого языка является IPL (Information Processing Language) от корпорации RAND. В частности, широкое применение получила версия языка IPL-V. Дальнейшим развитием этого языка стал разработанный Джоном МакКарти язык LISP (LISt Processing) для компьютеров IBM 704, версия LISP 1.5. В основе работы языка LISP лежит логика λ -исчисления Алонзо Черча. Преимуществом данного языка были высокоуровневая реализация и возможность эффективной работы с деревьями, обеспечивая поиск оптимальных стратегий перемещаясь по дереву решений. Позднее возникают COBOL (1959) (от COmmon Business-Oriented Language). Руководителем проекта была Грейс Хоппер. Язык PROLOG был языком логического программирования, направленным прежде всего на решение задач из области математической логики.

В сентябре 1956 года проходит, по выражению Мильнера [Millner,2003], «звездная» конференция в Массачусетском технологическом институте, в рамках которой произошло сразу несколько значительных событий:

1. Представлена завершенная версия программы «Логик-теоретик» Ньюэлла и Саймона.
2. Ноам Хомский представил свои идеи по трансформационной грамматике.
3. Джордж Миллер сделал доклад об ограничениях в кратковременной памяти.
4. Джон Светс представил возможности теории обнаружения сигналов для перцептивного распознавания образов и т.д..

Миллер назвал этот день «моментом зачатия» когнитивной науки [там же]. Сложно переоценить влияние программы «Логик-теоретик» и

трансформационной грамматики Нома Хомского на ход дальнейших исследований.

Гипотеза Ньюэлла-Саймона

Авторы гипотезы, анализируя процесс решения логических задач студентами, обратили внимание на то, что их испытуемые при решении разного рода задач пользуются некоторым общим набором правил, который можно было бы представить как «общую теорию разумного поведения». Эти приемы, не являясь универсальными, в ряде случаев, хотя и не всегда, приводили к решению. Примером такого правила может быть: «всегда старайся длинное выражение заменить на короткое». Ньюэлл и Саймон решили использовать эти стратегии. А созданные в соответствии с этими стратегиями программы они назвали «эвристическими», подчеркивая тем самым их отличие от алгоритмических программ. [Дрейфус, 2010, с.16] Результатом их работы стала статья [Newell, Simon, 1956], в которой они указывают, что целью исследования было понять как человек, например математик, может доказать ту или иную теорему, без применения методов, требующих больших вычислительных мощностей.

Основными принципами, заложенными в программу «Логик-теоретик были следующие» [Gugerty, 2006]:

1. Мышление представляет собой процесс оперирования абстрактными символами в кратковременной и долговременной памяти. Процесс изолирован от сенсорной информации извне.
2. Символы представляются частично как носители информации об объектах и событиях реального мира, но главным образом как результат взаимодействия с другими информационными процессами, таким образом чтобы управлять поведением всего организма.
3. Знания имеют иерархическую структуру.
4. Сложные проблемы решаются посредством эвристик, действительно эффективных, но не дающих гарантии решения.

5. «Логик-теоретик» работает в обратном порядке от теорем, которые должны быть доказаны, с использованием эвристик для поиска достоверных выводов до тех пор, пока не будет достигнут уровень аксиом.

Организацию информационного процесса Ньюэлл и Саймон представили как трехуровневую иерархию:

- a) «Инструкции» нижнего уровня, представляющие из себя элементарные операции.
- b) «Элементарные процессы», представляющие собой последовательно выполняющийся список инструкций.
- c) «Методы», представляющие собой последовательно выполняющийся список элементарных процессов.

Основных методов было предложено четыре:

1. «Замена», осуществляющая поиск и замену одних выражений другими.
2. «Отделение», метод, основанный на правиле логического вывода по принципу «modus ponens».
3. «Цепочка вперед», метод, реализовывавший транзитивные правила вывода: Если А есть В, В есть С, то А есть С.
4. «Цепочка назад», метод, действующий таким же образом, как и цепочка вперед, но в обратную сторону как попытка доказать, что А есть С, если В есть С, и А есть В.

Необходимо отметить, что Ньюэлл и Саймон подчеркивали тот факт, что их программа не моделирует работу человеческого мозга и не была попыткой воспроизвести его деятельность. Влияние программы оказалось огромным. Как пишут в своей книге Миллер, Гелантер и Прибрам «Планы и структура поведения» работа Ньюэлла и Саймона стала отправной точкой в процессе

возникновения когнитивной психологии [Милнер, 1965]. Кроме того, исследователи в своей работе предсказали и дали ответ на некоторые критические замечания, которые остаются актуальными и по сей день:

Во-первых, речь идет о проблеме сложности при симуляции когнитивных возможностей, ввиду наличия большого количества параметров, которые необходимо тем или иным образом квалифицировать, а также сложности самой по себе модели поведения.

Во-вторых, о проблеме «гомункулуса». Суть данной проблемы в том, что моделирование когнитивных способностей человека должно постулировать некоторые интеллектуальные, но, даже на сегодняшний день, плохо понимаемые ментальные процессы.

В-третьих, речь идет о том, каким образом проверить адекватность получаемой симуляции. В частности, Милнер [там же] основывается на возможности проверки адекватности прежде всего путем вербальной коммуникации.

Результатом этой и последующих работ Ньюэлла и Саймона стала выдвинутая ими позднее гипотеза, согласно которой утверждается, что физическая символьная система имеет достаточные и необходимые средства для решения основных интеллектуальных операций [Newell, Simon, 1976].

По нашему мнению, эта гипотеза не является чем-то совершенно новым, было бы правильнее считать это дальнейшим развитием идей Тьюринга и Шеннона, получивших первое инженерное подтверждение.

Основанием для выдвижения такого предположения явились результаты следующей после «Логика-теоретика» работы. Речь идет о компьютерной программе - [GPS] General Problem Solver (буквально – «Общий решатель задач»). GPS строилась на базе «Логика-теоретика». Как пишет Дрейфус [Дрейфус, 2010, с.18] в то время казалось, что философия наконец-то получила необходимые технические средства. Рагнар Фьеланд замечает, что такого рода надежды подогревались в том числе высказываниями аналогичными тем, что сделал Рассел, утверждая, что причинность должна уйти в прошлое и всюду

начнет править статистика [Fjelland, 2020]. Но действительность оказалась не такой радужной как того хотелось бы. Дальнейшая история подтвердила данный тезис.

Дрейфус [Дрейфус, 2009] отмечает несколько неудачных, или по крайней мере, имеющих завышенную оценку со стороны того же Маркуса Минского, примеров создания GPS-программ в 60-х годах прошлого века. Кроме того, Дрейфус критикует Минского за чрезмерное увлечение «эвристическими методами», в частности за допущение, что кодирование текста происходит с помощью эвристических методов. Действует ли человек по эвристическим программам или нет остается большим вопросом по его мнению. Приняв такое допущение, с необходимостью требуется предположить, что идеальная программа должна строиться по принципу от частей к целому. [Дрейфус, 2009, с. 94]

Далеко не все разделяли пессимизм Дрейфуса. Одним из тех, кто не соглашался или по крайней мере продолжал выражать наибольший оптимизм был Марвин Минский. В своей лекции по случаю присуждения Премии Тьюринга в 1969 году Минский продолжает утверждать, что для построения модели обучения достаточно математического аппарата. То, как функционирует организм физически не имеет значения [Minsky, 1970].

Трансформационная грамматика Хомского

Возвращаясь к знаменитой конференции 1956 года в MIT, обратим внимание доклад Ноама Хомского по трансформационной грамматике.

Центральной проблемой лингвистической теории Хомский считает удивительный факт несоответствия между языковыми знаниями, имеющимися в уме рядового говорящего, и теми скудными данными, которые были в его распоряжении, когда он осваивал родной язык [Хомский, 2005, с.7]. Из чего следует, по мнению Хомского, что должны существовать некоторые правила, которые могут точно определять какие выражения допустимы в языке, а какие нет. По мнению Сёрля это была *хомскианская революция XX*

века. Предложенная Хомским идея изучения языка путем создания формальных моделей повлияла не только на науки о языкознании, но и вселила надежду в исследователей в области ИИ. Ноам Хомский и последователи его идей в попытке ответить на этот вопрос исходят из предположения о рекурсивных механизмах, о том, что высказывание создается именно посредством использования применительно к уже известным выражениям конечного числа врожденных правил. Эти закодированные в мозгу правила, их реализация и есть функция мозга [Корниенко, 2018, с.98].

С точки зрения истории возникновения и развития символьных методов сложно переоценить значение идей Хомского. В то время казалось, что трансформационная грамматика открывала прямой путь к созданию программ-переводчиков, диалоговых программ и т.д.

Доклад Лайтхилла, «зима ИИ»

В целом, ситуация, которая складывалась в конце 50-х начале 60-х годов, полная в начале всеобщего энтузиазма, к середине 60-х стала становится все более и более пессимистичной. Перечислять все исследовательские программы нет необходимости ввиду того, что ни одна из них не достигла каких-либо значимых результатов.

Начало 1970-х можно охарактеризовать как «зима искусственного интеллекта». В немалой степени этому послужил отчет Лайтхилла [Lighthill, 1973] сделанный по запросу Британского совета по научным и инженерным исследованиям в связи с разногласиями между учеными Эдинбургского университета и Суссекса. Исследование возникло в связи с желанием со стороны Эдинбурга приобрести новую машину DEC10 американского производства. Результатом отчета стало резкое сокращение финансирования вплоть до обзорного совещания по интеллектуальным системам, основанным на знаниях, организованного SERC, которое прошло в сентябре 1982 года.

В целом доклад нельзя считать полным разгромом в области исследований Искусственного интеллекта: критике подверглись такие направления как машинный перевод и робототехника, при том, что исследования в области нейрофизиологии и психологии получили положительную оценку со стороны Лайтхилла. Одной из основных проблем, которые по мнению Лайтхилла препятствуют дальнейшему развитию - является проблема «комбинаторного взрыва». Суть проблемы заключается в резком, «взрывном» росте времени выполнения того или иного алгоритма при увеличении объема данных.

Эпоха LISP и экспертных систем

Несмотря на рост пессимизма в области Искусственного интеллекта несколько направлений все же продолжали свое развитие. Во-первых, речь идет о развитии языка LISP и возникновением LISP-машин. Во-вторых, росте интереса к началу 80-х годов к экспертным системам.

Годом начала истории LISP-машины можно считать 1973 года, когда в Массачусетском институте стартовала программа по созданию 24-битной архитектуры, оптимизированной для LISP. Позднее в Массачусетсе Расселом Нофтскером была основана компания «Symbolics Inc», занимающаяся разработкой LISP-машин и специализированного программного обеспечения. В 1982 году создается операционная система Open Genera. В 1988 выходит 40-битный «Ivory», первым представителем которого был XL400.

Одним из направлений использования языка LISP были попытки его применения для решения задач компьютерного синтеза программного кода [Koza, 1992] и использования языка для реализации схем генетических алгоритмов [Родзин, 2002]. К сожалению, ряд ограничений языка не позволили ему закрепиться в этом сегменте.

Другим активно развивающимся несмотря на все трудности направлением были экспертные системы и связанные с ними задачи машинного обучения, основанного на символьном подходе.

Символьное или символическое машинное обучение представляет собой ответвление инженерии знаний, одной из школ ИИ. В 1970-х у так называемых систем на основе знаний были достигнуты очень впечатляющие успехи. В 1980-х они быстро распространились, но потом быстро сошли на нет. Главная причина — печально известное «узкое горло» приобретения знаний: получать информацию от экспертов и кодировать в виде правил слишком сложное, трудоемкое и подверженное ошибкам занятие, поэтому для большинства проблем такой подход нецелесообразен. Оказалось, что намного легче позволить компьютеру автоматически учиться [Домингос, 2016].

Как правило системы, построенные на базе символьных методов, строятся на двух взаимодополняющих концепциях: деревья решений и наборы правил. А основной идеей обучения в рамках таких систем лежит принцип обратной дедукции [Домингос, 2016]. В начале 60-х годов специалисты в области ИИ пытались моделировать сложные процессы мышления, отыскать общие методы решения задач в общем виде. Но ряд проблем, связанных со сложностью и проблема комбинаторного взрыва, в частности, привели к пониманию того, что чем шире класс задач, которые может решить одна программа, тем беднее оказываются ее возможности при решении конкретной проблемы [Уотерман, 1989, с.13].

К 70-м годам пришло понимание, что должен быть другой способ сделать программу «умной». Основные усилия были переключены на разработку методов представления (т. е. способах сформулировать проблему так, чтобы ее легко было решить), и методов поиска (т.е. на способах управления ходом решения). Но и этот подход не принес значительных успехов.

Понимание того, что эффективность программной реализации зависит не только от формализмов, но прежде всего от знаний, которыми она обладает стало приходить только в конце 70-х годов специалисты, работающие в области ИИ, начали понимать нечто. Результатом этого было принятие новой

концепции, которую можно сформулировать следующим образом [Уотерман, 1989, с.14]:

«Чтобы сделать программу интеллектуальной, ее нужно снабдить множеством высококачественных специальных знаний о некоторой предметной области».

Началом широкомасштабных исследований в области экспертных систем можно считать Международную объединенную конференцию по искусственному интеллекту 1977 года, когда Э. Фейгенбаум высказал основополагающее соображения об экспертных системах, суть которых состояла в следующем [Хейс-Рот, 1987, с. 17]:

1. Большие возможности экспертной системы определяются теми знаниями, которыми она располагает, а не конкретными формализмами и схемами вывода.
2. Знания эксперта являются решающим для высококвалифицированной деятельности.

Понимание этого факта к концу 1980-х и привело к возникновению экспертных систем как совокупность методов и приемов, каждый экземпляр которых служит для решения проблем в некоторых узкоспециализированных областях. Знания, относящиеся к любой специальности, обычно существуют в виде общедоступных и индивидуальных:

1. Общедоступные знания – это факты, определения и теории, которые изложены в учебниках т.д.
2. Индивидуальные, или личные знания, в значительной степени состоят из эмпирических правил или эвристик [там же].

В качестве отступления необходимо отметить, что сама по себе правильность идеи создания именно человеко-машинных систем неоднократно отстаивали многие исследователи, такие как Финн В.К. и Венда В.Ф. В частности, Вендой В.Ф. была выдвинута в 1975 году идея гибридного интеллекта [Венда, 2020].

Фредерик Хейес-Рот, Дональд Уотерман и Дуглас Ленат в своей совместной работе [Хайес-Рот, 1987] указывают на существование нескольких причин, по которым в области экспертных систем особую роль обретают знания, а не формальные методы рассуждения:

«Во-первых, для большинства трудных проблем нет четких алгоритмов решений, поскольку многие важные задачи возникают в сложных контекстах социальных или физических явлений.

Во-вторых, причина преимущественного внимания к знаниям, а не к формальным методам рассуждения, носит прагматический характер: люди-специалисты достигают исключительно высоких результатов благодаря своим познаниям.»

На наш взгляд, данный аргумент представляется довольно слабым, хотя сами авторы пишут, что «в короткой истории существования экспертных систем правильность этого положения неоднократно подтверждалась» [Хейс-Рот, 1987, с. 15].

В-третьих, признание ценности знаний как таковых. Знания - дефицитный ресурс, уточнение и воспроизводство знаний способствует росту благосостояния.

Экспертные системы, продолжают авторы, существенно отличаются как от обычных систем, так и от других систем искусственного интеллекта. От традиционных систем экспертные систем отличает прежде *всего символьное представление, символичный логический вывод и эвристический поиск*. В то же время именно эти свойства, которые могут быть обнаружены в той или иной

системе искусственного интеллекта, приводят, казалось бы, к тому, что «простую задачу из области искусственного интеллекта можно решить, прибегнув к одному из формальных подходов, развитых при решении таких фундаментальных вопросов» [Хейс-Рот, Уотерман, Ленат, 1987, с. 16]. Таким образом, в большинстве случаев допускается, что в целом ряде случаев с системами ИИ могут успешно конкурировать системы, построенные с использованием классических алгоритмов, понимаемых как слабый или узкий ИИ.

Однако, экспертные системы все же имеют ряд принципиальных отличий от других систем искусственного интеллекта. А именно [там же]:

1. Выполняют задачи на уровне хорошего специалиста.
2. В них отдается предпочтение проблемно-ориентированным стратегиям решения задач по сравнению с более общими «слабыми методами» искусственного интеллекта.
3. Они используют знания о самих себе для того, чтобы делать заключения о том, как в них протекают процессы вывода, и чтобы дать объяснения или дать оправдания достигнутым результатам.
4. Они решают задачи, которые обычно относят к одному из следующих классов: интерпретация, предсказание, диагностика, отладка, конструирование, планирование, мониторинг, ремонт, инструктирование и управление.

В качестве примеров реализации такого рода проектов авторы указывают стэнфордские проекты DENDRAL и META-DENDRAL. Первая анализирует данные магнитного ядерного резонанса, масс-спектрографические и другие химические экспериментальные данные с целью вывода вероятностных структур неизвестного химического соединения. В процессе решения в ней используется эффективный вариант метода порождения и проверки.

Вторая система, META-DENDRAL, позволяет системе DENDRAL анализировать знания, обеспечивая генерацию и последующий отбор правил сегментации для органических структур.

Система PROSPECTOR открыла залежи молибдена, ценность которых, возможно, превзойдет 100 миллионов долларов.

Система R1 помогала покупателям сформулировать свои требования к конфигурации вычислительной системы VAX.

Система CADUCEUS вмещала в себя больше знаний о болезнях, чем любой медик и могла точно ставить диагнозы в тестовых примерах.

Система PUFF обеспечивала экспертные анализы данных в целях диагностики легочных болезней в Калифорнийском медицинском центре.

Последующие 90-е годы привели к значительному разочарованию в возможностях таких систем. Масла в огонь добавил вновь возникший интерес к нейронным сетям, обусловленный резким удешевлением мультипроцессорных систем и резким ростом производительности графических карт.

Еще одним интересным фактом в истории ИИ, прежде всего символьных методов, является ДСМ-метод, предложенный В.К. Финном в 1970-х годах XX века. Целью предлагаемого метода автоматического порождения гипотез является извлечение причинно-следственных зависимостей из базы фактов и использование базы знаний посредством формальной эвристики вида «индукция-аналогия-абдукция». Эта эвристика формализуется в виде правдоподобных рассуждений (точнее, когнитивных правдоподобных рассуждений), а ее компьютерной реализацией является система ДСМ (ИС-ДСМ) [Финн, 2011, с. 381]. Предлагаемый метод состоит из пяти компонент [там же]:

1. Формализованные условия применимости метода на основании казуальных моделей Аншакова и Финна.
2. Правдоподобные ДСМ-рассуждения, реализующие синтез познавательных структур - индукции, аналогии и абдукции.

Существенной особенностью ДСМ-метода является то, что ни индукция, ни аналогия, ни абдукция не рассматриваются изолированно. *При этом индукция не является перечислительной и содержит средства автоматической фальсификации.*

3. Квазиаксиоматические теории как средство организации базы знаний и базы фактов (КАТ). КАТ есть средство формализации открытых теорий различных предметных областей, открытого множества фактов и гипотез, а также открытого множества правил вывода, содержащего правила правдоподобного и правила достоверного вывода. Существенной особенностью КАТ является, с одной стороны, возможность расширения как множества аксиом, так и множеств фактов и гипотез. С другой стороны, особенностью КАТ является применение правдоподобных ДСМ-рассуждений как средства автоматического анализа массивов фактов и средства порождения гипотез о зависимостях причинно-следственного типа.
4. Дедуктивная имитация ДСМ-рассуждений, формализующих процесс автоматического порождения гипотез. Правила правдоподобного вывода, формализующие индукцию и аналогю, представлены в виде аксиом, используемых в качестве системы дедуктивной имитации ДСМ-рассуждений.
5. Собственно, сама информационная ДСМ-система. Система состоит из четырех подсистем: Решатель задач, База фактов, База знаний, Интерфейс пользователя.

Примерами использования ДСМ-метода являются:

1. Клинические исследования воздействия лекарственных препаратов на больных с различными индивидуальными особенностями [Финн В.К., 2011, с.89].

2. Задачи пополнения семантического словаря для использования в информационных технологиях, связанных с распознаванием речи [Рубашкин, 1998].
3. Задачи экономического прогнозирования [Гаджиев, 2015].
4. Задачи качественных оценок в социологических исследованиях [Климова, 2015] и ряде других исследований.

В настоящее время данное направление продолжает развиваться в ВИНТИ РАН и на кафедре математики, логики и интеллектуальных систем РГГУ под руководством самого автора.

Таким образом, возникнув в качестве самого первого метода ИИ, символьные методы претерпели значительные эволюционные изменения пройдя путь от декларативных методов, предназначенных для решения шахматных задач до большого класса задач в области экспертных систем. К сожалению, несколько кризисов, начала 70-х годов, конца 90-х, сильно повлияли на исследования в этой области. Нерешенные проблемы сложности обучения и комбинаторного взрыва привели к практически полному смещению фокуса исследовательских программ в область нейронных сетей.

Интерес к символьным методам продолжает поддерживаться относительно небольшим количеством исследователей. В частности, самим Виктором Константиновичем Финном.

На наш взгляд, ослабление интереса к символьным методам носит во многом конъюнктурный характер, связанный с доступностью финансирования исследований. В результате остается нерешенным ряд задач, и в первую очередь - проблема обучения.

§ Проблемы современного дискурса

На наш взгляд главными тенденциями, определяющими сегодня направление современного дискурса, являются две темы.

Во-первых, речь идет о понимании общего ИИ с точки зрения прохождения той или иной версии Теста Тьюринга.

Во-вторых, вокруг перспектив реализации общего ИИ с использованием одних только коннекционистских методов.

Сегодня существует большое количество возникающих и прекративших свое существование конкурсов, посвященных прохождению Теста Тьюринга, таких как премия Лёбнера (прекратила свое существование), премия Хаттера², проводимая в нашей стране номинация «Nanotalk»³ и т.д.

Большой обзор как истории попыток так перспектив прохождения теста Тьюринга можно увидеть в целом ряде работ, таких как «Parsing the Turing Test» [2008], посвященной 55-летию юбилею теста.

Именно в отношении этого теста Джон Сёрль дает свое понимание термина «сильный ИИ». Основным критерием, с его точки зрения, который мог бы классифицировать ИИ-систему как «сильную», является способность последней пройти тест Тьюинга. Сам философ относится довольно критически к такой постановке вопроса [Parsing the Turing Test, 2008, pp. 139-151]. В настоящее время существует несколько десятков теста Тьюринга. В качестве примеров можно привести «signal-тест» (MIST) Криса Мак-Кинстри⁴, обратный тест Тьюринга или Completely Automated Public Turing Test to Tell Computers and Humans Apart [Blum,2000] и т.д.

Одним из главных недостатков этого теста, на наш взгляд, является тот факт, что он направлен не на выявление наличия разумного характера поведения ИИ-системы, а на успешность поддержания требуемой иллюзии общения с живым человеком у членов жюри в течение определенного времени. Таким образом, тест предполагает реализацию не только и не столько рационального поведения, но прежде всего воспроизведение иррациональной стороны поведения. Примером тому может быть анализ победителей конкурса Лебнера, где, например, наличие ошибок специальным образом закладывалось

² <http://prize.hutter1.net/>

³ <https://nanosemantics.timepad.ru/event/1579161/>

⁴ <http://hps.elte.hu/~gk/Loebner/kcm9512.htm>

при реализации конкурсных программ [Parsing the Turing Test, 2008]. Такая постановка вопроса только уводит в сторону от задачи по созданию систем общего ИИ.

На наш взгляд одним из основных критиков Теста является сам Тьюринг. В частности, речь идет о 4-м пункте возражений относительно сознания. В качестве такой точки зрения Тьюринг приводит слова профессора Джефферсона о возможности написания машиной сонета, основанной не за счет «случайного совпадения символов» [Turing, 1950], а в результате собственных побуждений. По мнению математика, такое утверждение означает, что удостоверится в том, что машина мыслит можно только став машиной и осознав процесс собственного мышления. Но это невозможно. В качестве основной линии защиты Тьюринг выбирает тезис о крайнем солипсизме относительно самих людей: или мы сомневаемся в том, что кто-то может мыслить, или мы должны прийти к конвенциональному соглашению, «вежливо» соглашаясь, что наш оппонент тоже мыслит. Автор теста предлагает просто уйти от данной проблемы. Но, по нашему мнению, поступая таким образом, Тьюринг только дает почву для разного рода спекуляций и, по сути, ничем не обоснованных ожиданий. Прежде всего со стороны инженеров-разработчиков ИИ-систем. Сложившаяся сегодня ситуация, на наш взгляд, является подтверждением данного тезиса. Более подробно критический анализ Теста Тьюринга будет дан в Главе 3.

Даже беглый анализ списков публикаций и докладов основных мировых конференций в области ИИ говорит о недооценке роли философии в области проектирования ИИ-систем. Об этом наглядно свидетельствует даже беглый анализ докладов как зарубежных, так и российских конференций таких как «Annual AGI Conference»⁵, «opentalks.ai»⁶ и т.д. В качестве характерных примеров можно привести обзорный доклад И.О. Пивоварова⁷ на конференции «opentalks.ai» 2022 года и недавнюю работу С.А. Шумского

⁵ <https://agi-conf.org/>

⁶ <https://opentalks.ai.ru/>

⁷ <https://www.youtube.com/watch?v=u066n1mFJO4>

«Машинный интеллект. Очерки по теории машинного обучения и искусственного интеллекта».

Так, в своем докладе Пивоваров вводит сразу несколько понятий, которые на наш взгляд, представляют собой редукцию к бихевиоризму и немецкому, гегелевскому, идеализму XIX века. Задача подробного анализа доклада выходит за рамки настоящего исследования, поэтому коснемся лишь нескольких примеров, подтверждающих тезис об элиминировании философии из современного дискурса в области разработки ИИ-систем.

Давая определение понятию «обучение», докладчик характеризует обучение как «не требующий интеллекта» процесс, направленный на накопление «информации о результатах своего поведения» с целью изменения поведения агента; вводит понятие «картина мира» как суммы накопленной интеллектуальным агентом информации.

Оценка качества обучения происходит исключительно в зависимости от того поменялось ли поведение агента в нужную относительно ожиданий обучающего сторону или нет. Интеллект определяется как [там же]:

«...способность к долгосрочному планированию действий – ставить цели и достигать их (в условиях неопределенности с ограниченными ресурсами)».

Трактуя процесс обучения в бихевиористском ключе, докладчик упускает из вида некоторые важные положения. Во-первых, представляя процесс обучения как процесс простого накопления информации, автор тем самым закрывает доступ к производству синтетического знания. Таим образом «забывая», что обучение непосредственно связано с творчеством и не всегда имеет строго формализуемые цели. Во-вторых, упускает из виду как саму проблему цели, о чем уже говорилось выше, так и «проблему фрейма», о которой пойдет речь далее в параграфе «Критика возможности выхода ИИ за границы техники».

Сама по себе критика бихевиоризма хорошо представлена в работах Джорджа Сёрля [Searle, 1980, Сёрль 2004], А.Р. Лурия [Лурия, 2021], Б.М. Величковского [Величковский, 2022], Хьюберта Дрейфуса [Дрейфус, 2007, 2010] и ряда других авторов. Принципиально важной, таким образом, представляется критика взглядов докладчика с точки зрения недостаточности бихевиоризма как не учитывающего, например, наличие «сложных форм отвлеченного мышления» [Лурия, 2021, с. 17]. Более того, в том случае, если бы автор доклада смог согласовать свои идеи с идеями «хайдеггеровского ИИ» [Dreyfus, 2007], то скорректированная таким образом исследовательская программа могла бы стать более эффективной и обоснованной с философской точки зрения.

Онтология понимается Пивоваровым также в упрощенном виде как «набор объектов и связей между ними». Но такое определение онтологии скорее подходит для понятия тезауруса. Основная проблема в данном случае кроется даже не в смешении таких понятий как «онтология», «мир», «онтика» и т.д.. Понимание онтологии как онтики, с одной стороны, в значительной степени «упрощает» математическое описание, а с другой - редуцирует модель до набора случайных фактов. Хотя автор доклада справедливо отмечает недостаточность такой модели, но причину видит в отсутствии интенциональности и субъективности, без которых невозможна оценка ценности как оценки отклонений на каждом шаге в процессе достижения цели. Выход им видится в расширении онтологии до «картины мира», которая дополняется субъективными оценками мыслящего агента. Критику такого рода гипотез мы можем найти, например, у Хайдеггера в работе 1962 года «Бытие и время» [Хайдеггер, 2015], которая позже стала основанием для возникновения «проблемы фрейма» [Dreyfus, 2007].

Кроме того, такое толкование онтологии скрывает от исследователей целый класс проблем, связанных с проблематикой бытия, его сущностных форм, принципов устройства и содержания как с точки зрения классической, так и неклассической философий; оставляет сокрытым от математического

осмысления возможные модели возникновения новых сущностей, анализ модусов существования сущего т.д..

Соглашаясь с С.А. Шумским, автор доклада в качестве одной из наиболее перспективных идей концепцию предиктивного кодирования (predictive processing), имеющую, кроме всего прочего, все недостатки присущие байесовским моделям.

Не менее примечательной с точки зрения игнорирования накопленного философского опыта представляет собой работа С.А. Шумского «Машинный интеллект. Очерки по теории машинного обучения и искусственного интеллекта». Подтверждение этому, мы можем увидеть уже на первых страницах, где он в качестве целевой аудитории перечисляет «как математиков и инженеров, так и нейрофизиологов, психологов и лингвистов» [Шумский, 2019, с.11]. Как можно видеть, философы не включены в этот список. Сам по себе такой факт может оказаться незначительным, но, на наш взгляд, является показательным с точки зрения защиты выдвигаемого нами тезиса.

Задача подробного критического анализа работы автора, является отдельной большой задачей и выходит за рамки настоящего исследования ввиду большого количества такого рода мест. Поэтому, остановимся на нескольких характерных моментах, подтверждающих выдвигаемый нами тезис.

Так, Шумский утверждает [Шумский, 2019, сс. 102-103]:

«Основное преимущество глубоких нейросетей – их способность порождать иерархию все более абстрактных представлений. Каждый следующий слой нейросети формирует все более и более информативные признаки входных данных, как наиболее значимые комбинации признаков предыдущего уровня. <...> Глубокое обучение дало универсальную методику автоматического конструирования абстракций...»

Но такая гипотеза о возможности использования потоков битов и нейронных импульсов в качестве символов, представляющих внешний мир уже выдвигалась Алленом Ньюэллом и Гербертом Саймоном [Newell,1976]. Безусловно, сама по себе, гипотеза Ньюэлла-Саймона, и основанные на ней идеи символьных методов являются, на наш взгляд, необходимыми (но недостаточными) составляющими общего ИИ. А одной из ключевых целей настоящего исследования является обоснование роли символьных методов как необходимых. В то же время, отдельно взятый тезис о возможности генезиса символов только базируясь на статистических методах, является, на наш взгляд недостаточным и противоречащим идеям Фреге, Пирса и ряда других философов.

Абстрактное понимается автором безо всякой связи с идеальным, исключительно в свете способности глубоких нейронных сетей «порождать иерархии все более абстрактных представлений данных» [Шумский, 2019, с. 102]. Нам наш взгляд, такое понимание представляет собой простое обобщение, так как не учитывает современное философское понимание абстракции⁸. В частности, не учитываются такие проблемы как индуктивный характер идеализации, необходимость оценки оснований для включения тех или иных свойств и т.д., без которых абстрактные понятия теряют всякий смысл.

Далее, автор вроде бы соглашается с мнением Выготского о том, что смыслы отражают ассоциации слов не друг с другом, а с «сенсомоторными впечатлениями и элементарными переживаниями» [Шумский, 2019, с. 253], но на практике утверждает, что обучение — это процесс «сжатия данных» [Шумский, 2019, с. 220], в следствие чего, например [Шумский, 2019, с. 262]:

«...имея доступ к большим текстовым массивам, он [ИИ] может выучить любой язык и освоить научный жаргон любой области

⁸ <https://iphlib.ru/library/collection/newphilenc/document/HASHd77fccf081b2386490ced7>

знаний, постоянно наращивая свою компетенцию по мере обучения.»

На наш взгляд, такого рода упрощения полностью элиминируют весь критический философский опыт феноменологии, семиотики и т.д., о важности которых пойдет речь в следующих главах.

Подводя итог, можно сделать вывод, что предлагаемые автором идеи практически полностью соответствуют идеям Гилберта Райла, выдвинутым им в его работе 1949 года «The concept of mind». При этом полностью игнорируется вся последующая критика этих идей.

В связи с чем трудно не согласится с высказыванием В.Г. Кузнецова о том, что «сегодня решили, что без философии можно прожить» [Кузнецов, 2014].

Все это приводит, в частности, к тому, что такие проблемы как определение роли тех или иных технологических подходов, перспективы их интеграции; определение факторов, ограничивающих процессы обучения и функционирования автономного ИИ, поиск онтологических оснований процессов семиозиса, а также ряд других «предельных» проблем уходят далеко на второй план.

Глава 2 Анализ онтологического статуса ИИ

Необходимость критического анализа основных технологических подходов обусловлена в первую очередь необходимостью понимания пределов возможностей, узких мест тех или иных подходов. Такое понимание является отправной точкой в понимании роли (множества ролей), которые те

или иные методы могут играть в общей для тех и других задачи построения общего ИИ.

§ Сущность техники, пределы и граничные условия функционирования технических объектов

Эрнст Капп.

Основоположник философии техники Эрнст Капп в опубликованной 1877 году работе «Основные направления философии техники» характеризует феномен техники не «просто» как продолжение организма, но и как рефлексию, как отражение вовне внутреннего мира человека. Создаваемый человеком искусственный мир, в его понимании, является продолжением самого человека, причем делается это бессознательно. Построение этого искусственного мира происходит посредством «органопроекции», через воспроизводство живого как такового в механических формах. В качестве примера приводится паровоз, как требующий питания, имеющий систему выделения продуктов жизнедеятельности, возможность остановки всех функций и смерть. По мнению В.Г. Горохова [Горохов, 2012], прослеживается «триумф органопроекции», начавшийся Каппом и продолжающийся до сегодняшнего дня. И в этом высказывании Горохова, так же, как и рада других исследователей, таких как Макс Айд [Дессауэр, 2016], видится явная критика такой идеи.

С другой стороны, по мнению, например, Фридриха Дессауэра и П.А. Флоренского идеи Каппа, несомненно, имеют определенные зерна истины. Более того, по мнению О.М. Седых и М.А. Хаменкова [Седых, 2016, с. 132].:

«...если в начале XX в. идея органопроекции не находит широкой поддержки, в последнее время она получает новое значение (компьютерная метафора) и современные аналоги (идея внешних расширений человека М. Маклюэна)...»

Тем самым, авторы подчеркивают, что в этом же русле движется и развитие философской антропологии.

К этому можно добавить, таких авторов как Мишель Серр, который, по нашему мнению, говоря о значении виртуализации мира сегодня, по сути расширил модель органопроекции уже человеческого мозга [Серр, 2016].

П.А. Флоренский

Говоря о Флоренском, прежде всего необходимо отметить, что философ считает первым, кто пришел к мысли об органопроекции, Герберта Спенсера, тогда как остальные, включая Каппа, либо подхватили эту мысль, либо к пришли к ней сами, но позже Спенсера [Седых, 2016, с. 134]. Флоренский выступает против механицизма XVIII века, который мыслит механизм первичным, а организм - подобием механизма [Терешкун, 2013, с. 51]. Мыслитель считал, что человек и мир (Среда) взаимоподобны [Седых, 2016, с. 136].:

«В Среде нет ничего такого, что хотя бы в зачатке не имелось у человека, так же как и у человека нет ничего, что пусть и разрозненно не нашлось бы у среды.»

При этом граница между искусственным и естественным не стирается, вектор от человека к миру остается в полной мере, человек лишь расширяет, проецирует и конструирует.

Отсюда же по мысли ряда исследователей, таких как О.Ф. Терешкун, можно сделать вывод [Терешкун, 2013, с. 52]:

«Сегодня человек не рассматривается как машина, а машина становится проекцией органов человека на природный материал».

И если с первой частью такого вывода нельзя согласиться в полной мере (критика в отношении данного тезиса будет приведена ниже), то со второй частью трудно не согласиться. Особенно в свете бурно развивающимися сегодня когнитивных наук и явного приоритета коннекционистских подходов в области проектирования ИИ.

Таким образом, с самых первых шагов философия техники *уже была подвержена* склонности к поискам подобия техники и живых организмов. Тем самым, порой неосознанно, философия способствовала «оживлению» технических объектов в нашем воображении, порождая, тем самым, разного рода ожидания появления некой новой сущности, обладающей (сверх)сознанием, способной к творчеству и изобретательству – ИИ.

П.К. Энгельмайер.

Несколько позднее, уже в 1898 году, П.К. Энгельмайер подвергает серьезной критике работы Каппа, говоря о том, что идеи «органопроекции» имеют весьма ограниченный характер. При этом он расширяет понимание техники как целенаправленной деятельности, имеющей определенную организацию; указывает на то, что техника, с одной стороны, облегчает удовлетворение потребностей, но при этом и способствует возникновению новых потребностей. В его понимании техника есть «искусство вызывать намеченные полезные явления природы, пользуясь известными нам свойствами природных тел» [Горохов, 1997, 64]. Вслед за Францем Рело, он ставит вопросы о различии между техническим и научным мировоззрением, о влиянии техники на культуру. На наш взгляд, в понимании Энгельмайера не техника сама по себе оказывает влияние на человека и окружающий мир. В его понимании именно культура является первичной, ибо только «те народы, которые выработали у себя технику, только они и «делают» историю, остальные же историю «претерпевают».» [Горохов, 1997, 71]. Техника *относится им к сфере материальной культуры*, как «та искусственная

ближайшая природа, которой цивилизованный человек себя окружил и изолировал от природы настоящей» [там же].

На наш взгляд, в таком ракурсе понимания техники явно прослеживается преобладание тенденции изучения не столько вопросов, связанных с онтологическим статусом техники, сколько отношений техники и культуры, иными словами, в направлении анализа проблем философской антропологии и социологии.

Освальд Шпенглер.

Отдельно стоит остановиться на работе Освальда Шпенглера «человек и техника» 1931 года. Это необходимо хотя бы уже потому, что он практически единственный кто выводил технику за пределы человеческого, предлагал «не поддаваться искушению видеть цель техники в создании машин и инструментов» [Культурология, 1995, с. 457]. Техника определяется им как «тактика всей жизни в целом. Она представляет собой внутреннюю форму способа борьбы, который равнозначен самой жизни» [там же]. Философ идет дальше, когда говорит, что техника без орудий присутствует повсюду и не является прерогативой только человека. В его понимании существует «техника льва, перехитрившего газель». Любая целенаправленная деятельность в его понимании уже есть техника. И все это – «стороны одной деятельной, борющейся, одушевленной жизни» [Культурология, 1995, с. 458]. Техника в его понимании является орудием войны против природы. И если техника животного представляет собой «видовую технику» или инстинкт [Культурология, 1995, с. 465], то техника человека независима от человеческого вида. Техника есть результат выхода человека за пределы «принуждения вида» [там же]. И в качестве первых технических объектов Шпенглер видит оружие.

В то время как оружие животных естественно, то орудия человека – искусственны. «Здесь начинается «искусство» как противоположность природы. Всякий технический метод представляет собой искусство»

[Культурология, 1995, с. 469]. На наш взгляд, философ всячески старается подчеркнуть факт противоестественности, противоположности техники [Культурология, 1995, с. 476]:

«Творческий человек выходит из союза с природой и с каждым своим творением он уходит от нее все дальше» [там же]. Появляется работа мысли и руки, различие *замысла* и *осуществления*. Эта возникшая форма многообразной человеческой жизни понимается им как «культура».

Вместе с тем происходит переход от органического к организованному существованию, от естественных к искусственным группам, «от стаи к народу, сословию, государству» [Культурология, 1995, с. 476]. С приходом машин человек сам хочет быть Богом, целью человека становится поиск неорганических сил природы, приспособив их вместо машин и животных [Культурология, 1995, с. 484]:

«С появлением рационализма «вера в технику» делается чуть ли не материалистической религией: техника вечна и непреходяща, подобно Богу-Отцу; она освобождает человечество, подобно Сыну; она просветляет нас как Дух Святой. А молимся мы на филистер прогресса...»

Таким образом, человек становится полностью увлеченным могуществом техники, божественное меняется в сознаниях людей на техническое. Уже отсюда становится видно, что сегодняшний восторг относительно ИИ является не чем-то новым, а скорее результатом трансформации тех идей и верований, которые возникли еще в начале XX века.

Вместе с тем он ставит проблему отчуждения труда, роста напряженности ввиду обезличенности процесса, непонимания тех операций, которые совершает рабочий:[Культурология, 1995, с. 486]:

«Властелин мира сделался рабом машины, Она принуждает его, нас, причем всех без исключения, ведаем мы об этом или нет, хотим или нет – идти по проложенному пути».

Об этом позже - об отчуждении труда - вновь заговорит в 1954 году Жильбер Симондон. А о порабощении человека - Мартин Хайдеггер, но уже отбросив ту значимость влияния вождей, о которой говорил Шпенглер.

Рост механизации мира по мнению Шпенглера приводит к тому, что «цель машин исчезает за их числом и утонченностью» [Культурология, 1995, с. 485].

Всматриваясь в день сегодняшний, невозможно не заметить тех же настроений, которые царят сегодня в отношении техники и, в первую очередь, ИИ. Кроме того, наш взгляд представляется отдельной и интересной темой научного исследования вопрос о том факте, что период подчинения техникой человека *длится уже как минимум последние 100 лет*. Не могло ли человечество за этот период свыкнуться с этой мыслью?

Хосе Ортега-и-Гассет

Интересным и ярким мыслителем является Хосе Ортега-и-Гассет. Несмотря на то, что его работа «Размышления о технике» вышла в далеком 1935 году, его идеи продолжают быть актуальными и сегодня. Актуальность их прежде всего определяется тем, что испанский философ, так же как и Хайдеггер, в большей степени уделял внимание именно поиску ответа на вопрос: «что есть техника?». Вопросы же об отношении человека и машины и существующих проблемах в этой области, как нам видится, у Ортега-и-Гассета являются скорее производными от того, что следует понимать под техникой.

Он определяет технику [Ортега-и-Гассет, 1935, с. 171]:

«как реакцию человека на природу или обстоятельства, в результате которой между природой, окружением, с одной стороны, и человеком – с другой, возникает некий посредник – сверхприрода, или новая природа, надстроенная над первой».

Сами по себе действия, направленные на удовлетворение потребностей еще не есть техника, так как возможность действовать присуща и животным. Отличие заключается в том, что животные способны к *ограниченному* набору актов. Человек же преобразует природу. Цель технических действий в преобразовании обстоятельств, в уменьшении роли случайности при достижении целей и устранению или «уничтожению» потребностей и усилий, с которыми связано их удовлетворение» [Ортега-и-Гассет, 1935, с. 172]. Именно создание новых, более благоприятных обстоятельств и составляет сущность техники, при помощи которой и возникает «сверхприрода».

Если животное подстраивается под среду, не пытаясь изменить сами исходные обстоятельства, то человек по своей природе осуществляет *бунт* против природы. Именно неудовлетворенность миром с одной стороны - является человеческим уделом, а с другой - причиной возникновения такого феномена как техника. Степень этого неудовлетворения идет дальше, в результате чего человеку недостаточно «просто» удовлетворить те или иные потребности, более того - человек охотно отказывается от них при необходимости. Но при этом человеку гораздо сложнее отказать себе в разного рода излишествах. Человеку необходимо благополучие, удобство, а не простое присутствие в мире.

Следовательно, *технику нельзя рассматривать и изучать в качестве самостоятельного феномена*. И причина этого в том, что [Ортега-и-Гассет, 1935, с. 176]:

«...человеческие потребности являются таковыми лишь в своей исключительной связи с благосостоянием. А это крайне осложняет

положение дел. Откуда нам знать, что именно человек понимал, понимает и будет понимать под благосостоянием?»

Следовательно, и какого-то одного вектора развития быть не может, тем более заранее известного вектора.

Для того, чтобы получить ответ на вопрос «что такое техника» необходимо, прежде всего, понять факт того, «как одно сущее (то есть человек, если он желает существовать) вынуждено пребывать в другом — в мире или природе» [Ортега-и-Гассет, 1935, с. 184]. Это существование есть абстрактная возможность, реальность которой еще необходимо завоевать. И причина этому в том, что человеческое и природное бытие человека не совпадают друг с другом. Человек представляет собой своего рода «онтологического кентавра» [Ортега-и-Гассет, 1935, с. 186], часть которого погружена в природу, а часть – «выходит за ее пределы, то есть трансцендентна» [там же]. Поэтому человек не считает природное истинным бытием. С другой стороны, то сверхестественное, что в нем есть - еще не реализовано. Оно *может* быть реализовано. Реализовано в каком-то жизненном проекте. Программе.

И здесь, казалось бы, можно построить первый мостик между ИИ и человеком. Ведь ИИ также представляет собой программу, которая выполняется машиной, заставляя ее функционировать, в том числе совершать мыслительные акты. Понимая человека как программу, сторонники сильного ИИ сразу должны были получить в лице Ортега-и-Гассета долгожданное подтверждение своих идей, присовокупив к этому тезис Тьюринга-Черча [Клини, 1973, с. 430], гипотезу Ньюэлла-Саймона [Newell and Simon, 1976] и ряд других гипотез.

Но следуя логике философа, такого рода выводы все же не будут верными, ведь [Ортега-и-Гассет 1935:187]:

«Перед нами удивительное существо, чье бытие состоит не в том, что уже есть, а в том, чего еще нет, иначе – сущее в том, что еще не существует... человек – не вещь, а некое усилие быть или тем, или иным»

Смысл всего остального заключается в том, что они уже есть. Это относится и к звезде, и к животному. *Программа человека – это прежде всего совокупность целей, программа-стремление, программа изменяющаяся, бесконечная. Телеологичность в такого рода программе отсутствует в долгосрочной перспективе или носит условный, порой сиюминутный характер. В противовес этому, природа и мир, являясь тем, что они есть, представляют собой постоянное и определенное бытие, они являются вещами. По его мнению то, что обнаруживает человек не имеет отдельного от него существования, оно имеет свое содержание только внутри него, существуя как удобство или неудобство.*

Существует ли, *может ли существовать нечто похожее на желания излишества у машины?* Ответив на этот вопрос положительно - можно с уверенностью сказать о возможности выхода ИИ за пределы техники. Но уже сам факт того, что природа наших желаний уходит глубоко в область нашего подсознательного, нашей природы и не является частью только социального, говорит о том, что создание таких искусственных, желающих машин невозможно. Подтверждение этой мысли можно найти у Феликса Гваттари, который говорит о возможности аутопоэзиса у машины только в рамках человеко-машинных систем, где наше тело является тем интерфейсом, который связывает биосферу и механосферу [Кучинов, 2016]. Если же - нет, то, судя по всему, мы в попытке создания такого технического объекта по необходимости приходим к тому, от чего ушли - к созданию по сути копии того же человека.

Человек, как говорит Ортега-и-Гассет, с легкостью отдает право желать за него другим. Тем самым философ предупреждает человечество о том, что современный человек не волен выбирать между жизнью в природе и

использованием сверхъестественного. Более того, уже современный философу человек приходит к ложной мысли, что машина, подобно всей окружающей действительности существует сама по себе, как нечто что не было необходимости произвести. Техника в сознании человека начинает обладать статусом подлинности. Сегодня, можно говорить о тенденции укрепления этого заблуждения - желании наделить машину в лице ИИ не только статусом подлинности, но и наделить чертами полноценной, более того «сверх-», личности. С психологической точки зрения этот факт в еще большей мере облегчает отказ человека от собственной независимости в мире.

Карл Ясперс

В работе «Смысл и назначение истории» Ясперс пишет: «в раю техники быть не может» [Ясперс, 1991, с. 134]. Там, где нет потребностей, нет необходимости и в каких-либо средствах для их удовлетворения. Определяя технику как промежуточное средство, необходимое для достижения цели, философ наделяет технику такими характерными чертами как «рассудок» и «власть» [Ясперс, 1991, с. 134]. В основе техники лежит власть рассудка, основанная на исчислении и предвидении возможностей [там же]:

«Техника оперирует механизмами, превращает свои данные в количественные отношения».

Она является частью общей рационализации как таковой [там же].

«Техника – это умение, методы которого являются внешними по отношению к цели. Это умение – способность делать и обладать, а не созидать и предоставлять расти».

Осмысливая этот ряд определений и основных характеристик данных философом, на наш взгляд можно с уверенностью выделить те главные

пункты, которые помогут нам в дальнейшем ответить на вопрос о том, что такое техника и ИИ сегодня. ИИ как средство оперирования, как способность делать и обладать, является в полной мере техническим объектом. Подтверждением этому же служит и мысль о том, что техника как совокупность открытых человеком повторяющихся в любом количестве операций резко отличается от творческой деятельности. А те случаи, когда техника перестает быть опосредующим звеном, он называет не более чем искажениями.

Понимая технику таким образом, неизбежно встает вопрос о тех границах, которые ей должны быть определены. Такого рода границами, по его мнению, являются:

Во-первых, тезис о том, что «техника - средство, которое должно направляться определенным образом» [там же]. Техника находится на службе человека, этим определяется ее ценность. При этом философ делает замечание, что сам по себе акт изобретения может и не быть связан с полезностью как таковой, а возникнуть лишь из удовольствия создания чего-то нового, чего-то того, что не существовало ранее. Однако решающим является фактор применимости [там же]:

«Изобретатель в области техники не создает принципиально новых потребностей, хотя, удовлетворяя их, он их расширяет и разнообразит. Цель должна быть задана <...> Граница техники в том, что она не может быть сама по себе, для себя, но всегда остается средством».

Из чего мы можем сделать вывод о нейтральности техники. Сегодня (или в недалеком будущем) может показаться, что искусственные системы начинают действовать достаточно автономно для того, чтобы иметь предпосылки поиска в них чего-то большего. Но проблема автономности и самостоятельности, дающее одно из оснований для возникновения субъектности, осмысленности собственных действий и т.д. не так проста, как может показаться на первый

взгляд. для понимания этого необходимо попытаться ответить на такие вопросы как:

1. Существует ли какие-либо условия задающие накладывающие ограничения на процессы самообучения и функционирования техники в части самостоятельного принятия решений?
2. Если нет, то, где видимые пределы возможностей самообучения и функционирования ИИ-систем?
3. Если же граничные условия существует, то каким образом должны техническим объектом формироваться собственные смыслы?

Наблюдаемое сегодня кажущееся самостоятельное развитие техники, не имеющее видимых очевидных целей, может показаться одним из свидетельств зарождающейся самостоятельности. Но если всмотреться внимательнее, то здесь мы опять увидим человека. Философ пишет о том, что изобретатель может создавать новые образцы технических объектов руководствуясь исключительно собственным удовольствием от создания чего-то нового.

Во-вторых, тезис о пределах [Ясперс, 1991, с. 135]:

господства техники «...только над механизмом, над безжизненным, универсальным. Во власти техники всегда лишь механически постигаемое».

Действительно, можно согласиться с тем, что управляющая техникой рассудочная деятельность всегда безжизненна и способна управлять живым только если она оперирует им, как если бы оно было не живым. На первый взгляд может показаться, что это не относится к ИИ. Ведь задача ИИ относится к области обработки информации. Более того, целью этих действий с информацией со стороны ИИ является не просто ее преобразование, но принятие решения в рамках прежде всего социальных коммуникаций. Отсюда встает вопрос о необходимости осмысления не только «так-бытия» ИИ, но и

введения его в социальное как полноценного участника, а значит понимание того, как именно ИИ оперирует информацией.

В-третьих, тезис о «демонизме» техники. [Ясперс, 1991, с. 135] Под этим понимается нечто, что было создано, но не понято людьми. Это нечто оказывающее влияние на человека, но не понятое, нераскрытое. Отсюда возникает как страх, так восторженное преклонение перед техникой.

Еще одним важным моментом в понимании им техники является рост технического мышления. Мышления рационального, распространяющегося все дальше, преобразующего все сферы человеческой деятельности [Ясперс, 1991, с. 135]:

«Совершающееся преобразование распространяется и на науку; очевидным свидетельством этого является технизация медицины».

Сегодня к этому можно добавить практически все сферы человеческой деятельности. Тем не менее, философ говорит о невозможности о невозможности абсолютной технократии [Ясперс, 1991, с. 137]:

«Остается, однако, вопрос, как человек, подчиненный техникой, в свою очередь станет господствовать над ней. Вся дальнейшая судьба человека зависит от того способа, посредством которого он подчинит себе последствия технического развития и их влияние на жизнь, начиная от организации доступного ему целого до его собственного поведения в каждую данную минуту».

Таким образом, техника в понимании философа является прежде всего средством, с помощью которого человек может управлять силами, трансформирующими природу. Средством для создания чего-то, чего не существовало до этого, чего-то что создает новый, искусственный мир. И

человек, являясь создателем этого мира, несет обязательства по управлению этим миром.

Отсюда возникает главная опасность, которую спустя более чем полвека увидит Бернар Стиглер. Эта опасность заключается в возникновении отставания развития культуры от прогресса техники. Но наличие такого разрыва, на наш взгляд говорит «лишь» о потере контроля, но не о прекращении «подручности» техники. Машина, над которой потерян контроль оказывается в аварийном положении, но не обретает самостоятельность.

Мартин Хайдеггер

Определяя сущность техники, Хайдеггер соглашается с тем, что она выступает и как средство достижения цели, и как «известного рода человеческая деятельность» [Хайдеггер, 1993, с. 221]. По мысли философа, прежде всего требуется понять причину возникновения и существования технических объектов, определяемую им через *τέλος*. А телос, трактуя Аристотеля, он понимает как «вину», но не в современном, нравственном, аспекте, а в смысле законченности. Законченности, определяющей не конец процесса оформления материи, а начало существования вещи в качестве того, для чего она была создана. Именно хайдеггеровская трактовка четырех видов Аристотелевской причинности, как четырех видов вины, создает отправную точку дальнейшего исследования. Обретение формы и готовность являются только началом, определяющим как нечто начинает присутствовать в своем *ὑποκείμενον*, предлежании под-лежащего. Но в отличие от того, что существует «по природе», про-из-ведение ремесла берет свое начало не в себе самом, а в том кто его создал.

Отсюда, со всей очевидностью, напрашивается вывод о том, что невозможно понять сущности техники как того, что принадлежит вещному миру (и ИИ, если он является именно технической вещью) вне понимания всех тех видов вины, которые послужили ее началом. И здесь же находится исходная точка, от которой необходимо отталкиваться на пути поиска ответов

на вопросы о том, какова возможная степень автономности ИИ, о достижимости тех пределов, которые позволят или нет создать в будущем действительно полноценную искусственную личность, тот сверхразум, который сегодня ожидают увидеть многие приверженцы этой идеи.

Продолжая свою мысль, философ видит сущность современной техники прежде всего в добыче, энергии. Но разве не напрашивается сегодня похожая аналогия в случае ИИ, если заменить энергию на информацию? В сущности, *сегодня мы имеем дело с таким же производством, но производством информации*. Важно отметить, что речь в данном случае идет не о получении новых знаний, а именно о процессе накопления и переработки информации. Обсуждение того, всегда ли производятся в ходе этого процесса новые знания остается, как минимум, открытым и требует отдельного обсуждения ввиду своей неочевидности. Для нас пока достаточно наличия аналогии между процессами добычи энергии и информации.

Все, что произведено должно быть в распоряжении. Хайдеггер определяет это как «состояние-в-наличии». И именно этот факт состояния-в-наличии и определяет дальнейшую несамостоятельность машины и того, что она производит [Хайдеггер, 1993, с. 227]:

«...ибо она держится только тем, что поставлена на предоставление поставляемого ею».

Определяя сущность техники через понятие «по-става», он ставит человека и природу на один уровень, оба они становятся частью технических процессов. Философ определяет характер техники как производящий выведение из потаенности [Хайдеггер, 1993, с. 228]:

«Это поставляющее раскрытие всего может осуществляться только в той мере, в какой человек со своей стороны заранее сам уже вовлечен в извлечение природных энергий».

Постав — это способ существования техники, но не нечто родовое общее, объединяющее все орудия и средства производства. В этом смысле, постав является миссией или судьбой как для человека, так и техники. Такая миссия, «пребывающая в образе постава» [Хайдеггер, 1993, с. 231], является опасной так как в этом случае изгоняется всякая другая возможность, а именно возможность раскрытия потаенности в смысле произведения - *ποίησις*. Но здесь же и скрывается как *возможность спасения*, так и *понимание факта несамостоятельности техники* [Хайдеггер, 1993, с. 236]:

«Всякая миссия раскрытия потаенности выполняется как о-существование и в качестве такого... Осуществляющее, тем или иным образом посылающее на путь раскрытия потаенности есть как таковое спасительное. Ибо оно дает человеку увидеть высшее достоинство своего существа и вернуться к нему».

Именно за человеком философ оставляет возможность видения и саму возможность осуществления изменений. Из чего можно сделать вывод, что *человек в отношении техники и ИИ, как технической вещи, остается causa efficiens*. В этом, по нашему мнению, и кроется начало понимания сущности и места ИИ в мире.

Но постав, являясь сущностью техники, выходит далеко за рамки поставляющего производства. Естественнонаучный способ исследований, по мнению философа, становится таким же поставляющим процессом, в котором теория заставляет природу представлять себя как расчетную силу. На наш взгляд, сложившаяся сегодня картина в таких областях как Data Science, имеет прямую аналогию с картиной мира, описываемой Хайдеггером. Сегодня выводы, сформированные в результате работы формализованных алгоритмов, начинают формировать поведение людей. Расчётным становится не только представление природы, но уже сам человек начинает представляться как

расчетная сила. При этом поле деятельности современного человека может сужаться до размеров экрана телефона. Сама жизнь человека уже начинает определяться правилами этого формализованного, по сути, виртуального или игрового мира, в котором и ресурсы, процессы их добычи и воспроизводства замыкаются в рамках дополненной реальности. Процессы обработки информации все больше начинают принадлежать не сферам научного познания, а включаться в рутинный производственный процесс, машинами (средствами и орудиями труда) в котором являются разного рода интеллектуальные системы. Человек в этом процессе из активного исследователя-изобретателя превращается одновременно и в источник сырья, и в рабочего-оператора. Это еще одна отдельная тема, требующего отдельного обсуждения, но выходящая за рамки настоящей работы.

Предлагаемые аналогии можно и развивать, и оспаривать, но если принять их, то с достаточной степенью уверенности можно сделать шаги к обоснованию следующих выводов:

Во-первых, о том, что сегодня техника остается инструментальной и «подручной» человеку и не имеет собственной самостоятельности.

Во-вторых, о том, что ИИ принадлежит к технике – *исходя как из условий его возникновения, так и из возможности его существования в мире.*

В-третьих, о том, что угроза превращения человека в ресурс возникла не сегодня. Сегодня эта проблема становится более наглядной, но не более того. При этом мысль Хайдеггера можно продлить в том отношении, что *человек сегодня – это не только часть процесса, но и подобно природе – источник ресурсов.*

Фридрих Дессауэр

Фридрих Дессауэр определяет технику как «судьбоносную» силу, указывая, что [2016, с. 393]:

«...мы имеем дело не с тем, что изучается посредством упражнений, не с рутинной и не с внешними процессами типа тематического вечера, но с техникой как судьбоносной силой, с историческим, трансформирующим человека и человечество, изменяющим Землю фактором культуры (в самом общем смысле слова)».

Философ четко отделяет то, что понималось под техникой ранее, прежде всего в античном мире и тем, о чем необходимо говорить сегодня. В качестве примера он приводит понимание техники Сократом прежде всего через диалоги «Горгий», «Хармид» и «Кратил» [Дессауэр, 2016, с. 391]. Он поясняет это на примере «техники» виолончелиста, которая подразумевает, что ее [там же]:

«...можно освоить путем упражнений (игра на виолончели), а иногда внешний рутинный процесс.»

Сократическая модель технэ есть знание, основанное на мастерстве, прежде всего личное знание, понимание ремесла [там же]:

«Сократ ясно понимает мысль о «финалистском» или «телеологическом», то есть о господстве цели (если говорить со стороны человека) или задачи (если говорить со стороны вещи)».

То, что должно быть создано возникает из мысленного образа, эйдоса, и образ этот не произволен, а определяется человеком.

Развивая свою мысль, и говоря о «технике животных», он пишет о явном её отличии от техники человека. Это отличие заключается в том, что первая имеет коллективный и фиксированный характер, руководимый инстинктом. Примеры изобретательства среди животных «нам не известны». Паук всегда плетет одну и ту же сеть [Дессауэр, 2016, с. 393].

Современное же понимание техники является гораздо большим чем просто знание, так как «несет в себе творческо-формообразующее, «изобретательское» начало» [Дессауэр, 2016, с. 393]. Человек в нем представляется как homo Investigator, homo Inventor, homo Faber [Дессауэр, 2016, с. 394].

Homo Investigator, человек исследующий хочет знать, он хочет изучать причину и следствие, пользу и вред, он оценивает. И происходит это через «переживание собственных духовных и душевных процессов» [там же]; через создание в памяти представлений, над которыми он может работать, «раз»-мышлять.

Homo Inventor, человек изобретающий, организующий, наряду с потребностью в знании имеет и потребность в «комбинаторном формообразовании» [там же]. Это дает ему способность к творчеству в рамках тех знаний о природе, что у него имеются.

Homo Faber, человек обрабатывающий, осуществляет перенос представленного из интраментального пространства во внешний мир.

Таким образом, по его мнению, все технологии имеют триединый характер. Все они были однажды изобретены, и превосходят все, что [Дессауэр, 2016, с. 393]:

«...преднаходится в качестве созданного самой природой, все они исторически перешли из небытия к бытию».

Для дальнейшего понимания сущности техники философом важно отметить, что, по его мнению, все технические объекты «строго привязаны к законам природы» [там же]. Несмотря на эту «привязанность» технические объекты [там же]:

«...способны оказывать воздействие за пределами технической предметной области».

Основными признаками технических объектов являются:

Во-первых, «привязанность» технических объектов. Они выполняют свои задачи строго соблюдая законы природы.

Во-вторых, «финализм». Если для неорганического естествознания финализм не существует (исследователь так или иначе работает только с казуальностью), то в биологическом он безусловно присутствует. Но здесь существует глубокое отличие от финализма собственно технических объектов. В случае техники, такого рода финализм объективен, он «впечатан» в предмет творцом-человеком. «Технический предмет является техническим лишь в той мере, в какой он исполняет свою задачу» [Дессауэр, 2016, с. 396].

В-третьих, обработка руками человека. Так или иначе любой технический объект является результатом обработки человеческими руками, станками, приборами.

Но для понимания техники этого недостаточно. Также, необходимо дать пояснение таким понятиям как «цель» и «задача». Человек может иметь цель, пишет Дессауэр, а прибор - нет. Более того, не существует такого технического объекта [Дессауэр, 2016, с. 400]:

«...который бы мыслил осознанно в подлинном смысле этого слова, выбирал, принимал решения, осуществлял целеполагание, не существует; такого предмета нельзя ожидать и от самой совершенной кибернетики».

Прибор может лишь исполнять свою задачу.

Причина, по которой такое самостоятельное целеполагание невозможно, кроется в ответе на вопрос о том, как интраментальные цели превращаются в пространственно-временные формы, в том как представление идей превращается в «оформленные объекты нашего мира» [Дессауэр, 2016, с. 395]. Для этого, по его мысли, необходимо обратиться к истокам, то есть к «открытию». Так же, как и Ортега-и-Гассет, он относит жизнь человека к

постоянному напряжению между действительным и возможным; к постоянной попытке приведения того, что есть к тому, чем оно может быть. Именно здесь возникает главный водораздел, отделяющий возможности человека и техники. Технические решения должны быть «изобретены». И этот процесс изобретения больше, чем процесс «разработки». Отсюда возникает необходимость поиска технических решений. Областью такого поиска является не область наличествующего, но возможного. Для поиска и воплощения готовых образов решений необходимо иметь доступ к воображению, необходима способность извлекать оттуда образы готовых решений. Дессауэр, следуя Канту, назвал это царство предустановленных решений «четвертым царством» [Дессауэр, 2016, с. 400]. Это царство обосновывает и ограничивает технику. Говоря о «творении», подразумевается вся «полнота космического», включающая в себя не только «телесные субстанции», но и «существенно большую сферу латентных, то есть скрытых, еще не реализованных образов» [там же].

Отсюда же следует одно из главных ограничений, налагаемых Дессауэром на технические объекты, включая, «самую совершенную кибернетику». Процесс подлинного изобретения представляет собой [там же]:

«...нечто большее, чем нежели «прикладное естествознание», большее, нежели простое соединение элементов, большее, нежели просто комбинация».

Для осуществления творческого акта необходимо иметь сродство между потребностями человека и предзаданными образами. Необходима структурная аналогия между ними, выраженная в человеческом мышлении, человеческом духе» [Дессауэр, 2016, с. 401]:

«Насколько эти соответствия двух сфер можно прояснить в понятиях соответствующих корреспондирующих структур, таких как

«инвариантность», «трансформация», «гомо- и изоморфность», пока неизвестно».

Тем не менее, философ, допускает одну область, на которую можно было бы распространить функциональные возможности машин. Речь идет о задачах конструирования, где цели определены. Процесс, где момент неожиданности, уникальности или не проявлен так явно или отсутствует вовсе [Дессауэр, 2016, с. 402]:

«Конструктор остается в области известного, изученного, методически-систематического».

Подводя итоги Дессауэр, делает вывод, что «порядок финалистски связанных элементов в единство системы есть место расположения силы технических объектов и технологий» [Дессауэр, 2016, с. 404].

На наш взгляд, наиболее важными выводами, которые можно сделать из понимания техники рассмотренными выше авторами являются:

1. Признание особого рода финализма технических объектов. Суть основного отличия сводится к изначальной ограниченности финализма технических объектов ввиду отпечатанности формы задачи в момент ее изобретения человеком.
2. Признание того факта, что воздействие технического, выходит за рамки сил природы и оказывает влияние на социальное, на человека. Отдельным остается вопрос о действительности существования такого феномена как «ко-эволюция» человека и техники: существует ли один эволюционный процесс, относящийся исключительно к живому, к человеку, в рамках которого технические изменения носят исключительно внутрикультурный характер, являющийся частью одного, а не двух процессов? В последнем случае видимый феномен «эволюции»

техники, происходящий на первый взгляд спонтанно и не зависимо от человека, является всего лишь скрытым эффектом творческого многообразия. Причиной многообразия в этом случае служит разрозненность целей, которым следуют изобретатели. Впрочем, ответ также на этот вопрос выходит за рамки настоящего исследования.

3. Признание неспособности технических объектов к решению творческих задач, к изобретательству по ряду объективных причин. Прежде всего, ввиду недоступности для машин области человеческих потребностей.
4. Признание опасности «подчинения» техникой человека. Но подчинение это носит со стороны техники скорее пассивный характер, обусловленный степенью вовлеченности человека в глобальный производственный процесс, *добровольным* отказом со стороны самого человека от собственной свободы.

§ Место техники в социальном

Если исследования первой половины XX века главным образом сводились к осмыслению непосредственно сущности техники, то дальнейшие исследования, по нашему мнению, в большей степени строились вокруг проблемы места человека в созданном им самим искусственном мире, в котором техника начинает играть все более значимую роль. К такого рода исследователям относятся такие философы как Жильбер Симондон, Мишель Серра, Бернар Стиглер и целый ряд других философов.

Сегодня не человека необходимо освобождать от техники, а технику от человека, утверждал в 1958 году Жильбер Симондон [Симондон, 2011]. Освобождать в том смысле что необходимо наконец предоставить технике заслуженное место в процессах технического функционирования. Признавая, по его мнению, за техническим объектом способность непосредственно совершать операции синтеза между материей и формой, мы признаем его участие как трудовое. Понимая под трудом с одной стороны – медиацию между человеком и природой, а с другой – *непосредственно* операцию по

изменению структуры материала в соответствии с формой-интенцией, философ констатирует, что деятельность рабочего-оператора уже не является трудом в строгом смысле этого слова. Рабочий выполняет две полу-цепочки: подготавливает материал и подготавливает форму. И именно машина, в качестве технического объекта, конституирующего человеческое и природное, выполняет недостающую фазу. Без этой фазы деятельность человека не может быть признана трудом в строгом смысле слова. В случае операции над машиной человек совершает скорее до-технические, нежели технические действия.

Таким образом, владея машиной, но переставая понимать внутреннюю суть непосредственной операции, человек начинает отчуждать машину. Отсюда возникает, прежде всего, культурное отчуждение, вызванное непониманием, которое приводит в том числе к возникновению новой магии, наделяющей машины свойствами человеческой личности.

В этом же ключе следует понимать его мысль о том, что технический объект, сконструированный человеком, не просто создает медиацию между человеком и природой, он воплощает в себе устойчивое смешение человеческого и природного, осуществляет вставку человеческого в природный мир. Именно это же устойчивое смешение добавляет желание наделить технику антропоморфными чертами, забывая, что технический объект был до того помыслен и изготовлен изобретателем. Важно отметить, что сам по себе акт изобретения характеризуется им как не являющийся трудом. Изобретение – это не просто адаптивное и защитное поведение, но и ментальная операция.

Разделяя интериндивидуальное и трансиндивидуальное, он определяет те граничные условия, в которых существует техника. Являясь полноправным субъектом интериндивидуальных коммуникаций, техника становится только объектом, носителем информации между одним человеком - отправителем и другим человеком - получателем в рамках трансиндивидуальной коммуникации [Симондон, 2011]:

«Информация — это не то, что приходит безотносительным образом, а *значение*, возникающее из передачи форм — одной внешней по отношению к субъекту, другой — внутренней».

Следуя мысли философа, являясь смешением человеческого и природного, технический объект обладает способностью к передаче информации. На наш взгляд, это утверждение можно было бы расширить даже признанием за машиной способности, в рамках своего функционирования, выполнять некоторые операции над информацией, по крайней мере те, что подпадают под определение труда. На первый взгляд, в таком утверждении может возникнуть некоторое сомнение: труд, согласно Симондону, должен обладать фазой непосредственной операции над информацией. А такая трактовка может сама по себе привести к той идеи, что операции над информацией ведут, допустим не всегда, но к возникновению новых знаний. Но на наш взгляд, такое расширенное толкование только усиливает утверждение о том, что ИИ остается не более чем техникой: труд, как уже было сказано выше, характеризуется в том числе синтезом материи и формы-интенции. Эта форма, задаваемая изобретателем-человеком, может возникнуть только в акте изобретения. Но, как мы уже знаем, акт изобретения не является трудом и является частью трансиндивидуального мира, присущего только человеку.

На наш взгляд, основными выводами из анализа идей философа, является тот факт, что роль техники ограничивается возможностью выполнять, во-первых, некоторый аналог труда. Во-вторых, передавать информацию, в том числе не связанную с производственной деятельностью.

Размышляя о положении техники в современном мире, Бернар Стиглер в своей работе «Техника и Время...» пишет [Stiegler, 2016, p 2]:

«...бытие техники сегодня необходимо рассматривать как совокупность разнородных сил».

Таким образом, философ располагает современную технику между двумя мирами: миром классической механики и миром биологическим. Мирами, соединяющими в себе как механические, так и биологические основания. Развитие техники приобретает собственные черты, о которых, по его мнению, начал говорить еще Маркс.

Размышляя о существовании техники, Стиглер идет дальше Хайдеггера. В его понимании Хайдеггер не увидел очень важного, а именно - конституирующую роль техники. По его мнению, главным в осмыслении истории развития человека, его бытия у Хайдеггера является процесс техницизации знания. Результатом это является возникновение *ratio* и «поставы»; к этому же относится и концепция забвения, связанная с тем, что история бытия у Хайдеггера становится не более чем записями в техничности.

На наш взгляд, этот факт сегодня становится актуальным как никогда, с появлением разного рода публичных хранилищ истории частных жизней, таких как Facebook или Instagram. «Как», по выражению Стиглера, сегодня во многом победило «что».

Техника сегодня уже не понимается как увеличение бытия как физиса (*physis*).

«Техника становится современной, когда метафизика выражает и завершает себя как проект расчетливого разума с целью овладения и владения природой, которая сама по себе больше не понимается как физис» [Stiegler, 2016, p 10].

Соглашаясь с мыслями Жана Ладриера и Бертрана Гилле, философ пишет, что сегодня техника во многом отделилась от культуры. Более того, сегодня можно наблюдать противостояние техники и культуры. Противостояние это обусловлено тем, что современное индустриальное общество ставит во главу угла, основывается на постоянном росте скорости

инноваций. При этом скорость развития техники такова, что сегодняшнее развитие культуры отстает от технического прогресса [Stiegler, 2016, p 15]. Если Хайдеггеровский экзистенциальный анализ ставит существование мира перед Dasein, говоря что «Dasein приходит в мир только постольку, поскольку мир всегда уже предшествовал ему в своей фактичности» [Stiegler, 2016, p 16], то в действительности дело обстоит иначе. Для понимания этого он вводит понятие третьего вида «бытия» в форме «неорганически организованных существ», или «технических объектов» [там же]. Благодаря тому, что сегодня становится очевидным превышение скорости эволюции этих технических объектов над скоростью эволюции человека, становится очевидным конституирующая природа техники.

Вступая в мысленный спор с Руссо [Stiegler, 2016, p 112], он констатирует, что вся история человечества всегда неразрывно основывалась на технике, начиная с первых своих шагов, встав на две ноги и начав манипулировать инструментами, человек начал дальнейший процесс эволюции под действием техники. Таким образом, мы можем говорить о втором рождении и «изобретении человека» [Stiegler, 2016, p 115]:

«С появлением экстериоризации тело живого индивидуума больше не является только телом: оно может функционировать только с помощью своих инструментов».

Вслед за Леруа-Гураном, он говорит о ко-эволюции человеческого мозга и орудий труда [Stiegler, 2016, p 148]. Но на этом, согласие Стиглера и Леруа-Гурана заканчивается. По мысли Стиглера, последний отказывается признавать творческое сознание основой эволюции. В частности, «Леруа-Гуран выступает против двух видов интеллекта, технического и нетехнического» [Stiegler, 2016, p 156], при этом философ видит непоследовательность Леруа-Гурана в этом вопросе. В действительности, по мнению Стиглера, прослеживается факт ко-эволюции, необходимыми

условиями которого является наличие одновременно и творческого и технического сознания; памяти, заключенной в следах материальной деятельности. Именно этот вид памяти, уже-существующий и определяет главным образом эволюцию и ту темпоральность, которая так необходима Dasein. Но не менее важным для Dasein является и предпонимание и входящее в его структуру предвосхищение. И здесь кроется один из основных ключей, которые привели к Homo sapiens.

Во-первых, отражаясь как в «зеркале в уже-существующих следах этой материальной деятельности, через экстерииоризацию собственной памяти, у человека появляется потребность в возникновении способности к предвосхищению» [Stiegler, 2016, p 159].

Во-вторых, а точнее как о следствии этого, следует говорить о способности человека к рефлексии и предвосхищению, «рефлексивной интеллектуальности». Если быть более точным, то к значительному росту эффективности этих ментальных процессов, связанных с эволюцией коры головного мозга. Как пишет Стиглер [Stiegler, 2016, p 162]:

«...рефлексивный интеллект, который не только улавливает взаимосвязь между различными явлениями, но и способен воплощать символическое представление об этой взаимосвязи, был конечным приобретением позвоночных».

Таким образом, говорить о противопоставлении техники и человека не приходится.

В своей работе «Автоматическое общество» [Stiegler, 2015] философ показывает, что сегодняшнее общество детерминировано автоматизацией сверх меры, автоматизации подвергаются не только процессы, но и само по себе поведение человека. Но автоматичность определяется им не как противопоставление автономии. Бытие по Стиглеру, в том числе социальное, есть различные уровни автоматизмов [Кучинов, 2016, с.125]. Иерархически

организованные автоматы, действующие в интересах контроллера, порождают, по его мнению, будущее которого так боялся Хайдеггер. Будущее, в котором торжествует исчисляющий ум. И это не сбой техники, а проблема того вида техники, который начинает доминировать сегодня.

Важно отметить, что для философа человек, безусловно, остается главным действующим лицом этих эволюционных и исторических процессов. Техника, оказываясь неотделимой частью мира человека, включаясь в интересубъективное, является полноправным участником множества процессов, составляющих искусственно созданную природу, влияя на культуру и в определенной степени. Ввиду различия скоростей развития уже становясь не ее частью. Это же приводит и к неизбежности тех проблем, которые мы можем видеть сегодня. Неизбежность эта обусловлена той податливостью человеческого сознания под действием технических объектов, которая когда-то обеспечила эволюцию человека.

И все же говорить об отсутствии или потере инструментальности со стороны техники не приходится. Именно человек, через способности к предпониманию, предвосхищению может определять свою дальнейшую судьбу, ведь техника, даже имея большую скорость развития, это прежде всего экстериоризированная память, доступная человеку для анализа, оставаясь при этом средством, готовым к использованию, «подручным» человеку.

Подтверждение этих идей, на наш взгляд, можно найти у Мишеля Серра. В своем эссе [2016] он подтверждает мысли Стиглера о характере влияния техники на человека. Если в работах первой половины XX века Шопенгауэра, Ясперса, Хайдеггера и Симондона мы видим страх перед вовлечением человека и техники в социальное через производство с его всеохватывающим «поставом», то сегодня на первый план выходят проблемы коммуникации и информации.

Сегодняшние дети иначе познают и воспринимают мир. Это хорошо видно на примере пространства. Как пишет философ [Серр, 2016, с. 14]:

«...их пространство - топологическое, где все соседствует со всем, тогда как мы жили в метрическом пространстве, измеряемом расстояниями».

человек становится индивидом. Этот индивид не входит так же надежно как раньше в ту или иную социальную группу (культуре, языку, религии и т.д.). Современный человек не способен жить в группе, он разводится, не ходит в церковь и не относит себя к какому-либо классу. Следуя Серра, можно сделать вывод, что сегодня человек остается с миром один на один, ограничивая его устройствами, которые является для него основным окном в мир.

Однако, по нашему мнению, человек не может быть один. Именно потребность в социальном в немалой степени способствует «оживлению» техники, активизирует поиски путей и способов включения техники в процессы коммуникации как полноценного субъекта, наделенного в его, человека, представлении сознанием и свободой волей.

Конечно же, в этом индивидуализме есть и другая сторона [Серр, 2016, с. 17]:

«Новорожденный индивид — скорее хорошая новость. Положив на одну чашу весов неудобства того, что старые ворчуны зовут эгоизмом, а на другую — унесшие сотни миллионов жизней преступления, совершенные коллективным либидо и ради него, — я готов возлюбить индивидов, этих нынешних молодых людей.

Сегодня все мы превращаемся в «святых Дионисиев», голова которого лежит у нас в руках [Серр, 2016, с. 26]:

«Компьютер — ящичек у нас в руках — вмещает в себя и приводит в действие то, что мы называли нашими «способностями»: память, причем тысячекратно более мощную, чем наша; воображение, нашпигованное миллионами образов; и даже разум, ведь компьютерные программы могут

решить сотни задач, с которыми мы бы в одиночку не справились. Наша голова лежит перед нами, в овеществленном ящике мысли».

В результате мозг человека становится «чистой познающей субъективностью» [Серр, 2016, с. 28].

Овеществление операций нервной системы приводит сегодня к тому, что процессы нашего познания приобретают новые орудия [Серр, 2016, с. 32]:

«Познавательные процессы — такие, как память, воображение, логическое мышление, геометрия и прочие тонкости, — нашли свое продолжение, вместе с синапсами и нейронами в компьютере».

В дополнение к этому он отмечает значительный рост алгоритмизации мыслительных процессов, отличающихся от «дедукции геометра или индукции экспериментатора» [Серр, 2016, с. 70].

По нашему мнению, такого рода алгоритмизация приводит к действительному сближению хода течения мыслительных процессов современного человека и машинного интеллекта. Что, в свою очередь, приводит к возникновению иллюзии того, что техника и человек функционируют на одних и тех же картезианских принципах ментальных репрезентаций.

Из всего изложенного можно сделать следующие выводы:

1. Все больше входя в социальное, проникая в те области, в которых до этого было место лишь *человеку, техника, тем не менее, остается «подручной» человеку.*
2. Оставаясь «подручной», техника сегодня является одним из *конституирующих* современную культуру начал, приводя к росту *алгоритмизации* мыслительных процессов человека.

3. Происходящие сегодня изменения скорее следует отнести к изменениям *в способах и организации коммуникативных и мыслительных процессов самого человека*, не отрицая опосредованного влияния техники на процесс культурных, социальных и ментальных процессов в обществе.

§ Критика основных технологических подходов

Символьные методы

Суть основной критики символьных подходов на наш взгляд можно свести к следующим пунктам:

Во-первых, системы, построенные на базе символьных методов, строятся как правило на двух взаимодополняющих концепциях: деревья решений и наборы правил. В основе процесса обучения в рамках таких систем лежит принцип обратной дедукции [Домингос, 2016]. Еще в начале 60-х годов специалисты в области ИИ пытались моделировать сложные процессы мышления, отыскать какие-либо методы решения задач в общем виде. Как оказалось, существует целый ряд проблем, прежде всего связанных с экспоненциальным ростом сложности – проблема комбинаторного взрыва. Все это привело к пониманию того, что чем шире класс задач, которые может решить какая-либо программа, тем беднее оказываются ее возможности при решении конкретной проблемы [Уотерман, 1989, с.13].

Во-вторых, проблема автоматического порождения новых правил. Суть проблемы заключается не в поиске самого алгоритма порождения правил. Такие попытки существовали, и одним из успешных примеров можно считать ДСМ-метод. Основная проблема заключается в том, что [Гаддамер, 1988, с. 430]:

«...не существует метода, который позволил бы научиться спрашивать, научиться видеть проблематическое».

Таким образом, сегодня нет возможности расширения поля задач особенно в условиях неполной информации и отсутствия явно заданных целей.

В-третьих, ограниченность символьных представлений в возможности объяснения поведения, сформулированная Дрейфусом как проблема того, что [Dreyfus, 2007]:

«...наше повседневное поведение не может быть понято в терминах выводов из символических представлений <...> не может быть понято в терминах реакций, вызванных фиксированными особенностями окружающей среды»

Из этого философ делает вывод о необходимости отказа от внутренних символических представлений.

На наш взгляд, такое положение дел с одной стороны, действительно ограничивает пределы возможностей символьных методов в качестве самостоятельного подхода, требует как минимум решения проблемы автоматического порождения новых понятий, символов и правил, а с другой – дает обоснование смены приоритетов исследовательских программ от попыток создания полностью автономного ощего ИИ в сторону создания гибридных, человеко-машинных систем.

Коннекционизм.

Суть основной критики в отношении коннекционистского подхода можно свести к следующим основным пунктам:

Во-первых, речь идет о невозможности эксплицировать результаты работы обученной нейронной сети. Таким образом, идея «интеллектуальности» сводится к идее, реализуемой как машинный интеллект, представляющий собой не более чем процесс получения

статистически значимой матрицы соответствия входных и выходных сигналов.

Во-вторых, речь идет о проблеме определения смыслов, значимых паттернов. Например, на сегодняшний день все операции такого рода, например подготовка обучающих выборок, их разметка, производится человеком, в лице оператора или аналитика.

В-третьих, проблема объяснимости полученных решений. Вопрос в данном случае идет не о степени доверия к предлагаемым решениям со стороны пользователей. Проблема носит как социально-этический, так и философский характер. С социальной точки зрения суть проблемы хорошо передают опубликованные предложения Еврокомиссии⁹ по ограничению использования ИИ ввиду наличия проблем точности, надежности и открытости принимаемых ИИ решений. С философской точки зрения данная проблема отсылает к целому ряду проблемам так или иначе связанных с вопросами теории познания и достоверности научного знания. Подробный анализ этих вопросов выходит за пределы исследуемой тематики.

В-четвертых, мифологизация коннекционистских подходов. Суть проблемы состоит в попытках связать исследования сознания и исследования в области нейронных сетей. К такого рода исследователям В.Г. Григорьев [Григорьев, 2017, 121-124] относит прежде всего Дениэла Деннета. Противниками тезиса о том, что отношение между сознанием и мозгом носит нейросетевой характер являются такие авторы как Джордж Фодор и Зенон Пылишин.

Говоря о проблемах и мифах, возникающих вокруг коннекционистских методов, В.Г. Григорьев отдельно выделяет «4 мифа» Иштвана Беркли [Кузнецов, 2017, с. 121-124]:

⁹ https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682

1. Коннекционистские системы являются похожими на мозг. Наиболее значимыми и убежденными последователями этого мифа являются, по его мнению, Пол Черчленд и Дэниел Деннет.
1. Коннекционистские методы работают в реальном времени. В частности, Д. Руммельхартом утверждается: что наш мозг, хоть и работает медленнее, чем компьютерная система, получает преимущество «сжатия» времени за счет параллельной обработки. Беркли не соглашается с таким тезисом.
2. Предпочтительность коннекционистских методов по отношению к не коннекционистским. Черчленд, в поддержку этого мифа ссылается на успехи нейронных сетей в области распознавания образов. На самом деле, никакие из методов в условиях неполноты информации не имеют преимущества.
3. Миф об обучаемости коннекционистских методов. В частности, речь идет о том, что нейронные сети умеют переносить полученные навыки на новые случаи. Однако, даже если допустить принципиальную возможность такого «переноса», то речь идет не более чем как об аналогии. Вопросы генезиса, интерпретации и производства новых смыслов полностью остаются за пределами возможностей коннекционистских методов.

По мнению В.Г. Григорьева, сегодня такого рода мифы ушли в прошлое, однако анализ современного дискурса, на наш взгляд, говорит об обратном.

Таким образом, так же, как и в случае с символьными методами, существующие проблемы делают как минимум неочевидными перспективы создания систем общего ИИ только с использованием одних только коннекционистских методов.

Общий вывод, который можно сделать из приведенного критического анализа, заключается в том, что ни один из существующих технологических подходов не является достаточным для создания системы общего ИИ.

Одновременно с этим, на наш взгляд, имеются хорошие перспективы в области реализации исследовательских программ, объединяющих символичные, коннекционистские и т.д. методы.

Задача понимания «самоочевидного» и «самопонятного» всегда являлась одной из ключевых задач философии. Одной из таких задач сегодня является осмысление того, какое положение занимает сегодня ИИ и каковы его перспективы развития. Среди таких «самоочевидностей», на наш взгляд, сегодня прежде всего выступают идеи разного рода «сверх-»: сверхразум, сверхсознание, достижение сингулярности и т.д. Как уже упоминалось в начале, большое количество современных футурологов таких как Ник Бостром, Реймонд Курцвейл, предвещают в относительно скором будущем возникновение (сверх)разума подобного или превосходящего человека. В своей работе «Искусственный интеллект...» Бостром пишет, что в случае, если когда-нибудь разработают искусственный разум превосходящий человека, то [Бостром, 2016, с. 14]:

«...судьба нашего вида окажется в прямой зависимости от действий этих разумных технических систем – подобно тому, как сегодняшняя участь горилл в большей степени определяется не самими приматами, а людскими намерениями».

Несколько в ином ключе, и прежде всего, без картин, связанных с неизбежностью грядущего апокалипсиса, но с не меньшей убежденностью в неизбежности возникновения искусственного сверхразума размышляет и Михаил Эпштейн. В 2008 году он пишет об Интернете, как инкубаторе для такого сверхразума [Parsing the Turing Test, 2008, p. xi]. А спустя 10 лет, уже

рассуждает о концепции ноовитализма, срачивающего живое и техническое, где ИИ [Эпштейн, 2018]:

«...это особая ментальная сверхструктура, которую мы не можем локализовать как отдельный механизм, как не можем локализовать виртуальное пространство».

Несмотря на очевидные различия в концепциях Бострома, Эпштейна и ряда других исследователей, представляется очевидным, что ИИ часто начинает восприниматься как нечто, выходящее за рамки технических, искусственных объектов. В умах людей ИИ начинает обладать уже не просто свойствами субъектности и определенной автономности в рамках тех или иных моделей коммуникации, но и разумом, свободой воли, собственной субъективностью и субъектностью. ИИ в свете такого понимания становится полноценным субъектом, границы автономности которого выходят далеко за пределы тех, что предполагаются в работах Симондона, Стиглера и ряда других исследователей: границы автономности функционирования ИИ-систем перестают зависеть от человека.

Понимание того, соглашаться ли с такого рода тезисами или нет, на наш взгляд, во многом определяется тем онтологическим статусом, которым обладает сейчас и может обладать в перспективе та или иная ИИ-система. Только осмыслив онтологический статус ИИ становится возможным понимание того, каково место человека и машины и символических методов, в частности.

Все это приводит к необходимости ответить на вопрос о том, как, каким образом может существовать ИИ-система? Ответить на этот вопрос необходимо прежде всего в рамках анализа онтологического статуса ИИ. На наш взгляд, такого рода анализ должен ответить прежде всего на следующие вопросы:

1. Является ли ИИ техникой, «подручной» человеку? Если да, то каковы перспективы выхода ИИ за пределы техники?
2. Каковы граничные условия процессов самообучения и автономности функционирования ИИ-систем?
3. Каково место ИИ в социальном?

Отправной точкой настоящего исследования является тот факт, что при внимательном рассмотрении становится возможным построить достаточно четкие аналогии между тем миром техники, о котором писали Хайдеггер, Ортега-и-Гассет и Симондон, и положением дел сегодня, с момента возникновения таких феноменов как ИИ.

Казалось бы, постановка вопроса о том, является ли ИИ техникой может показаться на первый взгляд тривиальной. ИИ-системы являются, прежде всего, результатом инженерной мысли и, как следствие, техникой. Однако, с момента своего возникновения в работах Эрнста Каппа, философия техники так или иначе имеет склонность к обнаружению подобия между способами существования уже самих по себе технических объектов и живых объектов. Или (как в работах Жильбера Симондона, Бруно Латура и ряда других исследователей) наделяния технических объектов свойствами полноценных акторов, обладающих субъектностью в довольно широком понимании.

К этому следует добавить идеи таких авторов как Дэвид Чалмерс, чья концепция организованной инвариантности [Чалмерс, 2015] фактически пытается стереть границы между техническими объектами и человеком. С целью прояснения того, что представляет собой сущность техники далее предпринимается критический анализ взглядов различных авторов относительно сущности техники, и дается обоснование необходимости понимания ИИ как технического объекта. Определяя ИИ как технический, тем самым мы должны с необходимостью говорить об ИИ как о средстве, подручном человеку.

Мартин Хайдеггер определяет «бытие-в-мире» как познание [Хайдеггер, 2015, с. 61], без которого присутствие в качестве «бытия-в» невозможно. В то же время, относительно вещи можно говорить только о «бытии-при» [Хайдеггер, 2015, с. 55]. Оставаясь средством, бытийный статус которого определяется как «для-чего». ИИ, в этом случае, не обладает той необходимой полнотой «бытийного рода присутствия» [Хайдеггер, 2015, с. 71], обеспечивающей возможность генезиса.

Такое понимание ИИ является, на наш взгляд той отправной точкой, которая позволяет прояснить тот способ каким возможно производство нового знания, каковы границы автономности и пределы возможностей ИИ-системы в вопросах самообучения. Определяя систему как подручную человеку, мы, во-первых, по необходимости очерчиваем границу ИИ-систему как зависимую по необходимости от человека в вопросах самообучения. Во-вторых, проясняем пределы возможностей ИИ-системы в вопросах генезиса новых понятий, правил и знаков. На наш взгляд, главным вопросом на пути понимания того, остается ли ИИ-система только «подручной» человеку является анализ возможности ИИ-системой самостоятельного генезиса новых понятий, правил и знаков.

В заключение, приведем заключение В.И. Лекторского о том, что сегодня большинство исследователей, работающих в области ИИ, пришло к выводу что «нельзя ничего понять в сознании, психике, в ментальных процессах, если мы не учтем, что познающее существо включено в мир и находится в постоянном взаимодействии с ним» [Философия искусственного интеллекта, 2016, с. 41]. Говоря о таком феномене как ИИ, под миром прежде всего необходимо понимать социальное, о котором говорил Бруно Латур [Латур, 2020: 11]:

«...социальное не может мыслиться как разновидность материала или как некая сфера»

Отсюда мы можем сделать вывод о возможности утверждать, что без понимания того, что представляют собой участники социальных процессов, нельзя понять причины происходящего.

Современное прочтение некоторых идей вышеупомянутых и ряда других философов может в значительной степени помочь в понимании того, что есть сегодня техника и ИИ, в частности.

§ Критика возможности выхода ИИ за границы техники

В рамках настоящего исследования данный вопрос является одним из наиболее важных, так как ответ на него является, в том числе, ключом к пониманию *роли, пределов возможностей и граничных условий процессов самообучения и принятия решений в ходе автономного функционирования ИИ-систем*. Понимание этого необходимо для ответа на вопрос «каким образом, могут существовать символьные методы?».

На наш взгляд, именно такая, предикативная, постановка вопроса может дать понимание того каким способом, в каком окружении и в какой степени самостоятельно, некоторая система ИИ может реализовать процессы оперирования символами, связанными с ними значениями и смыслами.

До этого момента мы исследовали природу техники (и ИИ как технического объекта) задаваясь вопросом о возможной самостоятельности такого рода объектов на современном уровне развития, во-первых, через понимание техники относительно критерия «подручности». Во-вторых, через осмысление роли техники как части создаваемого человеком искусственного мира, все дальше отделяющегося от природного мира. В-третьих, через попытку осмысления места технических объектов в этом искусственном мире сквозь призму социального.

Но такого анализа недостаточно. Как было показано выше, даже до появления такого феномена как ИИ в отношении техники исторически

неоднократно проводились попытки переосмыслить ее сущность, в той или иной мере ставя под сомнение прежде всего «подручность» техники. С появлением первых электронных вычислительных машин вопрос о возможности выхода ИИ за границы «подручности», стал как никогда актуальным. По нашему мнению, одним из необходимых условий, без которых невозможен выход ИИ за границы «подручности» является не просто наличие некоторых мыслительных процессов, но наличие у ИИ собственной сознательной воли. В противном случае, следовало бы говорить не о выходе техники за пределы «подручности», а *о потере контроля* над техникой со стороны человека.

Проблемы в данном вопросе, начинаются уже с определений, касающихся ИИ как предмета исследования. Можно согласиться с С.Шапиро в том, что нужно выделить такие вопросы, характеризующие эту проблему:

1. Что понимать под термином ИИ, в чем различие между естественным, искусственным и машинным интеллектом?
2. Является ли вычисление достаточным для возникновения сознания?
3. Возможен ли ноэзис и семиозис без наличия сознания? (анализ данного и последующих вопросов является одной из целей настоящего исследования).
4. Возможно ли создать систему, обладающую (само)сознанием с помощью машинного интеллекта?
5. Возможно ли разрешение контroversы между доказательством и открытием истины [Шапиро, 1973, с.8].

Прежде всего, необходимо отметить, что на сегодняшний день существует практически единственная модель абстрактного вычислителя, представляющего собой основу функционирования машинного интеллекта - детерминированная машина Тьюринга, реализованная на основе архитектуры фон Неймана. Основная идея машины Тьюринга заключается в возможности

манипулирования символами *формально* и *синтаксически*. Варианты рассмотрения других алгоритмов, таких как недетерминированная машина Тьюринга, вероятностная машина Тьюринга, не принесет значительной пользы, так как все такого рода алгоритмы представляют собой варианты реализации все тех же формальных, работающих с синтаксисом автоматов.

Судя по всему, впервые термин «*сильный* искусственный интеллект» был введен Джоном Сёрлем в 1980 году как точку зрения, согласно которой определенным образом запрограммированные компьютеры 1) в буквальном смысле слова обладают когнитивными состояниями, 2) такие программы объясняют феномен человеческого познания [Searl, 1980]. Несколько позднее эту мысль перефразировал Дэвид Коул в статье «Аргумент китайской комнаты» как [Cole, 2020]:

точку зрения «согласно которой соответствующим образом запрограммированные компьютеры (или сами программы) могут понимать естественный язык и на самом деле обладают другими умственными способностями, аналогичными людям, чье поведение они имитируют.»

В работе «Общество разума», 1986 года, Марвин Минский прямо утверждает, что Уоррен Мак-Каллок и Вальтер Питтс «показали, что машины можно побудить видеть, рассуждать и запоминать» [Минский, 2018]. В 2008 году Сёрль определяет «сильный искусственный интеллект» как современную версию функционализма, суть которой заключается в гипотезе о том, что [Parsing the Turing Test, 2008, p. 142]:

«...основной механизм, создающий поведение, удовлетворяющее критерию Тьюринга, является вычислительным».

Дэвид Чалмерс полагает, что основной задачей исследований в области сильного ИИ является «воспроизведение ментального в вычислительных процессах» [Чалмерс, 2015, с. 389]. Важно отметить, что под ментальным он полагает не только и не столько сознательный опыт. В его понимании ментальное имеет двоякую природу. С одной стороны, это есть феноменальное понятие в качестве сознательного опыта, а с другой – через психологическое понятие – в качестве «казуальной, или объяснительной, основы поведения» [Чалмерс, 2015, с. 29].

Так или иначе, какого бы определения мы ни придерживались, понимание сильного или общего ИИ напрямую связывается с некоторой вычислительной системой, способной не просто конкурировать с человеком в интеллектуальной сфере, но и обладать теми же когнитивными феноменами, такими как сознание, мышление и т.д. Следовательно, исходя из целей, преследуемых исследователями в области общего ИИ, одним из основных критериев, обеспечивающих возможность выхода ИИ за пределы «подручности» необходимо считать наличие сознания.

В этих же целях Маркус Габриель определяет необходимость наличия у объекта, исследуемого на предмет наличия сознания, во-первых, интенционального, а во-вторых, феноменального сознания [Габриель, 2020, с. 108]. В качестве «интенционального» сознания понимается «сознание о чем-то», на что мы направляем наше внимание или внешнее восприятие, внешняя перспектива. «Феноменальное сознание» связано с внутренней перспективой. В этом случае имеются в виду наши «чисто субъективные», внутренние переживания [Габриель, 2020, с. 108]. Философ утверждает, и с этим, на наш взгляд, следует согласиться, что для наличия сознания необходимо присутствие обоих видов сознания.

В случае отсутствия феноменального сознания, у такого объекта нет переживаний, он не чувствует себя особым образом. В качестве примера приводится мысленный эксперимент с голубым кубиком. Безусловно, система ИИ или робот мог бы зафиксировать наличие объекта, имеющего цвет в

определённом диапазоне светового спектра. Таим образом, автор вполне допускает наличие у робота интенционального сознания. Проблема заключается в том, что ввиду отсутствия переживаний невозможно строить *образы самого себя*, а следовательно, невозможно обладать желаниями так необходимыми для возникновения осознанной воли. Такого рода машинам и генерируемым ими сообщениям, не хватало бы «релевантной феноменальной тонкой зернистости» [Габриель, 2020, с. 107].

С другой стороны, в случае отсутствия интенционального сознания, при наличии одного лишь феноменального, возникает обратная ситуация, при которой такая система попадает в «состояние абсолютно хаотического мелькания данных» [Габриель, 2020, с. 107]. В качестве примера философ приводит ситуацию с наркотическим опьянением или [Габриель, 2020, с. 110]:

«...обретшим форму времени импрессионизме, в котором мы больше не видим предметов, а переживаем только потоки цветов».

Похожие мысли мы можем найти еще у Дэвида Юма. Разделяя понятия «впечатления» и «идеи», философ «классифицировал формы внутреннего переживания», где под впечатлениями понимается «мешанина данных», а под идеями-упорядочивание этой мешанины [там же]. Так мы приходим к кантовскому высказыванию [Кант, 1994, с.89]:

«созерцание и понятия образуют элементы всего нашего познания, так что ни понятия без соответствующего им некоторым образом созерцания, ни созерцания без понятий не могут дать знание».

Однако, само по себе наличие интенциональности и феноменальности не является единственной проблемой. Не менее важной является проблема креативного мышления. Основную суть этой проблемы можно сформулировать следующим образом: каким образом было бы возможно в

рамках логики машины Тьюринга (являющейся исключительно символьной), осуществить высокоуровневую концептуализацию, в частности, осуществить генерацию новых смыслов. Иными словами - перейти от синтаксиса к семантике.

Возможно, к этому следовало бы добавить еще одну проблему. Речь идет об эпифеноменализме, утверждающем, что наше поведение детерминировано физическими процессами и наличие таких феноменов как квалиа никак не влияет на физические процессы. На наш взгляд, аналогичным образом следовало бы исследовать вопрос о наличии или отсутствия влияния семантических структур на физические процессы. Впрочем, эти вопросы так же выходят за рамки настоящего исследования.

Судя по всему, одним из первых кто наиболее четко сформулировал эти проблемы, является Джордж Сёрль еще в своем мысленном эксперименте «Китайская комната» [Searle, 1980]. Основными объектами критики философа являются тезис об эпифеноменализме и неизбежно следующий из этого бихевиоризм (позднее функционализм). В некотором смысле, частным следствием из этого является критика возможности перехода от синтаксиса к семантике в рамках машины Тьюринга.

Согласно пониманию Сёрлем сути логического бихевиоризма, для признания за ИИ наличия ментальных процессов, аналогичных биологическим системам (человеку и, как минимум, хордовым), было бы достаточно следовать следующей формуле [Parsing the Turing Test, 2008, p. 140]:

«...если машина может демонстрировать точно такое же поведение, которое человек демонстрирует в отношении некоторого когнитивного феномена, то в этом случае машина обладает этим когнитивным феноменом в том же самом смысле, что и человек».

Но такой вывод, по его мнению, был бы преждевременным. Для обоснования этого тезиса он прежде всего предлагает разобраться в понимании значения термина «искусственный». Необходимо четко понимать отличия между чем-то настоящим, но созданным искусственно и имитацией чего-то настоящего: «искусственные сливки на самом деле вовсе не сливки» [Parsing the Turing Test, 2008, p. 143]. Подобную аналогию можно применить и по отношению к «искусственному интеллекту».

Не менее двусмысленным является и термин «интеллект». По его мнению, следовало бы использовать выражение «ментальный интеллект», как дефиницию, подчеркивающую наличие ментальных или когнитивных процессов. Он присущ людям и животным. В свою очередь, «не-ментальный интеллект» подразумевает только процессы особого рода, такие как вычисление или передача данных. К обладателям второго рода интеллекта относятся книги и карманные калькуляторы [там же].

Таким образом, мы можем более четко специфицировать задачу, стоящую в области технологий ИИ: искусственное создание реального ментального интеллекта. Иными словами: пытаемся ли мы создать настоящую ментальную вещь или нет? В этом месте безусловно напрашивается необходимость ответа на вопрос о возможности существования небиологического ментального интеллекта. И прежде всего, путем поиска ответа на вопрос о том, каким бы мог быть его Dasein? Впрочем, такого рода вопросы также выходят за рамки настоящего исследования. Единственное, что необходимо заметить это то, что при такой постановке вопроса, тест Тьюринга следует специфицировать как тест на наличие именно человеческого ментального интеллекта со всеми вытекающими из этого последствиями.

Возвращаясь к мысли Серля, искусственное воспроизводство ментального интеллекта невозможно, так как бихевиористский и функционалистский принципы, заложенные в логику машины Тьюринга, не позволяют это сделать [Parsing the Turing Test, 2008, p. 144]:

«...вся система не имеет возможности присоединить семантику к синтаксису компьютерных символов».

Основная трудность здесь заключается в том, что при таком подходе не учитываются специфические особенности психических состояний, которые могли бы порождать причинность. По его мнению, в настоящее время существует только вера в то, что это возможно реализовать в терминах входных и выходных стимулов.

На сегодняшний день мы создали машины, оперирующие, по сути, двумя типами символов 0 и 1, неизбежно порождая вопрос о достаточности этих двух символов. В конце концов, нейроны мозга тоже находятся в состоянии активности или нет. Как пишет философ, этот вопрос вполне разумен [Parsing the Turing Test, 2008, p. 144]:

«есть простой и отвечающий на все ответ: нейронные активности являются частью *причинного* механизма, который вызывает сознание и познание как способности более высокого уровня».

Таким образом, мы выходим за рамки только вычислений, констатируя наличие более глубоких ментальных уровней, обуславливающих наличие сознания и самосознания. Программа, оперирующая символами «ничего не знает о необходимости активировать сознание или что-то еще» [там же]. Компьютерная программа определяет только класс эквивалентности, который не зависит от аппаратной части. Эту же мысль можно встретить у С.И. Шапиро [1973, с. 35]:

«...логическая модель отвечает на вопрос, как человек должен мыслить; психологическая – как он мыслит в действительности».

Продолжая свою мысль, Сёрль утверждает, что вычисление не является достаточным основанием для возникновения сознания, «мышление должно иметь нечто большее, чем просто символы» [Parsing the Turing Test, 2008, p. 144]. Примечательно, что в качестве одного из критериев, определяющих различие между машиной и человеком философом определяются процессы, лежащие в основе человеческого мозга. Процессы в мозге определяются передачей энергии, а в компьютере - абстрактными математическими процессами. Хьюберт Дрейфус так выражает эту же мысль [Dreyfus, 2007, p. 1142]:

«...необходимо рассмотреть возможность того, что воплощенные существа, подобные нам, берут в качестве входных данных энергию из физической вселенной и реагируют таким образом, чтобы открыться миру, организованному с точки зрения их потребностей, интересов и физических возможностей...».

В 2007 году выходит статья Дрейфуса, в которой говорится о том, что ИИ, если мы хотим его создать, должен быть еще более хайдеггиреанским, чем сегодня. Прежде всего, философ ставит вопрос о проблеме представления значимости и актуальности, без которых картина мира представляет собой набор бессмысленных фактов. Но проблема на этом не останавливается: как предупреждал Хайдеггер, ценности – это просто более бессмысленные факты. В качестве примера Дрейфус приводит Марвина Минского, который «был убежден, что представление нескольких миллионов фактов об объектах, включая функции, решит то, что стало называться проблемой здравого смысла» [Dreyfus, 2007, p.1138]. Однако в действительности, наличие миллиона фактов никак не решает другую проблему, проблему фрейма. Суть которой заключается в том, что программе необходимо каким-то образом знать, какие факты имеют значение в каждой конкретной ситуации с учетом того, что мир постоянно меняется. Для решения этой проблемы Минский

предлагал использовать ситуационный подход. Но система фреймов также должна как-то реагировать на изменения в мире, перестраивая степени актуальности тех или иных фреймов. Таким образом, задача становится бесконечно рекурсивной. Об этом же, еще в 1973 году говорил Шапиро, указывая на проблему изоморфизма при описании психических процессов [Шапиро, 1973, с.33]:

«...при взаимодействии с субъектом материальный объект непрерывно изменяется, и для сохранения соответствия необходимо адекватное изменение психологической модели».

В этом смысле формулировка Бенедиктом Спинозой мышления как движения по форме предмета является полностью соответствующей сути нашей проблемы. В конечном итоге, в 2003 году Минский признал на страницах журнала Wired [Minsky,2003]:

«...мозг ИИ мертв с начала 70-х годов, когда он столкнулся с проблемой знаний, основанных на здравом смысле».

В ходе доклада 1986 года, сделанного в Лаборатории искусственного интеллекта Массачусетского технологического университета, Дрейфус констатирует [Dreyfus, 2007, p.1140]:

«Значимые объекты...среди которых мы живем, не являются моделью мира, хранящейся в нашем сознании или мозге; они и есть сам мир».

Одним из важных выводов, которые можно сделать из этого высказывания - это вывод о констатации несостоятельности идеи Декарта об эпистемологической изолированности нашего сознания. Такой вывод ставит под сомнение целый ряд исследовательских программ в области общего ИИ.

В частности, автор на этом основании критикует, хотя и признает их первыми шагами на пути решения проблемы изолированности, исследовательские программы Родни Брукса, Фила Агре и Дениела Деннета.

В качестве альтернативы философ ссылается на Хайдеггера и Мерло-Понти. Для Хайдеггера готовность к действию не является жестко связанной парой «определенная функция» - «определенная ситуация». Скорее стоит говорить о готовности к действию как к побуждению, которое вызывает гибкую реакцию. Более важным является вопрос не «что?», а «как?». Эту мысль позднее развил Мерло-Понти введя понятие интенциональной дуги: приобретая навыки агент сохраняет их не в виде представлений, а в виде побуждений к ситуациям в мире. Как пишет Дрейфус [Dreyfus, 2007, p. 1147]:

«в нашей квалификационной деятельности нас тянет двигаться так, чтобы все лучше контролировать ситуацию <...> Не нужно знать, что такое оптимум, чтобы двигаться к нему».

В подтверждение своей мысли он приводит слова Тимоти ван Гельдера [там же]:

«...познание в сложных ситуациях может <...> включать представление и последовательную обработку, но такие явления лучше всего понимать как возникающие из динамического субстрата, а не как составляющие базовый уровень».

Такой вывод однозначно приводит к столкновению моделей непрерывной причинности с моделями, основанными на действии.

В качестве попытки, не совсем удачной, «примерить» эти подходы философ упоминает точку зрения Майкла Уилера [Wheeler, 2007], которая опираясь на то, что Хайдеггер именуется как «иметь дело с неподготовленным». В отличие от Хайдеггера, Уилер все же подразумевает специфичность

представлений относительно тех или иных конкретных ситуаций. Но этого нет у Хайдеггера, так как, по мнению Дрейфуса он предполагает, что у него [Dreyfus, 2007, p. 1149]:

«...все действия, в том числе неготовые к действию, происходят на фоне базовой, непредставительной, целостной, поглощённой интенциональности, которую Хайдеггер называет бытием-в-мире».

В отличие от Уилера, у Хайдеггера ориентированное на действие преодоление вообще не является репрезентативным.

В результате Дрейфус заключает, что все общепринятые нейромодели не только не помогают, а скорее мешают ответить на вопрос о том как из бессмысленного потока данных мы получаем знания. Причина этого в том, что так или иначе все они воспроизводят декартовскую модель, в которой получая информацию через органы чувств, мозг путем абстрагирования строит репрезентативные представления о мире. Иными словами, мозг – пассивный приемник.

В качестве реальной альтернативы философ приводит нейродинамическую модель Уолтера Фримена [Dreyfus, 2007, p. 1151]. Суть данной теории в том, что мозг не производит связывания информации в виде артефактов при попытке их интерпретирования. В реальности существует проблема того, как мозг выбирает значимую для него фигуру из окружающего шума. Другими словами, речь идет о проблеме интенциональности. В модели Фримена [Freeman, 1991, 2003] система восприятия основывается на прошлом опыте и возбуждении с целью поиска и получения вознаграждения за *прошлый* опыт.

Таким образом, прошлый опыт представляет собой уже не репрезентации, а сформированные ожидания в виде клеточных сборок Хебба [Dreyfus, 2007, p. 1152]. Рост опыта таким образом, становится *ростом степени различения* ожидаемых ситуаций. Фримен назвал их бассейнами

притяжения, соответствующими разного рода аттракторам, переходы в которых происходят за счет накопленной энергии напряжения, вызванной накопленными ожиданиями. «Попадание» в такого рода бассейн эмоционального отклика от действия, которое является наиболее соответствующим тому или иному аттрактору-ожиданию, приводит к уменьшению энергии напряжения системы. Таким образом, природа следует скорее принципу минимизации энергии ища наименее затратные пути.

Экспериментируя с кроликами, Фримен приводит в качестве примера действия, которые производит кролик: или он съедает морковку или уносит, если чувство голода имеет меньшее напряжение, чем желание спрятать еду ввиду положительного прошлого опыта, связанного с тем, что когда-то была похожая ситуация и было некоторое энергетически оптимальное действие.

Подводя итоги, можно выделить следующие выводы, сделанные Дрейфусом:

1. Существующие сегодня «общепринятые» подходы не способны решить проблему общего ИИ.
2. Решение следует искать в нейродинамических моделях хайдеггерианского ИИ, таких как модель Фримана.
3. Символьные методы необходимо рассматривать как своего рода «надстройку» над тем, что происходит на уровне нейродинамического субстрата, но не является его базовой функциональностью.

Подводя итоги, философ говорит о том, что хотя теоретическая возможность создания сильного ИИ возможна в парадигме воплощенного подхода, но существующее сегодня положение дел далеко от этого [Dreyfus, 2007, p. 1157]:

«...формирующийся воплощенный подход может стать шагом на пути к подлинно экзистенциальному ИИ. Хотя <...> все это далеко от программирования человеческого интеллекта».

И одной из главных задач на этом пути является проблема фрейма. Более того, для реализации воплощенного ИИ по мнению Дрейфуса требуется [Dreyfus, 2007, p. 1160]:

«...модель нашего особого способа быть внедренным и воплощенным таким образом, чтобы то, что мы испытываем <...> То есть мы должны были бы включить в нашу программу модель тела, очень похожего на наше, с нашими потребностями, желаниями, удовольствиями, болью, способами передвижения, культурными происхождениями и т.д.».

В качестве оппонентов такой точки зрения следует прежде всего привести таких авторов как Дениел Деннет, Джерри Фодор и, частично, Дэвид Чалмерс.

По мнению Чалмерса, реализация сильного ИИ, имплементирующего в себя сознание вполне возможна. В связи с чем он подвергает критике аргумент Китайской комнаты.

Свой тезис о возможности создания сильного ИИ философ строит на тезисе о существовании принципа организационной инвариантности. Суть принципа состоит в том, что [Чалмерс, 2015, с. 311]:

«...если имеется система, обладающая сознательными переживаниями, то любая система с точно такой же высокодетализированной функциональной организацией будет обладать качественно тождественными переживаниями».

Таким образом, возможна замена нейронов кремневыми чипами так, что эти гипотетические чипы были бы [Чалмерс, 2015, с. 310]:

«...наделены состояниями с теми же самыми схемами казуальных взаимодействий, что и у нейронов, эта система порождала бы точно такое же поведение».

Остается вопрос, насколько детализированной должна быть такая система? Не приведет ли это к требованию точного копирования, а не моделирования?

Необходимо отдать должное автору, развивая свою мысль на абстрактном примере реализации такой модели в виде комбинаторных автоматов (КоА), он пишет о том, что далеко не любой КоА может быть имплементирован в физическую систему. Так он указывает что, если взять КоА имеющий 1000 состояний, то шансы подобрать соответствующую ему физическую систему не превысят 1 из $(10^{1000})^{10000}$ [Чалмерс, 2015, с. 397]. На наш взгляд, такая вероятность фактически сводит к нулю возможность такой имплементации в практическом плане.

Относительно аргумента Китайской комнаты философ делает вывод о том, что аргумент Сёрля направлен «скорее против машинной интенциональности, чем против машинного сознания» [Чалмерс, 2015, с. 310]. Сёрль, по его мнению, считает, что сознание без интенциональности невозможно, а раз так, то система не обладающая интенциональностью не может иметь сознания. Чалмерс же считает, что проблема интенциональности не является критической для сильного ИИ. Более важной проблемой является проблема существования квалиа как основы сознательного опыта, основываясь на аргументах о «блекнущих» [Чалмерс, 2015, с. 316] и «скачущих» [Чалмерс, 2015, с. 332] квалиа.

Используя предлагаемый им принцип организационной инвариантности, заменяя одного человека в комнате изначально на одного китайца, с последующей заменой его нейронов крошечными демонами, каждый из которых снабжен набором инструкций и клочками бумаги (или кремневыми элементами), мы получим, по его мнению, полностью

инвариантную структуру. Такая структура будет образовывать [Чалмерс, 2015, с. 332]:

«...конкретную динамическую систему, казуальная организация которой прямо соответствует организации исходного мозга».

Оспаривая второй аргумент Серля относительно недостаточности синтаксиса для семантики, Чалмерс предполагает, что необходимо проводить различие между синтаксисом, как набором правил и их имплементацией в физическом объекте. Основным камнем преткновения в данном заочном споре становится выдвинутый Серлем и Стивенем Харнадом аргумент о том, что симуляция ментальных процессов не может быть идентична реальному процессу [Harnad, 1989]. Действительно в случае, если речь о некотором физическом процессе, таком как ураган, то такая симуляция останется симуляцией. Но в случае с феноменальными свойствами, казуальными петлями, являющимися свойствами системы, представляющей собой организованную инвариантность – дело обстоит иначе: «эти свойства представляют собой организованные инварианты» [Чалмерс, 2015, с. 408].

Основным недостатком, такого рода выводов, на наш взгляд является их интуитивность. Впрочем, и сам философ не отрицает этого факта. В частности, анализируя возражения противников сильного ИИ, он прямо указывает на проблемы креативности при следовании правилам и ограничения теоремы Геделя, но считает их недостаточными. Относительно первого аргумента его мысль заключается в том, что «указанное возражение не лишено оснований, хотя даже это не очевидно» ввиду того, что «класс вычислительных систем гораздо шире» [Чалмерс, 2015, с. 408].

Относительно теоремы Геделя аргументация звучит как отсутствие «оснований считать, что и люди могут знать об истинности соответствующих Геделевских предложений» [там же]. Мы сами не знаем - являемся ли мы формальной непротиворечивой системой или нет. А следовательно, почему бы

не допустить мысль о том, что так оно и есть: человек – это непротиворечивая формальная система, которая сама об этом не знает.

Гораздо более радикальную в этих вопросах позицию занимает Дэниел Деннет. Прежде всего речь идет о его попытках опровергнуть, или как минимум, придать некоторую относительность существованию «квалиа». Делается это прежде всего с целью прийти к несостоятельности «трудной проблемы» Чалмерса как таковой через идею «тонкого материализма» [Деннет, 2017, с. 172]. Смысл идеи заключается в том, что в нашем мозге нет чего-то похожего на «неуловимые» квалиа, но есть определенные структуры мозга, возникновение которых возможно объяснить и не прибегая к «квалиа». В рамках настоящего исследования критический анализ данного утверждения выходит за рамки настоящего исследования. Для нас важно отметить, что вне зависимости от того, чем бы эти структуры могли бы быть, на сегодняшний день нет какого-либо удовлетворительного понимания этого.

Относительно проблемы сознания философ не менее категоричен, делая ставку на принцип глобального пространства, выдвинутую еще Джоном Маккарти в 1959 году [Деннет, 2017, с. 210]. Развивая его мысль, Деннет предлагает «модель множества набросков» [Деннет, 2017, с. 212]. В основе этой идеи лежит «борьба за власть» над вниманием как ресурсом множества изначально бессознательных процессов, которые заявляют о себе путем «сообщений» в некотором общем глобальном пространстве-процессе. Тот процесс, который получает больше «славы», иными словами, оказывается более значимым в общем хоре процессов, становится сознательным и доминирующим для нашего «Я». Следует отметить, что при этом как такового «Я» - нет.

К сожалению, Деннет не дает четкого ответа на ряд ключевых вопросов таких как проблема описания восприятий от первого лица в терминах третьего лица; проблема понимания сущности того, что подставляет собой «Я». И это является слабым местом его теории, так как важность решения этих проблем определяется тем, что без их решения становится *крайне затруднительной*

возможность постановки строгого инженерного технического задания на проектирование системы общего ИИ.

Таким образом, сегодня нет достаточных оснований для того, чтобы говорить о возможности выхода ИИ-систем за пределы «подручности» по целому ряду оснований. Прежде всего речь идет о проблеме «хайдеггерянского ИИ», сформулированном Дрейфусом и недостаточности, по нашему мнению, аргументов Чалмерса и Деннета.

Общие выводы относительно онтологического статуса ИИ-систем:

- 1. ИИ-системы являются техническими объектами.*
- 2. ИИ-системы, как технические объекты, являются «подручными» человеку;*
- 3. На сегодняшний день отсутствуют условия для выхода ИИ-систем за пределы «подручности».*
- 4. Оставаясь «подручной», техника, и ИИ-системы, в частности, сегодня является одним из конституирующих современную культуру начал, приводя к росту алгоритмизации мыслительных процессов человека.*
- 5. Происходящие сегодня изменения скорее следует отнести к изменениям в способах и организации коммуникативных и мыслительных процессов самого человека, не отрицая опосредованного влияния техники на процесс культурных, социальных и ментальных процессов в обществе.*

§ Обоснование необходимости гибридного подхода как следствия вывода о «подручности» ИИ

Основными дискурсами, в рамках которых сегодня идет процесс осмысления техники, на наш взгляд являются дискурс о кризисе идей гуманизма и возникновение во второй половине XX века идей транс- и постгуманизма. К этому следует добавить наблюдаемый сегодня очередной возврат к идеям функционализма.

Концепция «пост» нашла свое выражение в «смерти субъекта» Мишеля Фуко, киборгах Донны Харуэй [2017], деконструкции классического субъективности и идеях деантропоцентризма, отвергающих тезис об исключительности человека [Шеффер, 2010]. Кроме всего прочего, такого рода идеи привели к значительному облегчению со стороны общества принятия самой возможности аутопозиса вне человека, к возможности самостоятельного и полноценного порождения новых смыслов в недрах машины. Иными словами, речь идет о, пусть и подсознательно, но признании в перспективе за ИИ права на обладание сознанием и свободой воли.

Примером такого рода феномена является появившийся сегодня дискурс о правах и этике ИИ. Но является ли это допущение не более чем иллюзией? В этой связи нельзя не согласиться с мыслью Ф.И. Гиренка [Гиренок, 2018]:

«Никогда человек не знал себя так плохо, как знает он себя сегодня, ибо сегодня знания о человеке скрывают от нас человека. Почему? Потому что они помещают человека в горизонт сущего, в котором существование человека отсылает к существованию вещей».

Установка на *ratio*, на авторитет позитивистской науки только усугубляет такого рода иллюзии.

Наступление NBICS-технологий, в значительной степени как обусловленное идеями трансгуманизма, только усиливает давление

техносферы. Можно сказать, что идеи Тейяра де Шадена и В.И. Вернадского сегодня деградировали до состояния техносферы. Причина этого кроется в алгоритмизации мышления самим человеком. В рамках настоящего исследования не представляется возможным дать глубокий анализ данной проблемы. Однако, на наш взгляд стоит выделить главную мысль, которая важна для определения роли и места символических методов в задачах общего ИИ: речь идет о значительной односторонности современного нефилософского дискурса (как минимум, за пределами философии) подменяющего тенденцию к алгоритмизации мышления успехами ИИ, тем самым только ускоряя процессы встраивания человека в техносферу.

Выход из этого может быть найден в переосмыслении роли, граничных условий процессов самообучения и автономности функционирования ИИ-систем. Как пишет Николас Луман [Луман, 2004, с. 66]:

«В этих условиях следует по-новому ставить и по-новому отвечать на вопрос о «резервной категории» (reserve category) для особенных, специфических человеческих качеств».

Сегодня речь идет о давлении на человека со стороны машин, требующих этих «резервных категорий». Но, природа этого требования, на наш взгляд, находится за пределами техники, которая, как мы уже выяснили остается «подручной» человеку. На наш взгляд, сегодня необходимо говорить не об обезличенной коэволюции человека и техники, а скорее о группах влияния, стоящих за теми или иными техническими системами.

Развитие когнитивных наук и явные успехи коннекционистских методов в ряде узких областей сегодня привели к значительному перекосу в сторону обсуждения проблем создания полностью автономных искусственных систем, попыткам полного исключения человека из жизненного цикла искусственной

интеллектуальной системы. Одной из причин этого, как уже говорилось ранее является алгоритмизация и формализация мышления самого человека, обусловленная степенью вхождения техники в разные сферы социального в качестве «организующего неорганического» [Хуэй, 2020, с. 330].

Второй, не менее важной причиной являются (ввиду уже самого по себе наличия рекурсии, реализованной в машине Тьюринга) попытки обнаружить феномена «творчества» или сознания у машинного интеллекта. Такого рода рекурсия порождает иллюзию кантовской рефлексии. Юк Хуэй пишет [Хуэй, 2020, с. 330]:

«...рекурсия функционирует подобно душе, которая возвращается к себе ради самопознания и при этом в любой момент выхода за свои пределы встречает ту или иную контингентность».

На наш взгляд, если бихевиоризм в ряде случаев пренебрегает рекурсивностью и практически полностью игнорирует контингентность, то последующий ему функционализм всего лишь дал больше «возможностей» для рекурсивности. Поддержка современной кибернетикой и когнитивной наукой идей функционализма базируется в немалой степени на идеях Алана Тьюринга, Людвиг фон Берталанфи и Норберта Винера. В работах этих исследователей понятие рекурсивности носит основополагающий характер в виде принципов обратной связи. Однако, говорить о том, что наличие обратной связи может преодолеть дуализм механицизма и витализма все же преждевременно. Говоря о «бедности кибернетической теологии» Юк Хуэй [Хуэй, 2020, с. 36] критикует, в частности, Симондона, говоря что неверно считать систему обратной связи обладающей «активной адаптацией к спонтанной целенаправленности» [там же].

Именно в этом, по мнению Хуэйя, видел Хайдеггер завершение философии - в «конкретизации систем» [Хуэй, 2020, с. 38]. Но является ли это концом философии или новым вызовом, стоящим перед человеком. Вызовом,

направленным на поиски выхода из состояния «поставы», переосмыслением отношения машины и человека?

Другим ярким примером такого положения дел можно считать идею «датаизма» придуманную журналистом Дэвидом Бруксом и впоследствии поданную Ювалем Ной Харари в качестве одной из основных угроз либерализму. Характеризуя это понятие Харари настаивает, что современные естественные науки «оспаривают» индивидуализм и свободу человека аргументируя это тем, что [Харари, 2021, с. 384-385]:

1. «Организм суть алгоритмы, и человек не индивидуум, он дивидуум. Иначе говоря, человек — это собрание алгоритмов».
2. «Алгоритмы человека не свободны». Они сформированы либо генами, либо средой, их решения либо детерминированы, либо случайны – но в любом случае они не свободны.
3. Из этого следует, что теоретически внешний алгоритм вполне способен познать меня лучше, чем я сам когда-либо даже в будущем познаю себя.

Более того, философ строит прогноз относительно того, что в будущем развитие когнитивных наук позволит «взломать человеческую сущность» [там же]. Впрочем, такая идея не является абсолютно новой, являясь продолжением функционализма в его понимании, например, Сёрлем. Обсуждение же практических результатов которых можно достичь, «взломав» нашу сущность также выходит за рамки настоящего исследования. В интересах нашего исследования, отметим, что, во-первых, на сегодняшний день эта проблема далека от решения, а во-вторых, даже признание простой «случайности» ставит необходимость учитывать субъективность человека как минимум в отношении целого ряда ноэтических проблем, в частности – ноэзиса и семиозиса как таковых.

Таким образом, несмотря на рост тотальности техники, и связанную с этим алгоритмизацию процессов мышления, сегодня необходима

корректировка исследовательских программ. Смещение фокуса исследований в область создания гибридных моделей является, по нашему мнению, более плодотворной задачей в долгосрочном периоде, чем игра в имитацию или развитие какого-то одного из технологических подходов.

Обоснование данного тезиса нами строится как минимум на пяти аргументах:

1. Аргумент экзистенциальный

«Всякое спрашивание есть искание. Всякое искание имеет заранее свою направленность от искомого» пишет Хайдеггер [Хайдеггер, 2015, с.5], таким образом *невозможно понять что-то, не отталкиваясь от субъекта и предмета поиска*. В связи с чем возникает вопрос: если человек отсутствует, то кто и для чего задает предмет поиска? На наш взгляд, нельзя забывать, что создание каких-либо физических интеллектуальных систем обусловлено потребностями человека. Именно человек является главным «Заказчиком». Кроме того, для понимания проблем ИИ следует выйти за пределы предметной области как таковой, аналогично тому, как, по словам Хайдеггера, для понимания биологии следует выйти «за пределы организма и жизни» [Хайдеггер, 2015, с.10]. Мы должны погрузиться в тот мир, и сформировать непосредственно основопонятия, дающие «толкование этого сущего на основоустройстве его бытия» [там же].

Таким образом, *без анализа и объяснения искусственного интеллекта как погруженного в мир человека невозможно ни создание прочной онтологической и эпистемологической базы исследований, ни точного формирования самой постановки задачи в ходе проведения каких-либо исследовательских программ в данной области*. Следуя этой логике сам факт возникновения какого-либо сознания или хотя бы чалмеровского зомби невозможен без «погружения» его в какое-то бытийное окружение. Не имея

его «отражений» в мире, невозможно понять, что из себя представляет исследуемый феномен, какого рода сущностью он является.

2. Аргумент коммуникативный

Прежде всего данный аргумент касается понятия истины и ее достижения. Существующее на сегодня множество теорий истины свидетельствует об отсутствии единого понимания в данном вопросе. Тем не менее, в соответствии с прагматическим принципом понимания, истина — это убеждение «с которым в конечном счете, суждено (is fated) согласиться всем, кто исследует» [Пирс, 1996].

С.В. Шачин определяет содержание коммуникативного разума Хабермаса в двух измерениях:

Во-первых, в «измерении индивидуальных сознаний» как результат хода индивидуального психологического развития, заключающийся в формировании объективного, социального и субъективного «миров» [Шачин, 2019, с. 154].

Во-вторых, в измерении «коллективного сознания общества» [там же], как результата универсализации своего индивидуального сознания до уровня общества. Развивая эту мысль в рамках понятия коммуникативного действия, Хабермас определяют истину как результат опосредования рациональностей личности и рациональностей высказывания.

Таким образом, истина представляет собой прежде всего продукт согласия, продукт «коллективного движения к истине» [Шачин, 2019, с. 155].

Соглашаясь с таким пониманием вопроса, на наш взгляд, *реализация каких-либо интеллектуальных систем невозможна вне рамок мира, и, прежде всего, мира человека.*

3. Аргумент социальный

По мнению Латура [Латур, 2020] сегодня уже невозможно просто провести ревизию того, что составляет сегодня социальные науки. Без анализа социальных связей становится невозможным объяснить целый ряд феноменов в психологии, биологии географии и т.д. Кроме того, сегодня необходимо говорить скорее о сборках, нежели об изолированных областях знаний.

Понятие «социальный» по отношению к тому или иному феномену определяется им как [Латур, 2020, с.11]:

«некое устойчивое состояние, комплекс связей, который потом может быть использован для описания какого-то другого феномена».

Это «социальное» является естественным компонентом в составе «сборок» различных наук, и прежде всего когнитивных.

Таким образом, попытки изолирования, пусть даже в исследовательских целях, человека и общества от физических интеллектуальных систем, не могут быть в достаточной степени удачными: в этом случае становится невозможным понять ни цели функционирования таких систем, ни мир, в котором они должны действовать.

4. Аргумент герменевтический

Одной из важнейших задач в области когнитивных наук и искусственного интеллекта является проблема понимания смысла текста. Это важно как минимум для организации вербальной коммуникации в системе «машина-человек». Важность решения этой проблемы является очевидной и подтверждением тому служит огромный пласт публикаций с заявлением о наличии существенных успехов в этой области. Однако, более детальный анализ показывает, что полученные результаты имеют довольно

ограниченный результат, особенно на фоне того количества усилий, которые были затрачены. Особенно данный факт заметен в области диалоговых систем. В данном случае заслуживают внимания даже не те или иные успехи или провалы в прохождении Теста Тьюринга, а проблемы организации действительно эффективной коммуникации в ходе процесса диалога машина-человек.

И здесь с особой очевидностью возникает герменевтическая проблема. Суть ее заключается в том, что все усилия в этой области, достигаемые с помощью коннекционистских подходов, лежат скорее в области филологии. Задачи же герменевтики, направленные на решение вопроса о том, как именно проходит процесс понимания текста, носят фактически противоположный характер. На это указывает Гадамер [Гадамер, 1988], говоря, что проблема герменевтики состоит не в правильном пользовании языком, а в «истинном взаимопонимании по тому или иному поводу».

И здесь возникает два труднопреодолимых препятствия, которые не позволяют говорить о возможности исключения человека из этого процесса.

Во-первых, это недостаточность одного только машинного интеллекта для решения задач понимания смысла текста как это декларируют, например BERT- модели [Delvin, 2019]. Гадамер в работе «Истина и метод» указывает, что «язык не является продуктом рефлексивного мышления, но участвует в реализации того отношения к миру, в границах которого мы живем». Мир не является предметом языка, «скорее то, что является предметом высказывания, всегда уже окружено мировым горизонтом языка» [Гадамер, 1988, с. 520]. Следовательно, использование только коннекционистских моделей, представляется недостаточным. *Модель должна быть расширена по меньшей мере до полной семиотической модели в понимании Д.А. Поспелова* [Поспелов, 2011].

5. Аргумент технологический

Приводя доводы в пользу потенциальной возможности создания искусственной системы, основанной на принципе организационной инвариантности необходимо констатировать, что существующий сегодня уровень развития техники не позволяет говорить даже о сколь-нибудь значимом качестве реализации феномена сознания в рамках тех или иных компьютерных систем. Важно отметить, что речь, в данном случае, идет не о росте производительности или степени параллельности вычислительных процессов, а о предполагаемой сложности такого рода системы. Прежде всего, речь идет об отсутствии понимания того, что представляет собой не только сам по себе феномен сознания, но и понять связь сознание-мозг. К.В. Анохин пишет [Анохин, 2021]:

«Для науки о мозге нет более важной задачи, чем объяснить природу субъективного мира – решить многовековую проблему разум-мозг (mind-brain problem). Мы можем обозначить ее как “главную проблему” (MAIN problem – Mind-brAIN problem), стоящую перед современной нейронаукой. <...>Эта проблема сегодня нередко обозначается как «трудная проблема» сознания.»

В качестве «трудной проблемы», автором понимается то, как она была поставлена Чалмерсом как [там же].

«...вопрос о том, как и почему физические процессы порождают субъективное сознание».

Решение данной проблемы в настоящее время отсутствует. Следовательно, сама по себе постановка задачи на реализацию сознания в рамках машинного интеллекта отсутствует как таковая.

Таким образом: *любая система ИИ, приближающаяся к имитации рационального поведения, должна быть системой гибридного ИИ.*

§ Определение граничных условий существования ИИ-систем

Определяя ИИ-системы как подручные человеку, и придя к выводу о необходимости гибридного подхода с необходимостью встает вопрос о существовании граничных условий. Исследуя проблемы самообучения ИИ-систем и автономности принятия ими решения в ходе их функционирования, возникает вопрос об оценке получаемых решений сточки зрения их ценности и применимости с учетом требования относительной несамостоятельности техники.

Определение, данное в «Энциклопедии и эпистемологии науки» характеризует граничные условия с нескольких сторон [Энциклопедия, 2009, сс.158-159]:

1. как понятие, характеризующее степень применимости определенных теоретических положений при объяснении конкретных эмпирических ситуаций;
2. как границы эффективности самой теоретической системы;
3. вводит требование о возможности любого объяснения с учетом некоторых общих законов, принятых научным сообществом;
4. граничные условия применимости научной теории определяются ее соотношением с какой-то другой концептуальной системой, выступающей в роли конкурента по отношению к данной.

На основании полученных ранее выводов о «подручности» ИИ-систем как относящихся к техническим объектам, требованиями гибридного подхода, на наш взгляд с необходимостью следует, что концептуальная система машинного интеллекта должна в обязательном порядке соотноситься с концептуальной системой человека как инстанцией более высокого порядка. Именно

концептуальная система человека и общества должна задавать те законы, на основании которых должно осуществляться объяснение и оценка значимости полученных ИИ-системой результатов.

Такое положение дел не отменяет возможности первичного самостоятельного генезиса новых понятий, знаков и правил со стороны машинного интеллекта, но требует дальнейшего согласования и утверждения полученных результатов человеком.

Отдельным вопросом остается определение условий, ограничивающих вариативность самого по себе генезиса. Примером такого рода ограничений может выступить требование о соответствии новых знаков языковым правилам и т.д. Более подробный анализ возможных условий и правил генезиса новых понятий, знаков представлен в Главе 4 настоящего исследования.

Таким образом, граничными условия, которым должен удовлетворять процесс самообучения, нарушение которых приводит к невозможности существования ИИ-системы или потере ценности полученных результатов однозначно задаются концептуальными системами человека в понимании их В.А. Смирновым.

Глава 3 Основные аспекты применимости символьных методов при проектировании ИИ- систем

Анализируя место техники, а также причины невыхода ИИ за рамки «подручности», можно сделать вывод, что рассматриваемое до этого противостояние символьных и коннекционистских методов не является основной проблемой. Как было показано выше, оба подхода имеют как свои достоинства, так и недостатки. И если сегодня коннекционистские методы имеют явное преимущество, то связано это прежде всего с ростом вычислительных мощностей, а не чем-то большим. Анализ достигнутых успехов, а также обзор конференций показывает, что основными тенденциями остаются, по сути, изолированные исследовательские программы, нацеленные на усовершенствование уже существующих методов. Простой сравнительный анализ обоих подходов вряд ли сможет добавить к этому что-то существенно важное. Попытки построения гибридных систем, как попытка объединения различных технологических подходов (например, в работах Домингоса) также носят ограниченный характер. При этом, практически не принимается в расчет важность построения непосредственно человеко-машинных систем, в которых обе составляющие части органически дополняют друг друга.

Прежде всего, для получения ответа на вопрос о роли символьных методов в задачах общего ИИ, необходимо сформулировать проблему, решение которой и могло бы являться ответом на поставленный вопрос. В нашем случае проблему можно сформулировать как:

Проблему необходимости и достаточности символьных методов для построения системы общего искусственного интеллекта.

В связи с чем, представляется принципиальным исследовать в теоретическом плане следующие аспекты, раскрывающие суть поставленной нами проблемы:

1. *Эпистемологический*. Существует ли принципиальная возможность инженерной реализации актов непосредственного генезиса новых понятий, знаков и правил вывода как необходимых условий, конституирующих возможность построения собственных концептуальных структур ИИ-системой;
2. *Гносеологический*. Существует ли принципиальная возможность обеспечения автономности процессов первичного генезиса новых понятий, знаков и продукционных правил со стороны ИИ-системы;
3. *Онтологический*. Как (каким образом) могут существовать символичные методы в рамках решения проблем 1 и 2.

Следует отметить, что для решения поставленной нами проблемы не всегда необходимо найти конкретное практическое решение, математическую функцию или что-то аналогичное. В нашем случае уже само по себе прояснение сути проблемы может являться принципиальным приближением к ее решению. На наш взгляд, в качестве отправных точек, следует принять в расчет два тезиса, которые по необходимости будут определять ход дальнейшего исследования:

Во-первых, тезис об отсутствии сколько-нибудь приемлемой теории сознания, техническая реализация которой дала бы возможность машинному интеллекту приобрести черты как минимум искусственного. В качестве примера, иллюстрирующего такое положение дел могут служить попытки Facebook по обучению нейронной сети на выборках трех разных типов [Quach,2022], или исследования в области совместного использования нейронных сетей и деревьев решений [Гридин, 2017; Фридман, 2019]. Оба примера демонстрируют тот факт, что в практическом плане не происходит

попыток создания новых, направленных на целенаправленное проектирования архитектур, потенциально воплощающих сознание. Вместо этого, осуществляются попытки комбинаторного, имеющего экстенсивный характер, расширения количества параллельно и последовательно функционирующих подсистем в соответствии с идеями Деннета.

Во-вторых, тезис о «подручности» современной техники, о котором говорилось выше. В целях дальнейшего исследования следует отметить, что работы Эрнста Каппа, Мартина Хайдеггера, Жильбера Симондона [Симондон, 2011] и Бернара Стиглера [Stiegler, 2016] говорят в большей степени об одной стороне проблемы «подручности»: о проблеме связанной скорее не с внутренними несовершенствами тех или иных методов, а с самим положением, занимаемым техникой как таковой с социокультурной точки зрения. Границы «подручности» в этих работах определяются с одной стороны самим способом, каким техника возникла, а с другой - той свободой в замещении труда человека принятию решений, которой наделяет ее сам человек.

Далее, необходимо упомянуть две важные проблемы, на которые обращает внимание В.А. Смирнов в своей работе «О перспективах анализа учения И. Канта о праве и морали средствами современной логики».

Во-первых, речь о проблеме соотношения природы и свободы. Размышляя о возможности применения логики к анализу норм права и морали, автор пишет [Смирнов, 1990, с. 68]:

«Великая проблема философии - совместить принципы детерминизма, на которых основано естествознание (да и не только естествознание, но и социальные науки), с принципом свободы, человеческой личности, человеческого достоинства. <...> С одной стороны, задача состоит в том, чтобы построить такое понимание детерминизма, закона, согласно

которому сами естественные и социальные науки допускали бы существование автономных систем, обладающих свободой.»

Одной из отправных точек для поиска путей решения этой задачи является, по его мнению, мысль о доступности для логического анализа не только акта суждения, но и акта волеия, имеющие свою особую форму императивов. Уже само по себе наличие таких форм, по Канту, дает основание для логического анализа, так как «формальная логика занимается любыми формами» [Смирнов, 1990, с. 69]. Следовательно, свободные волевые акты также можно подвергать логическому анализу.

Еще одним следствием из этого является следующий тезис [там же]:

«Не все есть знание. Мы иногда сводим многие акты, не являющиеся познанием, к познавательным, например, иногда и мораль, и право считаем знаниями. Однако существует много других вещей, отличных от знания.»

Таким образом, интеллектуальная деятельность состоит не только из познавательной, но и допускает другие ее формы, такие как императивы и предписания.

Еще одним важным замечанием, является тезис о том, что принципиально важным является в ходе дальнейшего исследования избежать ловушек корреляционизма. В первую очередь, речь идет об опасности излишнего психологизма, с одной стороны и спекулятивного идеализма – с другой. В случае принятия позиции радикального психологизма становится невозможным существование независимых от субъекта форм и объектов знания. В случае принятия позиции спекулятивного идеализма становится невозможным «прорваться к реальности, как она есть сама по себе, независимо от любых актов субъективности» [Мейясу, 2008].

§ Эпистемологический и гносеологический аспекты

Проблематика генезиса новых понятий и знаков

Исходя из тезисов Ньюэлла-Саймона, с неизбежностью встает вопрос о символах (знаках): их генезисе, способе их существования. Следовательно, проблему, обозначенную нами как эпистемологическая, можно сформулировать в виде двух вопросов:

- 1) как, каким образом, может происходить концептуализация понятий;
- 2) как, каким образом, может происходить генезис новых знаков.

Отправными точками дальнейшего исследования являются тезис о том, что, *существование символических методов возможно только в рамках тех или иных онтологий*.

Необходимость тезиса является следствием целого ряда проблем, в рамках которых говорить об онтологии в ее классическом понимании сложно по нескольким причинам.

Во-первых, ввиду существования целого ряда проблем, связанных с утратой со стороны классической онтологии своего первоначального смысла и появления множества несводимых друг к другу научных картин мира [Невважай, 2019].

Во-вторых, по причине, существующей сегодня проблем относительно понимания смысла самих региональных онтологий [Гуссерль, 2009, с.489]. Возникновения, например, концепций дисциплинарных (как некоторого «среза», задаваемого той или иной научной парадигмой [Невважай, 2019]) и структурных [Цофанс, 2009] онтологий. В рамках настоящего исследования, под многообразием онтологий будем понимать прежде всего дисциплинарную и структурную концепции.

Таким образом, представляется необходимым прежде всего провести анализ отношений в связке онтология-символ.

В рамках настоящего исследования знаки понимаются в соответствии с их определением, данным Чарльзом Пирсом [Пирс, 1983].

По мнению Квентина Мейясу, основным результатом давления на философию со стороны экспериментальных наук стало окончательное закрепление *корреляционизма*. В своем докладе «Время без становления» [Мейясу, 2008], под термином «корреляционизм» он понимает большой пласт, главным образом, постмодернистских учений XX века. Суть их сводится к тому, что [там же]:

«...нет объектов, событий, законов, сущностей, которые не были бы всегда уже скоррелированы с точкой зрения, субъективным доступом».

Вернув реализм в философский дискурс, противостоять этому должно было движение, получившее название «спекулятивный материализм (реализм)» [там же]. Философ полагает существование двух форм такого корреляционизма, таких как [там же]:

«...трансцендентальная, заявляющая, что есть некоторые универсальные формы субъективного познания вещей, и постмодернистская, отрицающая существование любой субъективной универсальности».

Не менее важным для настоящего исследования тезисом философа является *фактичность*, понимаемая как [там же]:

«...отсутствие *причины* для всякой реальности. <...> Можно говорить только об условной необходимости, и никогда об абсолютной».

Отсылая к картезианскому *cogito* как к условной, а не абсолютной необходимости, тем самым утверждается несостоятельность какой-либо субъективной метафизики как абсолютизации того или иного доступа к миру.

Корреляции являются контингентными: возможно как их возникновение, так и исчезновение. Отсюда философ делает вывод [там же]:

«Я могу мыслить независимый от всякого мышления X». Абсолютизация фактичности приводит к равным возможностям «порядка и беспорядка, становления и вечности».

Но быть фактичным не означает хаос. Условия фактичности, по мысли философа, подразумевают строгие условия бытия. Таким условием является *непротиворечивость*. В противном случае невозможны изменения, ибо они уже содержали бы в себе всё и не было бы причин для изменений. Тем не менее противоречие является необходимым атрибутом как христианского мировоззрения - как чуда, так и гегелевского Абсолюта. Это попросту удобно. И в этом одна из причин, по которой философ «не доверяет» метафизике в целом.

Из всего вышесказанного, по нашему мнению, можно с необходимостью сделать следующие выводы:

1. Доступ к миру не ограничен исключительно человеком и может существовать в рамках машинного интеллекта. Критикуя коррелятивизм, Мейясу открывает перспективу самостоятельного познания мира со стороны машинного интеллекта. Такой тезис, вместе с тем, заново поднимает целый ряд проблем, таких как проблема квалиа и т.д.
2. Принципиально возможно существование иной реальности, не зависящей от человека, но согласующейся с ней. Такой вывод можно сделать только в том случае, если мы соглашаемся с независимостью

существования мира в-себе от мира для-нас. Этот тезис будет важен в последующем при обсуждении гносеологического и коммуникативного аспекта.

3. Четвертое царство Дессауэра, в этих обстоятельствах, оказывается связанным не с Абсолютом, а скорее с нонсенсом Жилия Делеза.

Как нам представляется, анализ идей 1 и 3 может быть интересным предметом отдельной исследовательской программы и выходит за рамки настоящего исследования.

Во многом аналогичные идеи ранее были выдвинуты Аленом Бадью. Прежде всего речь идет об отказе от сущего как одного, как многое-без единого, что приводит к существованию только региональных онтологий [Долманов, 2012]. Кроме того, Бадью настаивает на присутствии аксиоматизации как необходимой. В частности, Долманов приводит слова Бадью [там же]:

«...единое и многое – это не «единство противоположностей», поскольку первое не есть, тогда как второе является самой формой всякой презентации сущего. Аксиоматизация требуется для того, чтобы, предоставленное неявному правилу счёта, многое породилось без понятия, то есть без предположения бытия-одного.»

Важно отметить, что при всем этом *субъекту остается право выбора связей элементов и событий «на свой страх и риск»* [там же].

Таким образом, Мейясу и Бадью дают некоторые основания для придания большей легитимности региональным онтологиям, обосновывая возможность их независимого существования.

В своей работе «Понятия и объекты» Рей Брассье подвергает резкой критике получившие сегодня широкое признание идеи корреляционизма, основными источниками которых он считает Иоганна Готлиба Фихте, Бруно

Латура, Теодора Адорно, Луи Альтюссера, Жиля Делеза и даже Грэма Хармана. Суть этих идей заключается в том, что [Брассье, 2017, с. 239]:

«...для корреляциониста невозможно отделить субъективное от объективного <...> В конечном счете корреляционизм – это не определенное философское учение, а расплывчатая и весьма разносторонняя стратегия дефляции традиционных метафизических и эпистемологических вопросов путем редукции проблематики «бытия» и «знания» к сцеплениям культурной формы, политической борьбы и социальной практики».

Выход из создавшейся ситуации философ видит в необходимости «вновь установить главенство связки «эпистемология-метафизика»» [Брассье, 2017, с. 239]. Несовпадение реального и чуждости объекта в научной репрезентации, по его мнению, не является проблемой, а напротив – является внутренней сущностью процесса научного познания, которая и заставляет вносить корректировки в существующие теории, обеспечивая их развитие. Развивая мысль философа, ответом на Вопрос 1 о возможности концептуализации понятий в рамках процессов внутри машинного интеллекта являются следующие тезисы:

1. Факт обнаружения несоответствия сложившейся концептуальной структуры и получаемой ИИ-системой информации является главным источником получения знаний об окружающей среде.
2. Полученные выводы и новые знания об окружающей среде не должны противоречить знаниям, полученным непосредственно от человека.

Даже не забегая так далеко и не претендуя на создание новых законов робототехники, в рамках настоящего исследования важно подчеркнуть мысль

философа о том, что попытки элиминирования связки онтология-метафизика, являются как минимум преждевременными.

Такая же неоднозначность, существует и в отношении связки метафизика-символ.

Как бы ни показалось очевидным на первый взгляд, но вопрос о проблеме символа внутри метафизики также не является предметом общего согласия: «метафизика есть наука, стремящаяся освободиться от символов» пишет Анри Бергсон в своей статье «Введение в метафизику» [Бергсон, 2019, с. 133]. Философ обосновывает это тем, что истинное знание по своей природе интуитивно, способно переносить нас внутрь предмета, чтобы слиться с тем «невыразимым», что есть в предмете. Роль символов, по необходимости, сводится к вспомогательной роли [Бергсон, 2019, с. 131]:

«они [символы] сообщают мне лишь о том, что сближает ее [сущность] с другими, а не принадлежит собственно ей».

Знаки, будучи «частичными обозначениями», а не «реальными частями» имеют смысл только внутри аналитических функций позитивной науки. Истинное же знание имеет интуитивный характер, обладая простотой, целостностью и представляет собой нечто «абсолютное, увиденное изнутри». Карен Свасьян пишет, что Бергсон непременно согласился бы с Гансом Файхингером в определении символа как фикции, правда с некоторыми оговорками [Свасьян, 2010, с. 53]:

«... если для Файхингера фикция - альфа и омега нашей мыслительной жизни, то Бергсон рассматривает ее как абберацию духа, культивируемую наукой; задача философии – вырваться из сети фиктивных символических мыслеплетений и достигнуть абсолютного».

Совершенно противоположной точки зрения придерживался Эрнст Кассирер. «Эмпирический методизм Спенсера» и «интуитивный методизм Бергсона» с порога отвергаются Кассирером как «*contrdictio in adjecto*» [Свасьян, 2010, с. 59]. В качестве наглядного примера Свасьян предлагает представить себе иррациональное или трансцендентальное число. Как, каким образом, возможно не имея идей, философствовать над конкретными фактами? В качестве аргументов, которые могли бы прозвучать со стороны Кассирера, Свасьян приводит следующую цитату [Cassirer, 1956, p. 94]:

«Сама история исчезла бы в беспредельной массе разрозненных фактов, не имея она общей структуральной схемы, позволяющей ей классифицировать, упорядочивать и организовывать эти факты».

Таким образом, *символизм является неотъемлемой частью метафизики и исследование вопросов генезиса символьных систем напрямую связано с вопросами метафизики.*

Начиная с момента своего возникновения, метафизика Аристотеля неизменно вызывала целый ряд дискуссий между рационалистами и эмпириками. Как известно, одним из основных камней преткновения является вопрос о скептицизме относительно реализма в аспекте достоверности научного знания. На своем начальном этапе, начиная с Нового времени, вопрос этот был более-менее однозначен и носил объективный характер, закрепленный в итоге Христианом Вольфом в качестве таксономии [Иванов, 2017]. В основе концепции Аристотеля лежал реализм, закрепляющий объективный характер высказываний о мире. Такая постановка проблемы могла бы в значительной степени удовлетворить сторонников искусственного интеллекта так как давала все основания для придания объективного и реального статуса сторонникам формально-логического подхода в построении онтологий.

Начиная с работ Декарта, Юма, Канта остро встал вопрос о необходимости критического переосмысления взглядов на рационализм. Со всей очевидностью встал вопрос о необходимости критического переосмысления основных положений классической метафизики. Появление уже в работах Декарта психического в качестве совокупности феноменальных актов, усиленное критикой классической метафизики Кантом, в итоге привело к пониманию того, что «связь субъекта и объекта опосредована феноменальными данностями» [Иванов, 2017, с. 37].

Важно отметить, что феноменальность актов восприятия и связанная с этим констатация их субъективности в значительной степени повлияла и на сторонников эмпиризма. В частности, по Кассиреру, эмпиризм не учитывает того, что единичное восприятие существует исключительно [Свасьян, 2010, с. 86]:

«...в силу своей принадлежности к «я» <...> возникает не задним числом из соединения множества восприятий, но изначально присутствует в каждом из них».

К сожалению, очередная попытка вернуть реализм, произведенная такими философами как Сол Крипке, Хилари Патнэм и Рей Брассье, не привела в полной мере к решению проблемы. Причиной этого является недостаточность обоснования необходимости внешних объектов как таковых. Здесь нельзя не согласиться с мыслью Д.В. Иванова [Иванов, 2017, с. 39]:

«Для того чтобы обосновать возможность полноценного реализма, требуется не просто показать зависимость ментальных репрезентаций от внешних, контингентных положений дел в мире, требуется прояснить природу концептуального познания.»

Как уже отмечалось выше, мы до сих пор не располагаем удовлетворительной концепцией, объясняющей то, каким образом сосуществуют мир и сознание. Отсюда, одной из ключевых проблем, имеющих непосредственное отношение к теме настоящего исследования, остается замкнутый круг, образованный кантовской мыслью о том, что [Кант, 1994, с.89]:

«...созерцание и понятия образуют элементы всего нашего познания, так что ни понятия без соответствующего им некоторым образом созерцания, ни созерцания без понятий не могут дать знание».

Одной из последних попыток преодолеть эту проблему являются исследования Джона Макдауэлла [McDowell, 1994]. Концепция Макдауэлла базируется на двух основных положениях.

Во-первых, наш опыт не располагает рациональными концептами, а следовательно – наши априорные рациональности опосредуют наши отношения с миром.

Во-вторых, апеллируя ко «второй природе» Г-Х. Гадамера, философ говорит о языке, как о внешнем мире, который также концептуализирует наше мышление, привнося в него в безусловной форме собственную рациональность. Как несложно заметить, в этом легко прослеживается влияние аналитической философии.

Впрочем, с концепцией Макдауэлла не согласны целый ряд исследователей таких как С. Глендиннинг и М. де Гэйнсфорд [Иванов, 2017], Хуберт Дрейфус [Scheer, 2013] и т.д. Так, Глендиннинг и де Гэйнсфорд указывают на факт игнорирования различения восприятия самого объекта и видимости объекта. Дрейфус, в свою очередь, говорит, что «ментальное» и есть, то, что рождает концепты. Критически к идеям Макдауэлла относился и Патнэм. Теория Макдауэлла, по общему убеждению, не решает проблему

рациональности. По общему мнению, идеи Макдауэлла, скорее, можно отнести к скрытому идеализму.

Возможно, к этому следовало бы добавить еще одну проблему - квалиа. Аналогично тому как Патнэм предложил эксперименты с «Двойником Земли» и «Мозгами в колбе», возможно следовало бы поставить такой же эксперимент с «Двойником человека», и «Машиной в колбе» для того, чтобы попытаться представить то, как могут функционировать квалиа исходя из существующих сегодня технических возможностей по мониторингу окружающего мира, таких как видеокамеры, микрофоны, тактильные датчики и т.д.? Какова могла бы быть метафизика и онтология существ, наделенных иными, но технически понятными нам квалиа, но при этом обладающих сознанием? На эту же мысль, хотя она и не является решением проблемы в полной мере, наталкивает высказывание Канта, приведенное Хайдеггером для иллюстрации «метода естествоиспытателей» [Хайдеггер, 1997]:

«Им «открылся свет что разум проникает лишь в то, что сам осуществляет по своему проекту, что он с принципами своих суждений, сообразными постоянным законам, должен предшествовать природе и принуждать ее отвечать на свои вопросы, а не тащиться за нею, словно на поводу». «Заранее наброшенный план» природы вообще впервые устанавливает первоначальный бытийный состав (Seinsverfassung) сущего, с которым должно соотноситься любое исследующее вопрошание.»

Впрочем, даже без анализа проблемы квалиа, положение дел в области построения онтологий остается далеким от своего решения. Подтверждением этого является вся история развития формальных онтологий от момента своего возникновения в работах Эдмунда Гуссерля и до сегодняшнего дня, в виде разного рода дисциплинарных [Невважай, 2009] онтологий т.д..

Сам по себе термин «формальная онтология» впервые появился в работах Гуссерля в 1901 году в «Логических исследованиях». По мнению

В.Л. Васюкова, Гуссерль не использует какой-либо формальный аппарат, его работы представляют собой скорее «анализ с целью выяснения интуитивных оснований и понятий» [Васюков, 2006, с. 306]. На сегодняшний день существует несколько концепций онтологий, и далеко не все из них имеют гуссерелевскую природу.

Среди них Васюков выделяет несколько таких концепций. Прежде всего, речь идет о Brentanovской традиции, определяемой им как «дескриптивная онтология сознания - онтология ментальных структур» [там же]. Последователями этой традиции являются Эдмунд Гуссерль, Казимеж Твардовский и ряд других исследователей. Основной установкой такого подхода является тезис о наличии таксономии разного рода конституэнт и различных форм их отношений. Сам Brentano выделял три вида онтологий: 1) онтология вещей, 2) онтология состояния дел, 3) онтология оценок. Позднее такая концепция была расширена Твардовским до понимания того, что вещи могут быть реальными и нереальными, возможными или нет: но все они являются объектами наших психических актов. К этому следует добавить концепцию региональных онтологий Гуссерля.

Основной смысл «не-гуссерелевской» [Васюков, 2006, с. 308] концепции формальных онтологий Нино Кокьярелла заключается в том, что формальная онтология должна изучать [там же]:

«логические характеристики предикаций, квантификации по переменным и различные теории универсалий».

Особое место в этом списке занимает онтология Станислава Лесьневского, представляющая собой одновременно и логическую систему формальной онтологии, «гибридный» характер которой [Васюков, 2006, с. 308]:

«...можно рассматривать как некий вид систем формальной онтологии, в которой характеристика онтологических понятий предполагает не теоретико-множественные и не мереологические термины, но лишь термины исчисления имен (в терминологии Е. Слупецкого)».

Таким образом, общим среди всех этих подходов остается, как минимум - открытый вопрос о механизмах построения онтологий, а как максимум - вопрос о самой природе производства знания, основанного на этих онтологиях, его достоверности. Кроме того, все эти подходы практически полностью оставляют за пределами исследований вопрос о природе собственно механизмов генезиса концептуальных понятий и связанных с ними знаков. Сегодня, в явном или снятом виде они оставляют именно за человеком право первенства в этом вопросе.

Следовательно, мы вновь возвращаемся к проблеме корреляционизма как отправной точки. Прежде всего речь идет о необходимости ответа на вопросы о том, является ли тождество и различия в объекте тождеством и различием в понятиях. Брассье называет это опасностью перехода от корреляционизма к идее концептуального идеализма [Брассье, 2017, с. 243]. Корень зла в этом вопросе философ, опираясь на Девида Стоува, видит в неверном и двусмысленном словоупотреблении «вещи», которое идет еще от Джорджа Беркли и его 23-го параграфа «Трактата о принципах человеческого познания» [Брассье, 2017, с. 245]:

«...это удается благодаря двусмысленности слова «вещи»: постигаемые и воспринимаемые вещи (то есть предметы, *ideata*) и вещи как таковые, *simpliciter* (то есть физические объекты). Именно это различие Беркли и пытается разрушить. Но он не может отбросить его с самого начала, не прибегнув к *petitio principii*, — отрицание этого различия и метафизическое утверждение, что существуют лишь сознания и их

предметы, должны быть следствиями рассуждения Беркли, а не его посылками.»

Таким образом, по его мнению, происходит подмена понятий со стороны Беркли. Возникшая в результате идеалистическая предпосылка, названная Стоувом «Перл» [Stove, 1991, p. 135-178], звучит как [Брассье, 2017, с. 247]:

«Нельзя познать независимую от разума реальность, не познавая ее.

Следовательно, нельзя познать независимую от разума реальность.»

Однако, как отмечает Брассье, независимость и недоступность являются совершенно разными понятиями. Существование чего-то независимо от нас никак не приводит к непознаваемости.

Исследование данного вопроса во всех его аспектах, утверждает Брассье, открывает возможность для понимания того, как могут сосуществовать друг с другом понятия и реальные объекты. Отсюда можно получить ряд выводов, которые могут представлять собой интерес для решения задачи общего ИИ. В качестве отправных точек для этого нам потребуются следующие тезисы философа:

1. Положение о том, что если нечто нельзя полагать без условий его полагания, то это не делает истинным тот факт, что слово (имя) не может полагаться как неполагаемое (если нет условия правильного использования понятия с ним связанного). При этом само понятие, соответствующее этому слову, конечно же требует условий своего полагания.
2. Понятие обладает значением исключительно благодаря мыслительной деятельности человека. Само по себе слово в структуре семиотической ситуации может существовать без понятия, но понятие конечно же нуждается в каком-то имени.

3. Сам по себе референт, существовал до того, как мы его помыслили и дали ему имя, существовал и будет существовать после того, как исчезнет последний носитель этого имени и связанного с ним понятия. «Разница между чтойностью вещей и фактом их существования не является концептуальной» [Брассье, 2017, с. 254]. Отсюда же следует ошибка Беркли по отождествлению физических объектов с переживаниями.
4. Разрешение проблемы объективного синтеза «по сути состоит в том, как разрешить вопрос об отношении между концептуальной мыслью и не-концептуальной реальностью. Но из того, что у нас есть понятие различия между Сатурном [понятием] и Сатурном [словом], не следует, что это различие является понятийным различием: понятие различия не равно понятийное различие» [Брассье, 2017, с. 256].

На наш взгляд, в дополнение к вышесказанному, следует рассмотреть взгляды Людвиг Витгенштейна и В.А. Смирнова. Речь идет о понимании факта и образа Витгенштейном, а также проблему «эмпирическое-теоретическое» в том виде, как она была поставлена Смирновым в его работе «Уровни знания и этапы процесса познания» [Смирнов, 1964].

В своем «Логико-философском трактате» Витгенштейн строит некоторую систему «мир – форма как субстанция мира - логическое пространство фактов – факт – образ – проекция - вещь». В его понимании существует некоторая взаимосвязанность этих понятий, которая образует совокупность фактов [Витгенштейн, 2011, 1.11], как отношений объектов [Витгенштейн, 2011, 2.01]. Философ пишет:

«Знать объект — это значит знать все возможности его вхождения в атомарные факты. Иначе мы не знаем объект (2.0123). Знание об объекте — это знание его внутренних качеств (2.01231)».

Отсюда, по нашему мнению, следует, что любое предложение есть прежде всего возможность как таковая. И возможность эта существует в рамках допустимой конфигурации мира, в которой все факты связаны друг с другом какими-то логическими отношениями. Таким образом, мир одновременно состоит из вещей, которые и определяют его фактичность, и в то же время в нашем сознании представляет собой конфигурацию различных форм так как он «изображается предложениями». Да, мы не можем говорить обо всем, мы не можем мыслить о чем-то, что мы не знаем. Рассел говорит в этой связи о «мистицизме» философа [Витгенштейн, 2011, с.31]. Но мистицизм этот относится к пониманию мира как целого: чтобы понять мир пришлось бы выйти за его пределы, но это невозможно. Таим образом, важным является то, что мир *может* существовать и за пределами наших мыслительных возможностей. Мы всего лишь не имеем возможности говорить об этом сколько-нибудь достоверно. Предложения, как *показывающие* смысл, способны посредством существующих выражений порождать новый смысл (4.03). Таким образом, производство новых смыслов в этом случае представляет собой расширение новых знаний о форме и качествах объектов, Источником является мир, который предоставляет поводы для новых логических конфигураций. Обнаруживая в мире, через чувственное восприятие то новое, что вписывается в существующее логическое пространство фактов, мы изменяем существующую конфигурацию мира.

В дополнение ко всему, на наш взгляд, существует некоторая двойственность в тезисах философа, которая может быть очень важна при решении задач общего ИИ. С одной стороны, мы видим явное преобладание модальности «возможности»: мы часто можем встретить установку «нечто может» с одновременной перформативной установкой «выражает», «показывает». Это указывает на принципиальную возможность существования новых конфигураций мира, тем самым, с одной стороны, обеспечивая доступ к ним. А с другой - стремление к логицизму на уровне форм, как способу конструирования мира в нашем сознании - дает основания

для уверенности в возможности воспроизведения со стороны машинного интеллекта части творческих способностей человека. Конечно же, при этом необходимо принять в расчет тезис философа о том, что познание это есть «философия психологии» (4.1121).

Таким образом:

1. Сама по себе констатация того, что мы имеем дело с проекциями, истинность или ложность которых только возможна уже является основанием для возникновения нового знания. При этом бытие вещей не ставится под сомнение. Сам мир предоставляет нам новые факты.
2. Познавательная ограниченность внутреннего мира того или иного субъекта (а она задается тем положением, что система не может познать саму себя) может быть расширена другим субъектом, наблюдающим первый через внешние качества и проектируя их на свое логическое пространство, отличное от первого.
3. Субъект может расширять собственный мир за счет постоянного процесса обнаружения себя, новых объектов и новых качеств объектов через чувственно воспринимаемые знаки.

Переходя к идеям В.А. Смирнова, важно отметить, что они являются во много развитием мыслей Витгенштейна. Он соглашается со многими идеями Витгенштейна. Прежде всего, он выделяет два плана-проблемы: «чувство-мысль» и «эмпирическое-теоретическое». Начиная свой анализ, он указывает на множественность способов употребления терминов «эмпирическое» и «теоретическое», прежде всего как противопоставление опытного и аналитического освоению действительности.

Анализируя уровни знания и этапы познания, философ начинает с того, что он подразделяет предикаты и индексные константы на термины, полученные из непосредственного наблюдения и термины, таковыми не являющиеся.

Первые - суть понятия, релятивизированные относительно субъекта, с некоторым чувственным коррелятом. Важно отметить, что они являются продуктом чистого опыта. Причина этого в том, что они получают свое выражение в виде предложений некоторого базисного языка, что приносит дополнительный элемент в непосредственное наблюдение. Сам язык, по его мнению, не является чисто конвенциональным понятием, так как языковая система связана с определёнными способами схематизации действительности, где именно язык лежит в основе акта наблюдения. Понятие наблюдаемого определяется через термины наблюдения. Формулирование результатов непосредственного наблюдения в терминах процесса наблюдения приводит к понятию диспозиционных предикатов, а в более широком смысле – к понятиям редукционных предложений.

Вторым типом подразделения терминов на «эмпирическое» и «теоретическое» является деление предложений на 1) фиксирующие факт и 2) выражающие закон. Эти подразделения не совпадают друг с другом: не всякое наблюдение выражает факт, некоторые будут выражать законы. В свою очередь, констатация факта может быть получена и не из непосредственного наблюдения, а опосредованно.

Проблема на этом не исчерпывается, так как сам по себе «объект» может представлять собой как «внутритеоретический объект» - объект науки, так и вещь или процесс реальности [Смирнов, 1964, §2]:

«С точки зрения материализма наука изучает в конечном счете объективную реальность, существующую вне и независимо от познания. Но из этого отнюдь не следует, что те объекты, относительно которых делаются утверждения, тождественны объективной реальности».

Все это - процессы идеализации и схематизации. Предвосхищая возможную критику со стороны прежде всего субъективных идеалистов, философ указывает, что имманентность объектов научной теории не подразумевает

того, что они являются фикциями. Да, они являются приблизительными моделями реальности, но сама реальность никуда не исчезает. Говоря о различиях между идеальными и эмпирическими объектами, мы должны действительно говорить об отличиях. Эмпирические объекты получаются путем огрубления и упрощения, мыслятся как себе тождественные, индивидуализированные. Но они все равно являются результатами схематизации и идеализации. Отличие теоретического и эмпирического объекта заключается только в степенях этой схематизации и идеализации. Отвергая наивный реализм, так же как это позднее сделают представители спекулятивного реализма, философ говорит о том, что объекты науки необходимо рассматривать как «приблизительные модели, прообразы действительности» [Смирнов, 1964, §2].

Из всего он делает следующий вывод [там же]:

«Подразделение объектов науки на эмпирические и теоретические в определённой степени соответствует проблеме «явление-сущность». Под эмпирическими объектами мы будем понимать не просто наблюдаемые объекты (это частный случай эмпирических объектов), а объекты, полученные в результате определенного способа подхода к действительности».

Принимая тезисы о 1) наличии мира вне-нас, 2) первичности языковой системы, обусловленной ей схематичностью и идеализацией, следует, по необходимости, теоретическое допущение о существовании возможностей и некоторых законов, которые могли бы обеспечить генезис символов со стороны машинного интеллекта. Безусловно, процесс этот творческий. Отсюда возникают вопросы 1) о принципиальной возможности генезиса новых символов и 2) об основаниях для такого рода процессов и путях их дальнейшей легитимизации в рамках тех или иных знаковых систем.

Фердинанд де Соссюр прямо пишет о «случайности языка» [Соссюр 1977, с. 17]:

«...язык — условность, а какова природа условно избранного знака, совершенно безразлично. Следовательно, вопрос об органах речи — вопрос второстепенный в проблеме речевой деятельности».

Язык, понимаемый таким образом, дает все основания для понимания природы первичного генезиса слов-символов языка как относительно свободного процесса, ограниченный скорее нашими квалиа, особенностями строения нашего речевого аппарата. На наш взгляд, можно соглашаться с теми или иными теориями, вставить на точку зрения Кратила и пришедшего ему на помощь Сократа, или «конвенционалиста» Гермогена из Платоновского диалога «Кратил». В любом случае, акт генезиса новых символов *остаётся в значительной степени контингентным событием.*

Любая языковая система может быть рассмотрена с различных сторон. Так, В.А. Смирнов выделяет три таких стороны [Смирнов, 1964, §3]:

1. синтаксическую - «как множество некоторого рода допустимых последовательностей символов»;
2. абстрактно-семантическую - «как множество языковых выражений, которым приписаны объекты из некоторой системы объектов» объекты из некоторой системы объектов»;
3. методологическую - «как множество выражений, относящихся к системам объектов определенной природы».

Генезис новых символов должен удовлетворять требованиям всех трех сторон.

На наш взгляд, процесс именованья, как процесс генезиса новых знаков (с последующей десигнацией) является более или менее свободным

процессом, главными ограничивающим факторами в котором являются скорее морфология и тезаурус того или иного языка. При этом:

1. наличие ошибок в процессе словообразования не может привести на этапе первичного именованя к значительным неудобствам с технической (рациональной) точки зрения и может быть скорректировано;
2. процесс первичного именованя носит конвенциональный характер, с учетом неполноты конвенциональности языка, о чем упоминалось ранее.

В этом отношении представляют интерес работы Г.С. Щура, в частности речь идет об идее синтаксических полей [Щур, 1991].

Обсуждение эстетических и ряда схожих проблем в рамках настоящего исследования является не столь принципиальным и выходит за рамки настоящего исследования. Отметим лишь, что в отдельных случаях эта проблематика может быть принципиальной.

Гораздо большую проблему представляет собой процесс генезиса новых концептуальных понятий. Следуя логике В.А. Смирнова именно семантические правила определяют допустимость тех или иных понятий [там же]:

«Понятие системы семантических правил означает в то же время принятие определённых систем допустимых объектов. Принятие способа речи коррелятивно принятию допустимых в качестве мыслимых объектов (систем объектов)».

Таким образом, сама форма тех или иных концептуальных понятий во многом должна определяться непосредственно системой языка.

На основании всего вышеизложенного, позволим себе сделать следующие выводы о том, как, каким образом, на наш взгляд может быть осуществлена принципиальная возможность генезиса новых понятий. Основанием для этого может служить два направления поиска:

1. обнаружение новых акциденций относительно существующих в онтологии экземпляров объектов;
2. лингвистически формализованный поиск разрыва между эмпирическими (в понимании В.А. Смирнова) и теоретическими объектами, чтойности и реального объекта.

Опираясь на вышеизложенное, позволим себе сделать следующие выводы относительно Вопроса 2 о возможности генезиса символов со стороны машинного интеллекта:

Во-первых, процесс именованя является в значительной степени свободным, контингентным актом. Само по себе присваивание имен со стороны, например, машинного интеллекта не несет никакой опасности. Безусловно, здесь необходимо учитывать правила словообразований и словоупотребления в тех или иных языках. Отсюда следует:

1. *ИИ-система имеет право давать имена новым понятиям, в соответствии с правилами установленными тем языком, на котором происходит процесс именованя.*
2. *Акт генезиса новых имен в значительной является свободным контингентным процессом, в рамках ограниченных тем языком, на котором происходит процесс именованя.*

Во-вторых, как уже упоминалось выше, легитимизация разрыва реального и чтойности объекта в научной репрезентации является

допустимой (по крайней мере в техническом аспекте). Признание этого с необходимостью приводит к реализуемости, по крайней мере теоретической, задачи генезиса новых понятий, как фиксации факта такого разрыва.

В-третьих, о необходимости возврата к *антропоцентрическим принципам*, как ограничивающим условиям генезиса со стороны машинного интеллекта с аксиологической, целевой и т.д. точек зрения. В терминах законов робототехники это могло бы звучать как Седьмой закон:

ИИ-система должна получить одобрение со стороны человека относительно корректности имен; при наличии разногласий робот обязан принять точку зрения человека.

Признав права машинного интеллекта на генезис имен, следующим шагом по необходимости становится проблема того, каким образом может происходить конструирование самих объектов. И здесь символьные методы, основанные на таких математических абстракциях как структуры и классы дают широкое поле возможностей по конструированию объектов в ходе процессов объективации, которые можно вслед за Дьёрдем Лукачем представить как фиксацию структур.

Проблематика генезиса новых правил

Исходя из определения того, что представляют собой символьные методы, а также того, как они реализуются сегодня в практическом плане (наборы правил и деревья решений) с неизбежностью встает вопрос об их природе. На наш взгляд, поиск ответов на этот вопрос необходимо искать прежде всего в двух аспектах:

Во-первых, необходимо провести логико-гносеологический анализ возможных механизмов генезиса новых продукционных правил.

Во-вторых, понять возможные пути обретения ими смыслов в приложении их к задачам общего ИИ.

Отдельно следует, хотя бы в первом приближении, рассмотреть вопрос о достаточности тех или иных логик. На наш взгляд, в той или иной мере ответы можно найти уже в работах Алана Тьюринга. Отдавая дань уважения ученому, необходимо отметить, что сам Тьюринг не считал свой алгоритм способным полностью имитировать человека по ряду причин:

Во-первых, данный алгоритм на практике должен был реализовывать только методы дедуктивного вывода. Изначально предполагалась его использование главным образом для замены человека-вычислителя при выполнении рутинных задач, поиска доказательств математических теорем и ряда других поддающихся алгоритмизации задач. Сам исследователь, уже позднее в своей работе «Вычислительная техника и интеллект»¹⁰, опубликованной в октябре 1950 года в философском журнале «Mind» *поднимал проблемы интеллекта и интуиции как немеханические*. Впрочем, ряд исследователей, такие как Ходжес [Parsing the Turing Test, 2008, p.11-15] считали, что Тьюринг допускал в будущем преодоление этих проблем с помощью машины Тьюринга. На наш взгляд, такое утверждение является достаточно сильным преувеличением. Впрочем, сам же Ходжес ссылаясь на позднейшую статью Тьюринга уже 1954 года, пишет:

«Его последняя опубликованная статья (Тьюринг, 1954) снова стала довольно популярной, появившись в «Penguin Science News». Здесь не упоминалась квантовая механика, но она вернулась к чистой математике вычислимости, которая недавно обрела новую жизнь с достижениями в алгебре, и дала ответ относительно теоремы Геделя. Его заключение было удивительно непохоже на то, что было высказано

¹⁰ <https://web.archive.org/web/20080702224846/http://loebner.net/Prizef/TuringArticle.html>

в 1950 году. Он сказал, что теорема Геделя показывает, что «здравый смысл» необходим для интерпретации аксиоматических систем, и на этот раз интуитивный «здравый смысл» не был заявлен как нечто такое, что машина может продемонстрировать так же хорошо, как и человек.»

Во-вторых, как позднее писал Сёрль, данная концепция реализует бихевиористскую модель, ограниченность которой Сёрль блестяще показал в своей «Китайской комнате» [Searle, 1980]. Схожих взглядов придерживаются, например, Дрейфус [Дрейфус, 2010], Роджер Пенроуз [Пенроуз, 2020] и ряд других исследователей.

В-третьих, сам Тьюринг в ходе интервью радио Би-би-си (1956), объясняя принцип, согласно которому любой механический процесс может быть смоделирован программой, запущенной на одной машине-компьютере, сделал важное предостережение, которое не было сделано в 1950 году. Машина, которую нужно смоделировать:

«...должна быть такой, чтобы ее поведение было в принципе предсказуемо с помощью расчетов. Мы, конечно, не знаем, как следует проводить подобные вычисления, и Сэр Артур Эддингтон даже утверждал, что из-за принципа неопределенности в квантовой механике такое предсказание даже теоретически невозможно».

В-четвертых, Тьюринг говорил о «непрерывности нервной системы». Он утверждал, что основа работы мозга в непрерывности его сущности, в противоположность дискретности в случае машины. При этом дискретная система может приближаться к непрерывной настолько близко, насколько это необходимо. На наш взгляд, этот его тезис требует отдельного философского осмысления с целью понимания влияния дискретности на мыслительную деятельность.

В-пятых, это понятие случайности, которое Тьюринг ввел без какого-либо серьезного определения. Его иллюстрация использовала "цифру Пи" как случайную последовательность, и это *par excellence* вычисляемый вход.

Данный список можно продолжить, например, недавней работой Рогнара Фейланда [Fjelland, 2020], а также исследованиями ряда других авторов. Человеческий интеллект не ограничивается теми или иными логиками, следовательно, без понимания природы той «темной материи», что лежит за пределами логического мышления, мы можем говорить либо о достижении пределов возможностей, либо об имитации со всеми вытекающими отсюда последствиями.

Понимание того, что представляют собой правила, являются ключом к пониманию того, как может происходить их генезис. Прежде чем приступить непосредственно к исследованию данного вопроса необходимо определиться с тем, что в рамках настоящего исследования нет необходимости придерживаться какой-либо определенной теории истины. Необходимость постулирования данного тезиса определяется тем, что в ином случае, мы будем вынуждены ограничиться только частью существующих сегодня логик, понимая логику как «науку о правилах рассудка вообще» [Кант, 1994, с. 90].

Еще в середине XIX века Джон Стюарт Милль, анализируя работы В. Гамильтона, говорит о существовании идеи воли как об ожидании, которое мы чувствуем таким образом «что когда существует причина, мы должны воспринять действие» [Милль, 2019, с. 290]. При этом он критикует Юма, так как, по его мнению, следует говорить не о привычке, а об убеждении в реальности существования связи «двух вещей в нашей голове», которую опыт, не опровергая без необходимости в будущем, сумеет точно воспроизвести в качестве порядка природы. В основе логики Милля лежат вещи, выраженные понятиями и их отношения. Одной из главных нерешенных проблем Милля, как уже говорилось выше, является психологизм, оставляющий логику в статусе искусства.

Витгенштейн, относит теорию познания к «философии психологии» [2011, 4.1121]. Так же как и Милль, философ говорит о высказывании как о связи как минимум двух вещей. Вещи, как уже отмечалось ранее, по мнению Витгенштейна не могут существовать вне связи с другими вещами. Но последний на наш взгляд, идет гораздо дальше Милля в целом ряде вопросов, в значительной степени преодолевая проблему психологизма, которая является одним из главных препятствий к пониманию того каким образом можно реализовать концепцию общего ИИ в рамках машинного интеллекта.

Говоря «aRb», мы говорим «то, что «a» стоит в определенном отношении к «b»» [Витгенштейн, 2011, 3.1432]. Факт, как соединение объектов, независим от других фактов. И уже здесь лежит принципиальная возможность для открытия новых фактов, их можно увидеть в противоречиях существующей конфигурации. Предложения могут изображать действительность, но не могут изобразить того, что они имеют общего с действительностью. Поэтому, они не могут изображать логическую форму – она сама в них отражается: «предложение показывает логическую форму действительности» [Витгенштейн, 2011, 4.121]. Этим философ закрывает возможность искать истинность и саму логику чисто аналитически, без обращения к действительности. Логика сама по себе лежит за пределами тех или иных логических систем. *Они являются лишь изображением действительности.* Отсюда, на наш взгляд возникает понимание того, что выражения, как составная часть более сложных предложений имеют свою первичную основу в мире, а не в языке. Следовательно, решая проблему генезиса новых правил, мы по необходимости должны обращаться к миру или другому субъекту, который действует уже в рамках своей конфигурации, когда он уже каким-то образом получил эти знания. Исключением из этого не может являться ни естественный ни машинный интеллект. Но это не исключает возможность совместного «аллопоэтического» проектирования новых конфигураций мира. И в этом же лежит причина, по которой мультиагентные системы должны рассматриваться как более эффективные. Более детальный анализ вопроса о

сравнительном анализе эффективности мульти- и моноагентных систем лежит за рамками настоящего исследования.

Витгенштейн пишет [Витгенштейн, 2011, 4.2]:

«Смысл предложения есть его согласование или несогласование с возможностями существования атомарных и неатомарных фактов».

Отсюда возникает несколько выводов, которые могут быть полезны для понимания возможных процессов генерации новых правил.

Во-первых, первоначальное конструирование (если мы говорим в терминах машинного интеллекта) высказывания не обязательно должно осуществляться в жестких границах истинности, оно может иметь более нечеткие границы. Например, более общего класса. Первоначальный процесс, так же, как и в случае с символами, может носить случайный характер или даже намеренно содержать в себе нонсенс. И только последующая верификация возможности его существования в рамках *потенциально модифицированной т.е. новой конфигурации мира*, является причиной принятия или отвержения нового высказывания. К сожалению, вводя таблицы истинности, Витгенштейн не дает ответа о природе их возникновения. Рассуждая в логике философа, мы можем лишь апеллировать к тому, что говорилось относительно вещей и атомарных актов. Принципиально описать их истинность мы не можем, так как для этого должны выйти за пределы языка, содержащего их. Истинностные возможности лежат уже в самом положении дел. «Логических объектов не существует» [Витгенштейн, 2011, 4.441]. Мы можем лишь говорить о том, что «считаем» те или иные предложения истинными, выражая тем самым свое отношение к действительности.

В случае несоответствия между собранной из атомарных фактов существующей логической картиной и действительностью мировой ситуации,

понимание причин возникновения новых знаний является достаточно прозрачным:

местом возникновения новых знаний является выявленное несоответствие существующей логической конфигурации мира и обнаруживаемой действительности.

Все это осуществляется в рамках высказывания философа о возможности конструирования мира с помощью логических строительных лесов [Витгенштейн, 2011, 4.023], в связке с тем, что «предложение должно в старых выражениях сообщить нам новый смысл» [Витгенштейн, 2011, 4.027].

Поясним это положение на базе работы Уилларда Куайна «Две догмы эмпиризма» [Quine, 1963]:

«Для Аристотеля вещи имеют сущность, но только лингвистические формы имеют значение. Значение — это то, чем становится сущность, когда ее разводят с предметом референции и отдают замуж за слово».

Предметом теории значения в этом случае становится исследование синонимии лингвистических форм аналитических высказываний как имеющих силу значения (в понимании Куайна) вне зависимости от фактов. Но такое положение дел является спорным, по мысли философа ввиду того, что либо они представляют собой тавтологии, либо являются «отчетом лексиколога о наблюдаемом явлении».

Тем не менее, существует особый тип деятельности, названный еще Карнапом «экспликацией», целью которой является не просто перефразировка, но улучшение определяемого за счет его упрощения или дополнения. Куайн выделяет три типа такого рода экспликации [там же]:

1. «определяющее выражение может представлять собой точную перефразировку определяемого выражения в *более короткий способ записи*, сохраняющий непосредственную синонимию предшествующего употребления»;
2. «определяющее выражение может, в духе экспликации, *улучшить предшествующее употребление* определяемого выражения»;
3. «определяемое выражение может представлять собой *впервые созданный способ записи*, впервые наделенный значением здесь и теперь».

Таким образом, генезис новых определений как процедура именования и конструирования новых пропозиционных высказываний может осуществляться в рамках уже существующих логических строительных лесов путем их «перестройки».

Идею конструирования мира по средствам языка развивает и Ойген Финк в работе 1957 года «Оперативные понятия феноменологии Гуссерля». Темой исследования главы «VI Медитация» Гуссерля, как пишет Финк, является конструирование мира. Если в обыденном употреблении язык достаточен, так как все уже обговорено, то относительно новой сферы «трансцендентной жизни» [Шестова, 2015, с. 71] дела обстоят не так хорошо. Как пишет Е.А. Шестова [там же]:

«Именно вопрос о том, как создаются и функционируют понятия, впервые, с точки зрения Финка, и обращает нас по-настоящему к вопросу о языке».

По мнению Финка, феноменологическая деятельность, это не просто еще один уровень философской рефлексии. Наблюдатель, погружаясь в феноменологию, расщепляет свое трансцендентальное Я на два: «конструирующее Я» и «Я, осуществляющее свою феноменологическую деятельность» или

«теоретизирующее Я» [Шестова, 2015, с. 73]. Так возникает новый модус сознания. Сфера феноменологии существует отдельно и не соотнесена с миром. По его мнению, «философия не участвует в конструировании мира» [Шестова, 2015, с. 74]. Именно в этом философ видит коренное отличие феноменологии от психологии. При этом отношения их носят «меонтический» характер: где греческое «μῆ» подчеркивает факт не полного отрицания, а «бытие не в том смысле». Феноменология не просто обнаруживает, но и конструирует «трансцендентальное бытие», в котором основные понятия создаются в результате феноменологической редукции, тем самым расширяя его относительно первичной психологической редукции. Психология, напротив, *уже* содержит в себе основные понятия, такие как «душа», «сознание» и т.д. так как опирается в полной мере на язык присущим «конституирующему Я» способом [Шестова, 2015, с. 75]:

«...высказывание как бы придает своему предмету статус так или иначе существующего».

Феноменологическая деятельность, как отторгнутая от мира, не имея той же языковой оформленности, требует собственной редукции. Более того, *она и не должна конструироваться так же, как сущее.*

При этом, хотя Финк и ставит проблему трансцендентального языка, но говорит о бессмысленности такого языка в отрыве от языка естественного. Мы можем говорить о том, что слава естественного языка могут отсылать к трансцендентальному смыслу, постоянно находясь в некотором динамическом равновесии.

Отсюда, на наш взгляд можно сделать первый вывод, который необходим для понимания теоретических пределов возможностей машинного интеллекта в вопросе генезиса новых правил и понятий:

1. Генезис новых понятий и правил возможен как перефразировка существующих выражений, представляющих собой комбинации уже известных атомарных и неатомарных фактов.
2. Верификация новых правил возможна только в рамках той или иной логической конфигурации мира (иными словами - ограниченно), путем как проверки взаимосогласованности новой структуры относительно существующей конфигурации, так и, одновременно с этим, на соответствие действительному положению дел.
3. Поиск соответствия действительному положению дел ограничивается конфигурацией той сенсорной системы, которой обладает система как источником чувственного восприятия.
4. Обязательным этапом верификации является поиск подтверждения истинности новых фактов в рамках коммуникации с другими системами, внешними по отношению к текущей (человек, иной экземпляр машинного интеллекта).

Настоящее исследование было бы неполным без обсуждения еще одной стороны проблемы, которую С.И. Шапиро определил как «контроверзу доказательства и открытия истины» [Шапиро, 1973, с. 8]. Исследователь приводит слова Д. Пойя [там же]:

«...будем учиться доказывать, но будем также учиться догадываться».

Одну из главных проблем он видит в отсутствии единства даже на уровне самого понимания вопроса. Так, по его мнению, противники алгоритмизации указывают на несовпадение критериев логической строгости и психологической целесообразности. Сторонники, в свою очередь, указывают на то, что несовпадение еще не означает несовместимость. Сам же исследователь считает, что «алгоритмичность вовсе не исключает эвристику» [Шапиро, 1973, с. 8]. И первым шагом на пути к такому пониманию является

расширение понятия алгоритма как строго формального процесса. Необходимость уйти от понимания алгоритма исключительно как процесса приведения «регулярных процессов решения задач к простым, максимально элементарным операциям» [Шапиро, 1973, с. 10]. Такая трактовка, по его мнению, не является «целесообразной». Человек мыслит «связанно», «умственными блоками», и излишнее дробление для него противоестественно. В результате чего, вслед за Л.Н. Ланда и Лотфи Заде он приходит к необходимости «расплывчатых» алгоритмов. В которых алгоритм понимается скорее как «предписание алгоритмического типа» [Шапиро, 1973, с. 11]. В частности, он пишет [Шапиро, 1973, с. 11]:

«Рассмотрение различных форм поведения и мышления приводит к заключению, что имеется некоторая достаточно «непрерывная» их градация, определяемая различной степени «жесткости» поведенческих и интеллектуальных процедур».

Именно эта концепция является для него надеждой на решение проблемы связи между «формально-детерминистской» и «неформально-эвристической» сторонами интеллектуальной деятельности. Таким образом, следуя Шапиро, нечеткие алгоритмы можно считать одним из необходимых, но не достаточных инструментов для разработки процедур автоматической генерации новых понятий и правил. Недостаточность такого рода алгоритмов обусловлена прежде всего тем, что они оставляют открытыми проблемы индуктивного вывода и связанных с ними вопросов интуиции.

Согласно Анри Бергсону, интуиция является одним из наиболее важных условий творческого процесса. Прежде всего это касается вопроса выбора, как «творчества возможного» [Бергсон, 2019, с. 189]. Выбор, если он не является автоматическим процессом, свойственным низшим организмам, является «потребностью» сознания. Интуиция и интеллект противоположны друг другу. Интуиция «движется по ходу самой жизни, интеллект идет в обратном

направлении» [Бергсон, 2019, с. 193]. Развитие интеллекта человека привело к ослаблению интуиции, хотя она и продолжает существовать «в неотчетливой и прерывистой форме» [там же]; погруженный в материю интеллект представляет свободу в форме необходимости. Тем не менее, именно интуиция, пусть и в остаточной или ослабленной форме, позволяет нам творчески уйти от ограниченности выбора. Сознание, не опираясь только на материю, на необходимость, способно создавать новое. В противном случае сознание, лишённое творческого выбора, «засыпает», оставляя один лишь автоматизм.

Отсюда можно сделать два вывода, которые будут необходимы для дальнейшего решения поставленной нами проблемы:

Во-первых, наличие теоретической возможности разделения 1) интуитивной (подсознательной, выходящей за рамки механической целесообразности) умственной деятельности и 2) рациональной, казуально обусловленной материей и ее формами деятельностью.

Во-вторых, такое обособление умственной деятельности задает возможность наличия творческой составляющей только у сознательной деятельности, тесно связанной с интуицией. В противоположность ей, чисто рациональная логическая деятельность не обладает таким свойством «ибо природа чисто логического существования такова, что оно по видимости самодостаточно и полагается силе имманентной истины» [Бергсон, 2019, с. 193].

Имея возможность «творческого выбора», человек, в отличие от животного, не ограничен фактически существующими вариантами. Следуя пониманию сознания Бергсоном, мы можем сказать о существовании потенциальных выборов за пределами фактического положения дел, т.е. о процессе не привязанном строго к наличной фактичности.

Несколько в ином ключе, но дополняющем понимание того, где может быть обнаружен механизм генезиса нового, рассуждает Карл-Отто Апель. Так в своей работе «Две фазы феноменологии ...» философ рассуждает о мире, порождаемом поэзией и языком. Поэтический мир, реализованный средствами языка, мыслится им в духе, положенном такими философами как Вильгельм фон Гумбольдт, Роман Ингарден и, главным образом, Мартин Хайдеггер.

Соглашаясь с Ингарденом, он пишет [Апель, 2001, с. 18]:

«...поэзия не является «бытийно-автономной» (как, например, природный организм); она является заключенным в языке интенциональным образованием, которое приобретает свою интенциональность от действительных суждений, которые относятся к реальности».

К сожалению, ни Ингарден, ни позднее Гуссерль не сумели, по его мнению, преодолеть платонизм, который оставляет «фундирование языковых структур значения в идеально-сущем» [там же].

Определяя «вторую фазу» развития феноменологии как синтез, философ предлагает обратиться к идеям Хайдеггера. Он прежде всего обращает внимание на то, что любое встречаемое нами сущее дает нам перспективу мира, каждый раз встречая его, мы вновь и вновь осмысляем весь мир во всей его полноте. И разорвать этот герменевтический круг нельзя так он является для нас «фундаментальным событием».

Для Хайдеггера, пишет он далее, «язык представляет собой структуру артикуляции бытия» [Апель, 2001, с. 25]. Если в повседневном употреблении трудно заметить просвет бытия, то в определенных обстоятельствах это становится заметно, например, когда «случай» подводится под «понятие класса». Как только мы обнаруживаем феномен, то мы тотчас же переходим к поэзии. Хайдеггер называет поэзию «совместное учреждение бытия». Приведем слова философа [Апель, 2001, с. 27]:

«В просвете бытия к миру, которое свершается через само творение, Хайдеггер нашел то всеохватывающее целое, ту область, к которой принадлежит творение, и которая не была бы ни чуждой ему, ни чем-то иным, к чему оно могло быть редуцировано. Само оно - творение - впервые открывает эту область».

Таким образом, мы можем говорить о том, что творение в своей целостности, как интенционально наполненное, раскрывает свой мир «и вступает в спор с миром, до той поры знакомым и признанным» [там же]. Творение искусства, являясь конечно же фикцией с точки зрения известного положения дел, можно рассматривать как вспышку смысла. Но «фиктивность» является таковой только с точки зрения историчности бытия-в-мире, т.е. конкретной истории [Апель, 2001, с. 27]:

«Поэзия теперь «исторична» [«geschichtlich»], но не в смысле той историчности [«historisch»], как она рассматривалась в 19-м столетии. Ее историчность состоит не в том, что она как обусловленное может быть редуцирована к тому, что уже осталось в прошлом во времени, а в том, что в ней и через нее абсолютное бытие «временится» из своей будущности и тем самым основывает историю, заново раскрывая человечности ее мир».

Если онтически, как фактически-случайное, творение остается фикцией, то в отношении бытия сущего оно становится реальным.

Таким образом, мы с полным основанием можем принять изначальную реальность результата поэтического акта как существующего, пусть и в рамках иной исторической экзистенциальности. И самое главное - как допустимое расширение поля пространства возможных выборов. Безусловно, остается открытым вопрос о возможном механизме первичного акта генезиса. Оставаясь только в рамках гипотезы о случайной генерации новых сущностей,

невозможно в практическом плане надежно реализовать такую процедуру. Необходимы дополнительные ограничивающие условия.

Для этого следует вернуться в практическую плоскость решения проблемы. В этом отношении прежде всего следует обратиться к работам В.К. Финна. На наш взгляд, предложенная исследователем концепция ДСМ-метода, представляет собой одно из наиболее приближенных к преодолению проблемы генезиса новых правил решение.

В основе процесса познания В.К. Финн ставит теоретические конструкции. Под ними исследователь понимает «теорию, фрагменты теорий и совокупности принципов и правил» [Финн, 2011, с.16]. Верификация таких конструкций должна происходить прежде всего экспериментально, в результате чего мы либо должны подтвердить теорию, либо зафиксировать «парадокс исследования». Именно парадоксы и служат основанием для корректировки существующих теоретических конструкций. Важным замечанием, на наш взгляд, является утверждение о неформализуемости вопроса о том, что является первичным – экспериментальные данные или теоретические конструкции «в силу синтетичности творчества и наличия инсайтов».

Ум определяется им как «способность к ориентировке в окружающем мире посредством продуктивного мышления» [Финн, 2011, с.19]. Отличительной чертой такого рода мышления является способность к индуктивному извлечению опыта, путем сравнения фактов и «извлечения из них сходных подструктур «носителей» причин и контрпричин событий». Основной цепочкой такого процесса является [там же]:

«Анализ опыта – гипотеза-сравнение ее с целью – верификация или фальсификация – коррекция гипотезы и пополнение данных и знаний – снова анализ опыта».

В свою очередь, «имитация в компьютерной системе» научного познания видится им как система, состоящая из 1) теоретических конструкций (База знаний), 2) экспериментально полученных данных (Баз данных), а также 3) рассуждений, аппроксимирующих эвристику путем применения последовательности правил. Под «рассуждением» Финн понимает [Финн, 2011, с.16]:

«...построение последовательности аргументов, вынуждающих принятие некоторого утверждения, которое и является целью рассуждения».

В рамках имитационной компьютерной модели База данных представляет собой совокупность объектов с заданной структурой: множества, кортежи, слова, графы и т.д. [Финн, 2011, с.25]

База знаний [там же] состоит из аксиом (например, булевой алгебры, алгебры кортежей и т.д.); аксиом связи исходных предикатов (задающих классы объектов, классы задач); аксиом задающих процедуры обработки информации, аксиом описывающих пространственные объекты; аксиом описывающих эффекты.

Правила вывода подразделяются на «правила достоверного» и «правила правдоподобного вывода». Правдоподобные выводы реализуются в некотором «генераторе гипотез», а достоверные – в «доказателе теорем» [Финн, 2011, с.26]. Задача этих правил вывода заключается в связывании аргументов (аксиом, принципов и т.д.) базы знаний и базы фактов (эмпирически полученных фактов). Наибольший интерес в теории Финна с точки зрения нашего исследования представляют собой правила правдоподобного вывода (ППВ). Следуя Пойя, Финн относит к ним рассуждения, применяемые для «формализации эвристик решения задач» [Финн, 2011, с.39]. ППВ подразделяются автором на 2 рода: 1) порождающие положительные примеры и 2) умозаключения по аналогии. В основании

процедур верификации новых правил автором положены принципы фальсифицируемости Карла Поппера. Финн пишет [Финн, 2011, с.26]:

«ППВ извлекают сходства объектов (ср. мысль Бергсона о работе интеллекта), используя положительные и отрицательные примеры изучаемого явления. Обнаружение сходства объектов на подмножествах примеров влечет выдвижение гипотезы. <...> Столкновение (+)-причин и (-)-причин, порожденных ППВ, дает «фактическое противоречие»».

Прежде всего, речь идет об ампликативных выводах, как выводах, основанных на догадке. Автор вводит понятие «когнитивных правдоподобных рассуждений», подразделяя их на три типа [Финн, 2011, с. 40]:

1. вероятностные (например, на базе байесовских правил);
2. приближенные (например, на базе нечетких множеств);
3. правдоподобные «являющиеся организацией различных взаимодействующих познавательных процедур».

В своих «Принципах конструирования интеллектуальных» систем В.К. Финн выделяет три типа предметных областей [Финн, 2011, с. 41]:

1. предметные области, для которых существует случайность событий обнаружения фактов в процессе анализа;
2. предметные области, в которых факты причинно обусловлены;
3. предметные области, для которых факты могут быть как причинно обусловленными, так и случайными.

В свою очередь, когнитивный процесс порождения нового знания и выдвижения гипотез может быть охарактеризован посредством трех теорий истины [Финн, 2011, с. 48]:

1. теории соответствия (формирование базы фактов);
2. теории когерентности (оценка выдвигаемых гипотез);
3. прагматической теории (оценка результатов).

Основным механизмом, обеспечивающим генерацию новых правил, является, по его мнению, «эмпирическая индукция» как «порождение причин эффектов на базе обнаруженных *сходств* фактов» [Финн, 2011, с. 48]. Иными словами, автор узаконивает неполную индукцию, допуская возможность обобщения на основе неполных данных. К сожалению, ДСМ-метод не решает проблему полностью. В частности, не решенными остаются такие проблемы как:

1. проблема формализации знания (особенно в сфере гуманитарных наук),
2. реализация новых структур, не имеющих аналогов: выход за границы тех ограничений, которые предлагает принцип конструирования по аналогии, каким бы ни было большим изначальное поле вариантов;
3. не получен ответ на вопрос, каков способ генезиса новых классов сущностей;
4. каков механизм порождения новых знаков.

Таким образом, проблема *творчества* (понимаемая как проблема генезиса) остается нерешенной в рамках ДСМ-метода. Тем не менее, ДСМ-метод представляется полезным по целому ряду причин:

1. ДСМ-метод предлагает в значительной степени проработанный механизм генезиса новых правил по аналогии, на основе анализа сходств и различий;

2. ДСМ-метод позволяет более ясно увидеть возможные подходы для получения граничных условий, в которых могла бы существовать искомая эффективная функция первичной генерации новых правил;
3. ДСМ-метод уже сегодня демонстрирует возможную роль и способ существования такой функции, пусть ограниченное возможностями процедур, основанных на поисках сходства и различия;
4. в значительной степени подтверждает, в практическом плане, эффективность использования принципов фальсификации Поппера в задачах генезиса новых правил.
5. ввиду наличия подтвержденной эффективности ДСМ-метода является целый ряд успешных примеров, например, в области анализа медицинских данных [Финн, 2011, с. 89].

Существует, на наш взгляд еще одно обстоятельство, которое может прояснить различие и, как следствие, определить пути, которыми следует идти для преодоления ограничений ДСМ-Метода. Речь идет о понимании мышления Хайдеггером и Делезом. Так Делез различает два вида вещей: «не затрагивающие мышление» и «заставляющие мышление» [Делез, 1998, 174].

К первым относятся объекты узнавания. Конечно же мышление необходимо в этом случае, но работа мышления здесь ограничена тем, что уже присутствует в нем. В таких вещах нет места сомнению, но этот процесс не имеет «ничего общего с размышлением». К такого рода вещам, по нашему мнению, относятся те, на которые направлен ДСМ-метод. Метод поиска сходств и различий позволяют лучше «увидеть» известные вещи, и, в некотором смысле – сконструировать похожие [там же]:

«Все истины такого рода гипотетичны, поскольку они не способны вызывать акт размышления в мышлении, так как они уже предполагают все, о чем спрашивается».

Совсем иначе обстоят дела в том случае, если нет узнавания или мы имеем дело с «сомнительными» вещами. Как пишет философ, здесь необходимы [Делез, 1998, 174-175]:

«...добрая воля мыслящего и добрую природу мышления. <...> В мысли первичны взлом, насилие, враг».

В этом случае мы имеем дело с творчеством, как следованию правилу Витгенштейна о котором речь пойдет ниже.

Подводя итоги, мы с необходимостью приходим к тому, что одной из оставшихся нерассмотренной ключевых проблем остается проблема творчества, как выхода за рамки поиска сходств и различий: *создание процедуры генезиса новых сущностей за пределами возможностей по комбинаторному соединению ранее известных фактов.*

Иными словами, речь идет о творческом (поэтическом) акте генезиса, за пределами узнавания и связанным с ними поиском сходств и различий.

Проблема творчества. Суть проблемы творчества, применительно к теме нашего исследования, заключается в поиске эффективных процедур, обеспечивающих процесс генезиса не только непосредственно новых правил как соответствий фактов, но синтеза новых концептуальных структур. На наш взгляд основными условиями возможности такого синтеза является «Великая проблема философии», сформулированная В.А. Смирновым как:

возможность свободного выхода за рамки существующих теоретических конструкций.

Соглашаясь с тезисом Дрейфуса, о «хайдеггерианском ИИ» как о «воплощенном ИИ», позволим себе в практическом плане сузить понимание проблемы до необходимости реализовать, хотя бы частично, возможности

генезиса новых сущностей в качестве свободного творческого акта. При этом оставляем как главное ограничение тот факт, что (несмотря на наличие разного рода сенсоров) машинный интеллект остается патнэмовскими «мозгами в бочке». Тем не менее, даже ограниченное решение, обеспечивающее имитацию творчества средствами символьных методов, в значительной степени будет влиять как на понимание роли этих методов, так и на круг задач, в которых системы машинного интеллекта могут действовать автономно.

В качестве практического механизма, позволяющего имитировать свободу, на наш взгляд, может выступить случайность или «квазислучайность». В некотором смысле подтверждение этому тезису мы находим уже в словах Апеля [Апель, 2001, с. 29]:

«...поэзия, именно благодаря фиктивной свободе (а не полной онтологической независимости!) по отношению к фактическому, возносит бытие к истине, это фактическое <...> принимается в расчет эмпирическими науками.»

Здесь философ говорит о возможности поэзии. Генезис нового, творческий процесс изобретения нового, безусловно относится к этой области как учреждению бытия сущего в своем собственном бытии. Важно отметить, что по мнению философа существование творческого акта исторично, но не в смысле обыденном, а исходя из того, что оно также «временится» [там же], являясь, что важно, редуцируемым к прошлому таким же образом.

В качестве итога всего вышесказанного позволим себе выдвинуть рабочую гипотезу, которая могла бы быть положена в основание процесса генезиса новых сущностей таких как правила, понятия и символы (знаки). Данная гипотеза включает в себя два следующих тезиса:

1. *Началом генезиса новых концептуальных структур, понятий и правил является (псевдо)случайный акт генезиса первоначального ансамбля правил на основе парадокса или нонсенса, в понимании их Жилем Делезом (о чем речь пойдет далее).*
2. *Если в вопросах, так или иначе связанных с проблемой генезиса новых символов, фактор случайности в значительной степени ограничен рамками феноменологии, эстетики, квалиа (физиологией, или конструкцией сенсоров в случае машинного интеллекта) и т.д., то в задачах генезиса новых правил такого рода ограничений или не существует, или они несущественны.*

Следуя терминологии Финна, обозначим процедуру генезиса новых понятий, правил и знаков как «квазитворческий генератор». Приставка «квази», необходима для определения не столько границ, сколько пределов свободы со стороны машинного интеллекта относительно процесса генезиса. Последующая верификация результатов безусловно необходима и может быть реализована аналогично тому, как это осуществляется, например, в рамках ДСМ-метода.

В основании такого генератора нам видится прежде всего возможность отхода от позитивистских методов и полноценного включения категории случайности.

Случайность - эвристика – индукция. Прежде всего, обратим внимание на фактор осознанности и неосознанности мыслительного процесса. Так С.И. Шапиро обращает внимание на феномен связи между осознанными и неосознанными процессами. Речь идет о снижении степени осознанности умственного процесса с ростом скорости течения этих процессов. Автор дает этому феномену имя «свернутости» [Шапиро, 1973, с. 229]. С точки зрения нашего исследования, такого рода «свернутость» или «развернутость» процесса не представляют интереса. Нас интересуют те интуитивные

умственные процессы, результатом которых является акт инсайта как казуально недетерминированного, случайного акта генезиса новых правил и концептуальных структур. Безусловно, речь идет о деятельности, имеющей внешние условия, такие как наличие целей, а также Баз знаний и Баз фактов в терминологии В.К. Финна.

Сама же по себе проблема эвристики на наш взгляд может быть реализована с помощью принципов, аналогичных тем, что заложены, например в широко известном генераторе псевдослучайных чисел, а также ряда дополнительных принципов, о которых речь пойдет ниже. Таким образом наши рассуждения приводят нас к идее об эвристике (и связанном с этим инсайте) как квазислучайном процессе.

В этом отношении представляет интерес следующая мысль В.Ф. Спиридонова [Спиридонов, 2014, с. 105]:

«Эвристики в отличие от других стратегий независимы от материала решаемой задачи <...> они лишены функции решения, то есть с помощью использования одних эвристических стратегий решить задачу не удастся. Их действие лежит в иной плоскости — переформулирования проблемной ситуации, выделения в ней значимых условий, фиксации сделанных ошибок, изменения креативного состояния решателя и т. д. Таким образом, функцией эвристики в ходе решения является помощь в анализе и понимании задачи, облегчение процесса порождения новых идей, то есть наступления инсайта.»

Важно подчеркнуть, что эвристические методы не гарантируют получение правильного результата. Иными словами, процесс получения эвристического решения, по крайней мере человеком, в определенной степени является случайным процессом [Спиридонов, 2014, с. 107]:

«Процесс решения задачи с помощью эвристических стратегий напоминает не поступательное движение к цели, но, скорее,

использование ловушек и силков в охоте на крупного и сильного хищника».

По нашему мнению, проблема здесь та же, что и в случае с полной индукцией – невозможность полного перебора вариантов. Тем не менее, даже такая «негарантированность» полученных результатов с эволюционной точки зрения оказалась эффективной стратегией. Перенос такого инструмента в логику машинного интеллекта, на наш взгляд является важным моментом в решении задачи создания общего ИИ.

Одной из проблем, которые обнаруживаются при такой постановке задачи, является необходимость максимально сузить поле возможных вариантов, тем самым повышая вероятность обнаружения необходимого решения. Попробуем сформулируем первые четыре принципа которые должны быть положены в основу квазитворческому генератору:

1. отсутствие гарантированного результата – норма;
2. случайность;
3. преднамеренное нарушение правил.
4. Нонсенс и «пустое место»

Иными словами:

Первичный акт генезиса должен происходить путем преднамеренного нарушения существующей структуры правил Базы знаний или (логических строительных лесов) за счет реализации принципа случайности при отсутствии требования обязательности получения результата. В качестве триггера, вызывающего начало акта генезиса и, одновременно с этим, верифицирующего факт того, что акт свершился выступает нонсенс или обнаружение «пустого места» в понимании Делеза.

Отсутствие гарантированного результата – норма. Такой тезис обеспечивает допущение принципиальной возможности самой попытки в условиях недостатка и, в некоторых случаях – значительного, отсутствия информации и каких-либо объективных условий для принятия полностью объективного и казуально обусловленного решения. В качестве вытекающей отсюда проблемы укажем необходимость поиска стратегии максимального уменьшения количества потенциально возможных вариантов.

Случайность. На наш взгляд именно эксплуатация в техническом плане такой категории как случайность может быть использована в качестве одного из наиболее эффективных инструментов, эмулирующих феномен инсайта. В этой связи, представляет интерес классификация типов случайности, предложенная О.А. Самойловой. В частности, речь идет об онтологической и гносеологической случайностях [Самойлова, 2015, с. 137]. Онтологическая случайность понимается как реально существующая в мире, природе случайность. Гносеологическая случайность обуславливается ограниченностью нашего знания. Последняя подразделяется автором на:

1. эмпирическую (ошибки измерения);
2. редуцированную (обусловлена наличием динамически равновесного хаоса);
3. концептуальную случайность (обусловлена ошибочным отражением реальности ввиду ограниченности теории);
4. ноуменальную (обусловлена пределом возможностей сознания и познания);
5. ошибка предпочтения (обусловленная давлением предыдущего опыта, взглядов и т.д.).

С точки зрения нашего исследования интерес представляют прежде всего концептуальная случайность и ошибка предпочтения. Самойлова видит два вида концептуальной случайности.

Во-первых, случайность, обусловленная неверным посылом, незнанием «природы изучаемого явления, связан с историчностью процесса познания, зависит от устоявшейся научной парадигмы» [Самойлова, 2015, с. 138].

Во-вторых, «это ошибка интерпретации, неверное истолкование полученных фактов, ошибочное обобщение эмпирического материала: слишком широкая индукция, ложные или, напротив, не выявленные закономерности» [там же].

Важным отличием концептуальной случайности является субъективность ее природы, наличие некоторого несоответствия фактов и нашей теории (конфигурации мира). Результатом этого является принципиальная возможность манипулирования элементами той или иной теоретической конструкции «без ущерба» для действительного положения вещей. Следовательно, наличие предпосылок для применения эвристического подхода и эксплуатация случайности в условиях неизвестности. Положительным моментом в этом случае является тот факт, что случайность может быть достаточно жестко ограничена рамками исходного теоретического построения.

Такой вывод, на наш взгляд, подкрепляет и выделение в отдельный вид ошибки предпочтения. Это связано с тем, что существование такого вида случайности дает возможность поиска и выделения некоторых правил, которые, с одной стороны, сами могут быть предметом манипулирования, а с другой – ограничивающим поле вариантов фактором. Оба вида случайностей – концептуальная и ошибка предпочтения – относятся Самойловой к устранимым, следовательно, управляемым случайностям.

Прежде чем продолжить дальше, необходимо напомнить о существовании проблемы «комбинаторного взрыва», которая неизбежно возникает в случае попытки решить что-либо с позиции «грубой силы» как

попытки осуществить полный перебор возможных вариантов. Мы можем обнаружить множество неудачных попыток решения данной проблемы. Например, Дрейфус пишет о принципиальной разнице между шахматной программой «Мак Хэк» и гроссмейстером [Дрейфус, 2010, с. 46]:

«Анализ шахматной программы «Мак Хэк, составленной Р. Гринблатом, наглядно иллюстрирует разницу между тем, как человек схватывает шахматную позицию, и тем, как машина производит сложный перебор. <...> В одной из партий турнира программа однажды крепко «задумалась», она затратила на ход 15 минут и посчитала 26 000 вариантов, в то время как человек может посчитать их только 100, от силы 200.»

Конечно же, со времени описываемого Дрейфусом эксперимента прошло много десятков лет, и сегодня машина (правда, если быть более точным, то команда разработчиков + суперкомпьютер Deep Blue) смогла обыграть в шахматы действующего чемпиона мира. Тем не менее, суть проблемы комбинаторного взрыва остается актуальной.

На наш взгляд, гипотеза о реализации принципа случайности как базового механизма генезиса новых правил может быть эффективной, при условии искусственного сокращения количества вариантов, что допускается принципом отсутствия гарантированного результата. А также за счет ряда других возможностей, о которых пойдет речь ниже.

Определяя акт генезиса как «случайный», мы хотим подчеркнуть, что он осуществляется с учетом граничных условий, которые казуально определяются рамками действительного положения дел как существующей системы логических строительных лесов Витгенштейна. В рамках такой системы задаются свойства, атрибуты и факты, определяя тем самым диапазоны вариативности новых объектов или правил, как теоретических конструкций.

Принимая тезис о случайности, мы с необходимостью приходим к необходимости решения проблемы множества альтернатив или комбинаторного взрыва. Как верно отмечает Саймон, реальные миры редко бывают аддитивными и трудно поддаются факторизации [Саймон, 2013, с. 84]. Сам исследователь фактически апеллирует к абдукции, оставаясь в тех же ограничениях, что и ДСМ-метод. Кроме того, на наш взгляд одной из основных нераскрытых проблем исследователя является проблема собственно генерации новых правил.

Как нам представляется, основными инструментами могут быть принципы преднамеренного нарушения правил, нонсенса и «пустого места» о которых речь пойдет ниже. Одной из важных отличительных черт этих принципов является имманентно заложенная в них возможность ограничения количества альтернатив.

Принцип преднамеренного нарушения правил. Относительно необходимости и даже неизбежности нарушения правил, позволим себе сослаться на Витгенштейна и его «Философские исследования» [Витгенштейн, 2018, § 201]:

«Наш парадокс был таким: ни один образ действий не мог бы определяться каким-то правилом, поскольку любой образ действий можно привести в соответствие с этим правилом.»

Нонсенс и «пустое место». У Жюль Делеза мы находим следующие слова [Делез, 1998, с. 175]:

«В мире есть нечто, заставляющее мыслить. Это нечто - объект основополагающей встречи, а не узнавания. Оно может быть постигнуто в различных эмоциональных тональностях <...> Но в своем первичном

качестве, независимо от тональности, оно может быть почувствовано. Именно в этом смысле оно противостоит узнаванию.»

По нашему мнению, относительно машинного интеллекта, когда нам необходимо найти некоторый аналог «чувствования», мы можем говорить об имитации «удивления», как аналога пустого результата, который возвращается поисковой подсистемой. Кроме того, причиной, порождающей такое чувство, а также фактором, подтверждающим то, что акт генезиса свершился могут выступить парадокс и нонсенс.

В основе понимания смысла Делезом лежат два тезиса. С одной стороны, смысл воспринимается им как «результат действий и страданий тел» [Делез, 2015, с. 48], а с другой - как особенное измерение предложения «*выражаемое предложения*, не сводимое ни к чему иному сущее, чистое событие, которое упорствует и обитает в предложении» [Делез, 2015, с. 32]. Отсюда мы приходим к событиям, как результатам действий и страданий тел. Выражениями этих событий являются бестелесные эффекты, которые мы и фиксируем, выражаем и подчиняем друг другу в нашем сознании. Они есть логические и диалектические атрибуты. Таким образом, вслед за стоиками, философ разделяет два плана бытия [Делез, 2015, с. 15]:

«Стоики соотносят причины с причинами и утверждают связь причин между собой (судьба). Они соотносят эффекты с эффектами и определенные связи между эффектами. Но эти две процедуры осуществляются не одинаково: бестелесные эффекты никогда не бывают причинами друг друга, они лишь «квазипричины», подчиняющиеся законам, выражающим, возможно, в каждом случае относительное единство или смесь тел, от которых эти эффекты зависят как от своих реальных причин».

Онтология Делеза дуальна, она включает в себя мир тел и мир эффектов. Только тела могут взаимодействовать в реальности.

Несколько по-иному выразил эту мысль Курт Хюбнер. Мир эффектов, выражаясь словами философа, является «контингентным а priori». Определяя различия между трансцендентализмом и операционализмом, он заключает, что [Хюбнер, 1994, с. 35]:

«...если для Канта способ такого производства объекта является а priori необходимым и неизменным, операционализм выводит его чисто методологически, исходя из задачи подчинения природы на основании её исследования. Соответственно а priori Канта может быть названо необходимым, а priori операционализма — контингентным, или произвольным».

Возвращаясь к Делезу, скажем, что смысл есть мгновенный «двойник предложения: вроде кэрролловской улыбки без кота» [Делез, 2015, с. 49].

Парадокс философом определяется как утверждение двух смыслов сразу [Делез, 2015, с. 9], сопровождаемое «смещением глубины, выведение события на поверхность и развертывание языка вдоль этого предела» [Делез, 2015, с. 19]. Среди всех типов парадоксов, с точки зрения генезиса новых правил на наш взгляд наибольший интерес представляет парадокс «неопределенного размножения» [Делез, 2015, с. 54]. Суть его заключается в том, что в какой-то момент мы обнаруживаем какой-либо эффект (ансамбль эффектов), который получает имя в результате акта десигнации (связывания слова и образа эффекта «это-то»), но при этом мы не можем его объяснить с помощью последующих имен более высокого ранга в процессе сериации. В результате образуется то, что философ обозначил как «нонсенс». Нонсенс есть имя, выражающее свой собственный смысл [Делез, 2015, с. 93]. Смысл замыкается на самом себе.

Таким образом, мы подошли к тому принципиальному понятию, которое, на наш взгляд является ключевым для понимания как минимум форм генезиса новых правил – сериации.

Каждое предложение можно рассматривать как сериальную форму или последовательность имен $n_1 \rightarrow n_2 \rightarrow n_3 \rightarrow \dots$, где каждое последующее имя обозначает смысл предшествующего [Делез, 2015, с. 49]:

«...каждое имя сначала берется с точки зрения того обозначения, которое оно осуществляет, а затем — того смысла, который оно выражает».

Нонсенс - N_n - есть случай, когда серия прерывается, так как элемент последовательности оказывается обладающим своим собственным смыслом, не обозначенным другим именем. Нонсенс «одновременно и говорит о чем-то, и высказывает смысл того, о чем говорит» [Делез, 2015, с. 93].

Философ выделяет две формы нонсенса:

1. имя есть «множество, включающее самого себя в качестве элемента» [Делез, 2015, с. 95];
2. имя «есть элемент, раскалывающий предполагаемое им множество — множества всех множеств, полковой брадобрей» [там же].

Отсюда философ делает вывод о том, что наличие нонсенса является основанием для двух видов последующего синтеза: регрессивного и дизъюнктивного. Первый является основанием для возникновения нового смысла, второй - для расщепления смысла на два.

Обнаружить нонсенс возможно только имея как минимум две серии, одна из которых является «означивающей», а другая - «означаемой». Такими сериями могут быть серии вещей и событий, либо серия предложений и серия вещей, которые описываются этими предложениями. Отдельно стоит выделить в качестве возможных пары двух предложений. Там самым

расширяется зона поиска смыслов за пределы вещей, до области языка и теоретических конструкций. В последнем случае смысл может быть получен вне привязки к вещам. Истина и смысл в понимании философа не одно и то же [Делез, 2015, с. 94]:

Иначе «...зачем тогда совершать восхождение от области истины к области смысла, если бы все дело заключалось только в том, чтобы отыскать между смыслом и нонсенсом отношение, аналогичное тому, что существует между истиной и ложью?»

Более того, смысл есть эффект, который *производится*, порождая остальные измерения предложения (сигнификацию, манифестацию и десигнацию) [Делез, 2015, с. 129-130]. При этом «смысл как условие истины» остается безразличным к истине и лжи [там же].

Таким образом, не имея возможности повлиять на законы мира действительного, мы можем конструировать из мира эффектов свой первоначальный мир в относительно широком диапазоне альтернатив, в котором смысл становится *порожденным* эффектом или событием, зависящем прежде всего от квазипричинности эффектов. *Соотнесение с истиной как соответствие с миром вещей является отдельной операцией*, где только последующая верификация может показать соответствие причин и эффектов. Что бы ни представляли собой эти серии они всегда неравнозначны. Эта неравнозначность и является «местом», где рождаются нонсенс, парадокс и смысл.

Одним из главных отличий означающего и означаемого является «привилегия» последнего [Делез, 2015, с. 56]:

«...не быть соотнесенным с независимым термином, ибо смысл как выраженное не существует вне выражения».

Расширим серийный подход за счет того, что *всегда* будем рассматривать как минимум две серии, даже если они могут меняться местами в последствии. Рассматриваемые серии *всегда* имеют несовпадения, при этом важно подчеркнуть: серии одновременны. Природа смещения или несовпадения кроется в том, что внутри каждой из серий термины имеют абсолютное место. Из этого возникает отсутствие или избыток того, что философ определяет как парадокс или «парадоксальная инстанция» [Делез, 2015, с. 59]. Именно она обеспечивает коммуникацию серий, будучи «смещенной в отношении самой себя» [там же].

Таким образом, проводя сравнительный анализ серий мы можем выявить эти несоответствия, локализованные как минимум несовпадением в занимаемом положении той или иной инстанции-элемента в этих сериях. Можно сказать, что мы имеем дело с некоторыми топологиями. И можно полагать, что применяемые математические методы в этом случае во многом могут быть выражены средствами математической топологии.

Таким образом, в дополнение к ранее сделанным выводам, мы можем добавить следующие:

1. Процедура первичного генезиса новых правил представляет собой сложный контингентный *a priori* процесс, состоящий из двух шагов: 1) стохастического (псевдослучайного) акта генезиса ансамбля новых правил-кандидатов; 2) процесса верификации полученных правил, с целью отбора правил, непротиворечащих эмпирическим данным.
2. Генерация правил может осуществляться как 1) генерация одиночного правила, 2) встраивания правила (ансамбля правил) в существующую серию 3) генерация новой серии правил.
3. Результатом процесса верификации может быть ансамбль более чем из одного правила, прошедшего верификацию (следствие парадоксов Делеза).

4. Реализация процесса первичного генезиса правил-кандидатов может быть схожа с реализацией алгоритма генерации псевдослучайных чисел, ограниченная свойствами, атрибутами и диапазоном допустимых значений исследуемого класса.
5. Ограничивающими факторами, частично препятствующими возникновению «комбинаторного взрыва», являются механизмы, эксплуатирующие концепцию нонсенса Делеза, поиска избытков и недостатков в исследуемых сериях.

§ *Онтологический аспект*

На первый взгляд может показаться, что сама постановка вопроса о каком-либо «месте» или регионе бытия, в отношении правил и понятий является в значительной степени надуманным. Тем не менее, в случае успеха в поисках и очерчивании границ его существования, можно было бы сделать заключение о существовании и определить его границы. Итогом этого, стали бы, как минимум, три вывода:

1. О принципиальной возможности существования такого места, ведет к возможности объективации в виде некоторой структуры, имущей характерные признаки, отвечающие за генезис новых концептуальных структур, правил и символов.
2. С точки зрения символьных методов, речь идет как минимум о существовании некоторого региона бытия, который возможно имитировать, пусть и с целым рядом ограничений.
3. С точки зрения коннекционистских моделей уточнило бы задачу поиска либо проектирования структур, ответственным за процессы логического вывода.

По мнению В.А. Смирнова [Смирнов, 1985], ни метафизический эмпиризм, ни рационализм, опирающийся на идеи предустановленной гармонии не сумели решить проблему достоверного аподиктического знания.

Таким образом, мы приходим к необходимости исследования того, каким образом активное познание может быть реализовано в виде каких-либо концептов в рамках машинного интеллекта. Прежде всего, в рамках настоящего исследования, нас интересует не столько проблема выводимости истины как таковая, сколько возможные идеи, которые могли бы быть реализованы в рамках машинного интеллекта в качестве коррелятов соответствующих регионов бытия.

Рассуждая об истории развития логики, Е.Д. Смирнова пишет о постепенном «превращении формальной логики в символическую» [Смирнова, 2021, с. 4]. Таким образом, исследование символьных методов ИИ становится тесно связанным с исследованиями в области логики, и прежде всего, с вопросом ее обоснования, который лежит, в первую очередь, в области теории познания. Этот вопрос можно поставить следующим образом [Смирнова, 2021, с. 5]:

«Является ли логика наукой эмпирической или теоретической? Имеет ли она собственный базис, или ее основания лежат в психологии, в теории познания, в математике?»

Таким образом, вопрос о месте или регионе бытия символьных методов, и связанных с этим понятий и правил лежит прежде всего в той же области, что и основные вопросы логики, и начало поисков во многом связано с критикой психологизма. Более того, как пишет Е.Д. Смирнова, в конце XIX психологизм стал господствующим направлением.

Действительно, рассуждая о том, как, каким образом существуют «Законы Причины и Действия» Милль говорит об «особых процессах духа в теории определения объективной истины» [Милль, 2019, с. 290] и о

необходимости перехода от психологии к логике. Одновременно с этим, философ пишет:

1. «логика есть наука об отправлениях разума, служащих для оценки очевидности» [Милль, 2019, с. 64];
2. «при всяком мышлении имеют место <...> процессы образования понятий, памяти и т.п.; но для логики нет необходимости обращать на это внимание» [Милль, 2019, с. 65].

Таким образом, можно заключить, что философ остается на позиции психологизма и не считает необходимым преодолеть это, сводя логику в основном к чисто практическим задачам из области «очевидности». В его понимании логика – это прежде всего искусство.

По мнению А.А. Шиян, для Гуссерля [Гуссерль, 2015]:

«...категориальная идеальная предметность существует независимо от нашего сознания и дается нам уже в простом восприятии, просто мы этого не замечаем».

Предметности суждения являются продуктом деятельности нашего сознания и не существуют в нашем мире. Таким образом, Гуссерль производит идеализацию логики, в которой логические связи относятся к сфере идеального. А.А. Шиян подчеркивает, что для философа «логические законы существуют объективно, но особым образом» [там же].

Одновременно с этим, идеальность логических связей у Гуссерля соотносится с «предметностью» категориального суждения, истоки которой лежат в актах творческой спонтанности., где спонтанность не есть случайность, но самопроизвольность Именно творческая спонтанность на основании предметности восприятия производит новый тип предметности,

что приводит к субъективизму, сочетаемому «с элементами платонизма» [Смирнова, 2021, с. 5].

Здесь важно отметить факт наличия у философа отдельного региона бытия, *«питающегося» из регионов восприятия человека, но существующего в достаточной степени самостоятельно.*

Не менее критично Е.Д. Смирнова относится к работам Яна Лукасевича, относя его к другой крайней точке зрения. Согласно взглядам последнего, логика рассматривается как теоретическая наука [Смирнова, 2021, с. 6]:

«...об объективных, идеальных связях и отношениях <...> логика вообще не имеет отношения к изучению законов и форм мышления, и ее трактовка как науки о законах и формах мышления есть возврат к психологизму.

Среди прочих концепций необходимо отметить существование таких альтернатив как объективно идеалистические и конвенционалистские подходы. Крайней формой последних, можно считать «принцип терпимости» Карнапа».

Определяя роль и место логической семантики, Е.Д. Смирнова делает два следующих вывода:

1. «Ни логика, ни логическая семантика не создают принципиально новых концепций истинности. Логическая семантика заимствует учение об истинности из теории познания, обрабатывает его для решения своих задач» [Смирнова, 2021, с. 7].
2. «Логика, как и математика, не являются эмпирической наукой. Ни первая, ни вторая не обосновываются ссылкой, что некто так рассуждает или вычисляет. <...> Сказанное не означает, что между логикой <...> и психологией (психологией мышления и психологией творчества) не

может быть взаимодействия. Однако ни психология, ни Computer Science не могут обосновать логику (равным образом математику и методологию)» [Смирнова, 2021, с. 7].

Важным моментом в анализе региона символических методов является тот факт, что развитие формальной логики привело к ее превращению в символическую. Работы Готлиба Фреге, Чарльза Пирса, Бертрانا Рассела и целого ряда других исследователей с необходимостью показывают связь этой новой, символической логики, с идеями активности познания, а также тесной связью с ролью языка и «других знаковых систем» [Смирнов, 1985].

Одним из наиболее важных механизмов в схеме познавательной деятельности Смирнов видел диалектический материализм. Знание понималось им не просто как результат воздействия внешних объектов на органы чувств, но как результат их взаимодействия, как [там же]:

«...систему с внешней обратной связью, т.е. не только объекты воздействуют на субъект, но и субъект воздействует на познаваемые объекты <...> Познание происходит на базе ранее добытого и социально закрепленного знания. Субъект познает мир сквозь призму выработанной предшествующими поколениями концептуальной системы, с помощью различного рода семиотических систем, и прежде всего естественного языка и специализированных языков науки».

При этом философ не соглашался с упрощенным пониманием теории отражения как копирования или зеркализации. В частности, он не соглашался с точкой зрения Людвиг Витгенштейна относительно понимания последним природы отражения как *проекции текущего положения дел в систему знаков*.

О необходимости начинать исследования в области логики начиная с языка говорил уже Милль. Определяя логику как часть искусства мышления, философ определяет язык как «одно из главных орудий или пособий мысли»

[Милль, 2011, с. 72], подразделяя знания на те, что происходят непосредственно из интуиции, и остальную, «большую часть знаний». Более того, рассуждение в ряде случаев вообще невозможно без посредства слова.

Как уже упоминалось выше, ученик и последователь Гуссерля - Финк - настаивает на *первичности* естественного языка. Вне зависимости от того, обращаемся ли мы к «раннему» или «позднему» Витгенштейну, язык остается главным действующим лицом в вопросах познания.

Естественный язык по необходимости является одним из основных претендентов на роль структуры, регион бытия которой и определяет положение символических методов как прямых наследников. *Именно законы, так или иначе укорененные в естественном языке, являются базой для символических методов.*

Объединяя выводы, полученные относительно эпистемологического и гносеологического аспектов проблемы, на наш взгляд, с необходимостью, можно сформулировать посылки:

1. Если и не полностью, то в значительной степени регион бытия процессов познания пересекается с регионом бытия символических систем.
2. В свою очередь регион бытия символических систем пересекается с регионом бытия естественного языка и языков науки.

которые, приводят к следующему предварительному выводу, дальнейшее исследование которого представляется важным для понимания роли символических методов:

Совпадение этих регионов приводит к возможности воспроизводства их в рамках машинного интеллекта как языковых моделей.

§ Информационный аспект

Говоря о проблеме общего ИИ нельзя обойти вниманием проблему коммуникации, и обмена информации, в частности. Во введении к работе «Внутри мыслящих миров» Ю.М. Лотман, с одной стороны, признает сложность и многоаспектность понятия интеллекта и интеллектуальной деятельности, а с другой стороны отмечает, возможность решения задачи выделения «некоторого инварианта интеллектуальности» путем обнаружения «ряда «мыслящих объектов»» в том случае, если ограничиться семиотическим аспектом проблемы. В частности, философ пишет [Лотман, 2014, с. 6]:

«...если ограничиться семиотическим аспектом, то задача эта представляется решимой».

Определяя с этой точки зрения интеллектуальную способность, по мнению Ю.М. Лотмана, можно свести ее к следующим функциям:

- a) Передаче имеющейся информации (текстов).
- b) Созданию новой информации, то есть создание текстов, не выводимых из уже имеющихся, а обладающих определенной степенью непредсказуемости.
- c) Памяти: способность хранить и воспроизводить информацию (тексты)».

Здесь мы находим сразу несколько важных аспектов и понятий, таких как «информация», «передача информации» или коммуникация.

Опуская такие проблемы как декларируемую еще Л.С. Выготским необходимость для функционирования интеллекта другого интеллекта [Выготский, 1982, с. 115], а также положение о необходимости наличия асимметрии для «мыслящих» семиотических структур [Лотман, 2014, с. 8] и ряда других аспектов, сосредоточимся непосредственно на прагматической

части вопроса: проблеме, сформулированной в пунктах 2 и 3: процессах производства и хранения информации.

В качестве отправного тезиса нашего исследования определим мысль Витгенштейна о том, что в пределах логики нельзя сказать ничего нового. Именно на этот факт обращает внимание Лотман, когда подчеркивает избыточность такой характеристики как «новизна» относительно обмена сообщениями. Процесс передачи сообщения, согласно Лотману, представляет собой процедуру перевода с языка L_1 , текста T_1 на язык L_2 и текст T_2 . В случае, если есть возможность «обратного» перевода без потери смысла, т.е. получения исходного текста T_1 новизна сообщения отсутствует [Лотман, 2014, с. 21]. В этом случае отсутствует сам акт обмена как таковой.

Таким образом, одним из необходимых условий акта обмена является наличие некоторой семиотической системы, а также наличие двух языков, имеющих отличные друг от друга семиотические структуры [Лотман, 2014, с. 206]:

«Диалог подразумевает асимметрию, асимметрия же выражается <...> в различии семиотической структуры (языка) участников диалога...»

Различие этих семиотических структур обусловлено неизбежно возникающими отличиями в предыдущем опыте. Более того, если мы даже ограничимся моделью «машина-машина», то, со всей очевидностью, полной идентичности можно добиться *только в первый и при этом одновременный*, момент начала деятельности этих систем.

Может показаться, что использование некоторого искусственно ограниченного языка может решить проблему различия семиотических систем. Однако в этом случае возникает парадокс. Суть его в том, что использование некоторого искусственного языка, во-первых, не может исключить неизбежного возникновения рассогласованности между семиотическими системами адресата и адресанта. В данном случае, не

принципиально, идет ли речь о моделях «человек-машина» или «человек-человек» Во-вторых, искусственность языка, которая на первый взгляд, могла бы снизить искажения, в действительности сужает предметную область обмена информацией. В этом случае [там же]:

«...семиотические структуры утрачивают способность обсуждать другие, присущие им в естественном состоянии».

Таким образом, даже попытка ограничиться искусственными языками не избавляет от проблем асимметрии семиотических структур.

На принципиальную важность новизны, а также семиотическую природу информации, опирались (но несколько в другом аспекте) и такие исследователи Клод Шеннон и Леон Бриллюэн.

Так, предлагаемая Шенноном теория информации скорее относится к атрибутивному подходу: в его понимании, это мера, снимающая неопределенность в приемнике информации. Автор теории дает возможность осуществить однозначное и, что немаловажно, симметричное преобразование информации, первоначально не имеющей материальную сущность, к физическим процессам, с помощью которых возможен перенос сигналов посредством понятия, близкого к понятию термодинамической энтропии, инструментальное измерение последней является уже инженерно решаемой задачей.

К «недостаткам» такого подхода относится невозможность расширить данную теорию до семиотической.

Насколько нам известно, на сегодняшний день не существует сколько-нибудь успешных попыток со стороны его последователей расширения предложенной им теории до семиотической. Тем не менее, идеи Шеннона представляют большой интерес в рамках обсуждения коммуникативного аспекта, о котором пойдет речь далее.

В целях настоящего исследования важно отметить тот факт, что теория информации Шеннона оперирует понятиями как «новизна» и «алфавит языка» в качестве объекта кодирования.

Так же, как и Шеннон, Бриллюэн в 1964 году говоря об информации, прибегает к идеям энтропии, негэнтропии, отталкиваясь от идей Больцмана и Карно [Бриллюэн, 2010]. Но в отличие от Шеннона, Бриллюэн сосредотачивает свое внимание не на проблемах передачи информации, а на самом феномене. Исследователь пишет [Бриллюэн, 2010, с. 31]:

«...негэнтропия эквивалентна информации <...> принцип Карно распространяется на эти превращения».

В этом его понимание близко к идеям Шеннона. Но исследователь идет дальше. Он подчеркивает, что в этом нет ничего житейского и такое понимание информации «никогда не спутаешь с наукой или знанием» [там же]. Очень важным отличием от Шеннона, является то, что исследователь говорит о возможности накопления информации, как и энергии, ее устаревании в условиях замкнутых систем (не получающих новую информацию извне).

Исходя из того, как может быть построена та или иная физическая система, мы можем утверждать, что она находится в том или ином неустойчивом состоянии [Бриллюэн, 2010, с. 35]:

«...предоставленная самой себе эта система эволюционирует к устойчивому состоянию, характеризующемуся минимальной энтропией. Первоначальная информация мало-помалу утрачивает свою ценность, а негэнтропия постепенно убывает. <...> Однако существует другой тип информации, носящий непреходящий (permanent) характер: великие открытия, научные законы».

Таким образом, исследователь различает два вида информации:

1) «скоропреходящая» (transitory), такая как прогноз погоды, биржевые котировки и т.д., и 2) непреходящая (permanent). Результатом мышления ученого, полагает автор является «в действительности создание новой информации» [Бриллюэн, 2010, с. 36]. Такого рода выводы являются крайне важными так как позволяют провести в достаточной степени прочные аналогии между информацией и элементами концептуальных структур. Иными словами, это позволяет построить модель, связывающую понятия, семиотические системы, понять природу жизненного цикла информации и механизмы ее передачи.

По нашему мнению, негэнтропийный подход позволяет дать объяснение сути природы предсказательной силы знания, и научных законов, в частности, опираясь на прямую связь негэнтропии и неопределенности.

Как философская категория информация, стала формироваться в 60-70-е годы XX века. На наш взгляд все позднейшие концепции так или иначе основывались именно на идеях Шеннона и Бриллюэна. С философских позиций все существующие подходы к информации можно, разделить на субстанциональные, атрибутивные и функциональные.

К последователям субстанционального подхода относится прежде всего Винер и его последователи, считающие, что информация представляет собой не энергию или материю, а нечто третье. Таким образом, информация рассматривается ими как некая самостоятельная сущность.

К последователям атрибутивного подхода относятся И.Б. Новик, А.Д. Урсул и др., считающие информацию всеобщим и неотъемлемым свойством материи. Так И.Б. Новик считал, что «информация представляет собой качественную и количественную характеристику организованности, упорядоченности отражения» [Новик, 1969].

А.Д. Урсул раскрывает сущность информации исходя из категорий различия и отражения [Урсул, 1973]. Философ обосновывает мысль о том, что

информация есть только там, где существует различие, и отсутствует там, где его нет.

Последователи функционального подхода к информации, такие как П.В. Копнин, Д.И. Дубровский, Б.С. Украинцев и др., напротив, полагают, что информация есть свойство лишь самоорганизующихся систем, связанное с их функционированием [Лысак, 2015].

Вне зависимости от того, какой из концепций дать предпочтение, все они дают основания для построения некоторой теоретической конструкции, в рамках которой информация представляет собой упорядоченные в виде структур системы, в узлах которых находятся факты-события как коррелятов «эффектов» Делеза; именно в процессе их сигнификации, десигнации и манифестации эти эффекты получают свои имена, значение и т.д..

Не менее важным является понимание информации с точки зрения новизны и ее сути в качестве необходимого условия для производства нового знания. Обнаружение новой информации связано с предположением о возможном наличии асимметрии у предполагаемого адресата. Суть этой асимметрии связана прежде всего с эффектом новизны, который должна вызвать передаваемая информация в системе адресата. Что, с необходимостью, приводит к обнаружению адресатом нонсенса, разрешение которого приводит к возникновению новых смыслов, имеющих в нашем случае выражение в форме новых понятий, правил и знаков.

К этому же, на наш взгляд, необходимо присовокупить проблему контингентности фактов и невозможности получения абсолютной истины. Мейясу на этот счет выразился следующим образом [Мейясу, 2008]:

«Я рационалист, и разум ясно показывает, что невозможно продемонстрировать необходимость законов. Так что нам остается верить разуму и принять это: законы не необходимы — они факты, а факты контингентны — они могут меняться беспричинно. Время не управляется

физическими законами, поэтому сами законы управляются этим сумасшедшим временем».

Условие априорной контингентности в еще большей степени усиливает тезис о невозможности сохранения идентичности двух мыслящих систем, независимо от того, речь идет о моделях «человек-человек» или «машина-машина».

Таким образом, мы можем сделать следующие выводы:

- 1. Обнаружение факта асимметрии концептуальной структуры адресата, относительно концептуальной системы адресанта должна пониматься как нонсенс, обуславливающий генезис новых понятий и правил.*
- 2. Та информация, что распознается в качестве нонсенса в концептуальной системе адресата, является триггером, запускающим производство (генезис) новых или корректировку существующих концептуальных понятий и правил.*

§ Коммуникативный аспект

Наиболее общая модель коммуникативного акта, будь то линейная инженерная модель Шеннона-Уивера [Shannon, 1949], или модель предложенная Р.О. Якобсоном [Лотман, 2014, с. 30], может быть рассмотрена сквозь призму техники. Мишель Фуко выделяет у Юргена Хабермаса три основные типа техник [Фуко, 1996, с.431]:

1. техники, позволяющие производить вещи;
2. техники, позволяющие использовать системы знаков;

3. техники, позволяющие определять поведения индивидов, предписывать им определенные конечные цели и задачи¹¹.

Именно такая трактовка коммуникации - как техники (использующей системы знаков) - дает возможность провести анализ элементов, входящих в каждый такой акт, с «технической» точки зрения задач общего ИИ.

Основными конструктивными элементами такой модели согласно, например, концепции Р.О. Якобсона и развитой в последствии Ю.М. Лотманом, являются цепочка: адресант (коммуникатор) ← некоторый кодирующий механизм ← канал передачи ← декодирующий механизм ← адресат.

Прежде всего рассмотрим первый элемент – адресанта - с точки зрения целей, которые он преследует, инициируя акт коммуникации. Джордж Герберт Мид определяет коммуникацию не просто как отправка-получение каких-либо сигналов, направленных на привлечение внимания (аналогично животным), а как обязательное условие обретения собственной индивидуальности. Важным условием такого рода коммуникации он называет опыт, как «коммуникацию в смысле значащих символов» [Мид, 1997, 167]. Целью коммуникации становится не просто согласование поведения, а согласование опыта, включая саморефлексию как пересборку своего Я. Именно в этом вновь полученный от другого опыт гарантирует согласованность действий индивида в обществе.

Важной отличительной чертой всех вышеупомянутых теорий является то, что в процессе коммуникации происходит не просто обмен фактами, осуществляется желание изменить поведение другого, происходит достижение согласия или даже истины, пусть и конвенциональной. Все это конечно же имеет место. Но наиболее важным здесь является то, как

¹¹ В целях исключения искажения исходного смысла предлагаемого автором (кем именно) мы считаем целесообразным привести данную им классификацию дословно не взирая на объем соответствующего фрагмента. Необходимая атрибуция, отсылающая к исходному тексту предпослана ему.

происходит этот процесс: через трансформацию как собственного, так и чужого опыта; трансформацию конфигураций персональных миров индивидуумов. Происходит это прежде всего путем символического обмена знаниями, полученными опытным путем за счет рефлексии, абстрагирования и т.д..

Юрген Хабермас определяет, научную коммуникацию как «рефлексивный процесс достижения взаимопонимания между личностями» [Шачин, 2019, с. 77]. В качестве отправной точки своей теории Хабермас положил понятие действия. В работе о Хабермase С.В. Шачин в этом отношении так передает мысль философа [Шачин, 2019, с. 95]:

«Действия могут быть инструментальными, то есть непосредственно изменяющими что-то в объективном мире в соответствии с выбранным человеком планом, либо коммуникативными, нацеленными на взаимопонимание между личностями с целью достижения общего определения ситуации и согласования их планов...».

Кроме того, отдельный интерес представляет собой анализ понятия «притязания на значимость». Философ выделяет 5 способов актуализации [Шачин, 2019, с. 104-105], подробное рассмотрение которых выходит за рамки настоящего исследования. Отметим, что сама по себе такая постановка вопроса предполагает рефлексивный и логический анализ со стороны адресата сути полученного сообщения, а не простое «добавление» полученной информации в некоторое хранилище данных.

Таким образом, сущность коммуникативного акта по необходимости предполагает ее семиотическую природу. Более того, в случае если мы предположим реализацию общего ИИ в рамках только коннекционистских методов, последние с необходимостью должны будут каким-либо образом эмулировать символичные методы, как необходимые инструменты, обеспечивающие процедуры логического вывода. Аналогичные рассуждения

являются в равной степени применимыми для анализа природы получателя-адресата.

Таким образом, необходимость символьных методов в процессе формирования сообщения обуславливается их символической природой. И такая актуальность присуща не только для таких достаточно гетерогенных систем моделей как «человек-машина», но и «машина-машина». Наглядным подтверждением этому является то огромное количество протоколов обмена данными которое существует сегодня в IT-индустрии сегодня. Со своей стороны выделим как минимум две основные причины:

1. необходимость унификации данных ввиду разнообразия структур баз данных и знаний, а также стандартов хранения данных;
2. необходимость сокращения объема передаваемой информации в целях оптимизации использования вычислительных ресурсов.

Последняя проблема также не является исключительно «машинной» или технической. Точность и краткость, а также ясность изложения имеют ключевое значение в процессе человеческого общения.

Относительно задачи унификации, отметим, что она с необходимостью является символьным процессом, или как минимум требующей какой-либо эмулируемым в случае коннекционистских методов, так как напрямую сопряжена с процессами оперирования символами специфичным для символических систем способами. Более подробное исследование данного вопроса выходит за рамки настоящего исследования.

Ю.М. Лотман под кодом понимает прежде всего язык. Следовательно, принимая, концепцию Базы знаний и Базы фактов В.К. Финна, в целях коммуникации возникает задача выражения содержания этих баз средствами некоторого языка, общего для адресанта и адресата. Здесь необходимо сразу отбросить возражение об избыточности такого шага в случае искусственных систем. Основными причинами для этого являются два следующих аргумента:

1. Кроме модели «машина-машина», существует модель «человек-машина», где в ходе осуществления коммуникативного акта существует необходимость преобразования информации, вызванная с принципиальными различиями их концептуальных структур.
2. Даже оставаясь в рамках модели «машина-машина», невозможно добиться полной гарантированной идентичности двух искусственных семиотических мыслящих систем ввиду того, что каждая из систем с необходимостью будет 1) иметь различный опыт, начиная с первой секунды своего существования; 2) технический прогресс с необходимостью приведет к отличиям в их технических реализациях.

Таким образом, необходимость решения проблемы языка и кодирования информации для ее передачи средствами этого языка существует с необходимостью.

Сама по себе проблема создания и последующей эволюции таких языков представляет собой отдельный предмет научных дискуссий. С точки зрения настоящего исследования представляет интерес сама уже постановка задачи о необходимости кодирования знаний и данных, как индивидуальную концептуальную структуру некоторого экземпляра общего ИИ.

Одним из основных критериев эффективности процесса кодирования, по нашему мнению, является обеспечение инвариантности смысла при любых трансформациях текста (или некоторой серии знаков). Ю.М. Лотман отдельно выделяет в качестве критерия эффективности работы такой системы то, насколько «адекватно» происходит передача информации. В этом случае текст выступает как «техническая упаковка» сообщения [Лотман, 2014, с. 19].

Так или иначе, процесс формирования сообщения является задачей по трансформации некоторых смыслов, в некоторый текст, или экспликации семантического содержания в форму синтаксической структуры, отчуждаемой от концептуальной системы адресанта.

Относительно «работы» механизма декодирования, отметим, что в целом он аналогичен механизму кодирования, но направленная в противоположно: в сторону обратной экспликации синтаксической структуры в связи с ее семантическим содержанием. Сам по себе перевод осуществляется с помощью наличествующей в той или иной культуре условной системы эквивалентностей. В связи с этим, М.Ю. Лотман выделяет проблему возможной взаимной непереводаемости кодов, содержащихся в концептуальных системах адресата и адресанта. [Лотман, 2014, с. 30].

Из всего вышеизложенного мы можем сделать следующие выводы:

- 1. Акт коммуникации невозможен исключительно в рамках коннекционистских подходов, по меньшей мере по той причине, что целью такого акта является изменение собственных представлений о мире, сформированных в виде внутренних концептуальных систем или индивидуальной конфигурации мира.*
- 2. Процедура формирования сообщения, связанная с необходимостью кодирования-декодирования информации, представляет собой процесс, специфический для символьных методов.*
- 3. Акт коммуникации, по крайней мере в рамках своей рациональной составляющей, с необходимостью возможен только в рамках семиотических систем, основное содержание которых представляет собой обмен информацией, закодированной в символической форме.*

Заключение

Концепцию связи между рациональностью и символической языковой природой нашего сознания в истории новоевропейской науки впервые, по-видимому, выдвинул В. фон Гумбольдт [Гумбольдт, 1985]. Ее развитие можно проследить в работах Э. Сепира и Б. Уорфа, Л. Витгенштейна и ряда других исследователей. Тезис о символической природе нашего сознания, как базовом условии реализации рационального поведения, получил дальнейшее воплощение в гипотезе Ньюэлла-Саймона относительно общего ИИ. По их мнению, символьная система является необходимым и достаточным условием для того, чтобы реализующая её машина считалась разумной.

В настоящем исследовании мы исходили из следующих тезисов:

1. Тезис В.К. Венда о приоритетности построения гибридных человеко-машинных систем, предназначенных для усиления и расширения возможностей человека.
2. Тезис В.Г. Кузнецова о том, что можно представить основные проблемы в области создания общего ИИ - 1) представления знаний и 2) поиска решений задач в пространстве возможных альтернатив - в качестве «головоломки» (в терминологии Т.Куна).
3. Тезис В.К. Финна и Ж.-Л. Люгера о необходимости в задачах создания общего ИИ ограничиться областью рационального, не вторгаясь в сферу эмоционального и психического.
4. Тезис Ж.-Л. Люгера о невозможности отказа от первичного обучения ИИ-системы.

В качестве результата, необходимого для обоснования роли символических методов, нами проведено обоснование выдвигаемого нами тезиса о необходимости понимания ИИ как техники, «подручной» человеку.

Мы показали, что такая постановка задачи даёт принципиальную возможность решения проблем построения систем ИИ в достаточно широком пространстве задач, ограниченных рамками гибридного, человеко-машинного взаимодействия.

В ходе настоящего исследования получены:

Во-первых, выводы относительно философских оснований, определяющих роль символических методов в задачах общего ИИ.

Во-вторых, сформированы правила генезиса новых понятий, правил и знаков (Приложение 1), закрепляющие специфическую роль символических методов в задачах общего ИИ.

В-третьих, сформулированные правила обладают практической значимостью, в ходе реализации задач по созданию прикладных систем общего ИИ.

Промежуточные выводы:

Выводы относительно философских оснований, необходимых для формирования правил генерации новых символов:

1. Совпадение регионов процессов познания и регионов бытия символических систем приводит к возможности воспроизводства их в рамках машинного интеллекта как языковых моделей.
2. Дано обоснование шага именованности как следующего после шага генезиса нового понятия.

Выводы относительно философских оснований, необходимых для понимания роли символьных методов в коммуникационных процессах:

1. Обнаружение факта асимметрии концептуальной структуры адресата, относительно концептуальной системы адресанта должна пониматься как нонсенс, обуславливающий генезис новых понятий и правил.
2. Та информация, что распознается в качестве нонсенса в концептуальной системе адресата, является триггером, запускающим производство (генезис) новых или корректировку существующих концептуальных понятий и правил.

О роли символьных методов в коммуникационных процессах:

3. Процедура формирования сообщения, связанная с необходимостью кодирования-декодирования информации, представляет собой процесс, специфичный для символьных методов.
4. Акт коммуникации, по крайней мере в рамках своей рациональной составляющей, с необходимостью возможен только в рамках двух семиотических систем.
5. Акт коммуникации невозможен исключительно в рамках коннекционистских подходов, по меньшей мере по той причине, что цель такого акта является не просто обмен информацией, а изменение собственных представлений о мире, сформированных в виде внутренних концептуальных систем.
6. Коммуникативный акт реализуется специфичными для символьных методов средствами, как процесс передачи структурных элементов тех или иных концептуальных систем.

Общие выводы:

1. Основной ролью символьных методов как необходимых компонентов в задачах общего ИИ является обеспечение процессов как порождения новых понятий, правил и символов (знаков) как элементов концептуальных систем.
2. Символьные методы являются первичными триггерами, которые запускают тот или иной «мыслительный» процесс, обеспечивают процесс порождения новых смыслов, а также оценку полученных результатов.
3. Гипотезу Ньюэлла-Саймона необходимо уточнить в части понятия «разумного в широком смысле» поведения». А именно: ограничиться пониманием разумного поведения как рационального.
4. Любая система ИИ, приближающаяся к имитации рационального поведения: а) должна быть гибридной (человеко-машинной), б) должна включать символьные методы в качестве основных механизмов порождения новых знаний.
5. Концептуальные системы человека должны определять граничные условия процесса самообучения ИИ-системы, и их нарушение приводит к невозможности существования ИИ-системы или потере ценности полученных результатов.
6. Механизмы символьных методов могут обеспечивать:
 - a. процессы обучения и самообучения символических систем на основании обучающих выборок, аналогично тому, как это происходит при коннекционистских подходах;
 - b. процессы генезиса новых представлений, знаков и правил;
 - c. процессы коммуникации в рамках моделей «человек-машина» и «машина-машина».

Завершая подведение итогов, наметим перспективные направления дальнейших исследований. В результате проведенных исследований можно

по-новому взглянуть на механизмы обучения (в том числе, на область машинного обучения ИИ-систем).

В частности, представляется важным и необходимым исследовать факторы случайности и дискретности в процессах генезиса новых имен, процессах генезиса новых понятий правил как модели творческих способностей человека.

Также, на наш взгляд, возможно, исследуя проблематику парадокса и нонсенса, намеченную и в общем виде развёрнутую Делезом, оформить их критерии в виде конкретных правил, дополняющих, расширяющих и уточняющих правила, указанные нами в составе промежуточных итогов.

Приложение 1 Правила генезиса новых понятий, правил и знаков.

Выводы относительно машинной реализации рационального поведения в условиях недостаточности информации:

Правила генезиса новых понятий — элементов концептуальной системы:

1. Факт обнаружения несоответствия между текущим состоянием концептуальной структуры ИИ-системы и поступающей извне информации является необходимым условием активирующим процессы генезиса новых понятий.
2. Первичный акт генезиса может происходить путем преднамеренного нарушения существующей концептуальной системы за счет а) реализации принципа случайности при, б) отсутствии требования обязательности получения результата.
3. В качестве триггера, вызывающего начало акта генезиса и, одновременно с этим, верифицирующего факт того, что акт свершился выступает нонсенс или обнаружение «пустого места» в понимании Делеза.
4. Полученные ИИ-системой выводы и новые знания об окружающей среде не должны быть одобрены человеком, и не противоречить знаниям полученным непосредственно от человека.

Правила первичного генезиса новых символов:

1. ИИ-система имеет право давать имена новым понятиям как элементам концептуальных систем.

2. ИИ-система обязана получать одобрение со стороны человека относительно корректности имен; при наличии разногласий система обязан принять точку зрения человека.
3. Генезис новых имён активируется фактом завершения генезиса новых понятийных элементов концептуальной системы.
4. Акт генезиса новых имен является свободным контингентным актом внутри ограничений, установленных языком, на котором осуществляется именование.
5. Принцип контингентности может быть реализован как псевдослучайный процесс, реализованный на принципах существующих алгоритмов генерации случайных чисел.

Правила первичного генезиса новых продукционных правил:

1. Процедура первичного генезиса новых правил представляет собой сложный контингентный a priori процесс, состоящий из двух шагов: 1) стохастического (псевдослучайного) акта генезиса ансамбля новых правил-кандидатов; 2) процесса верификации полученных правил, с целью отбора правил, непротиворечащих эмпирическим данным.
2. Генерация правил может осуществляться как 1) генерация одиночного правила, 2) встраивания правила (ансамбля правил) в существующую серию 3) генерация новой серии правил.
3. Результатом процесса верификации может быть ансамбль более чем из одного достоверного правила, прошедшего верификацию (следствие парадоксов Делеза).
4. Принцип контингентности может быть реализован как (псевдо)случайный процесс, реализованный на принципах аналогичных существующим алгоритмам генерации случайных чисел, ограниченный свойствами, атрибутами и диапазоном допустимых значений исследуемого класса.

5. Ограничивающими факторами, частично препятствующими возникновению «комбинаторного взрыва», являются механизмы, эксплуатирующие концепцию нонсенса Делеза, поиска избытков и недостатков в исследуемых сериях.
6. Если в вопросах, так или иначе связанных с проблемой генезиса новых символов процесс в значительной степени ограничен рамками феноменологии, эстетики, квалиа (физиологией, или конструкцией сенсоров в случае машинного интеллекта) и т.д., то в задачах генезиса новых правил такого рода ограничений или не существует, или они несущественны.

Список литературы

1. Arnheim F., *Intelligence Simulated*. - "Midway", University of Chicago, June 1967
2. Badiou A. *L'être et l'événement*. P.: Seuil, 1988.
3. Barr A. *Handbook of Artificial Intelligence* / E. Feigenbaum – Los Altos, CA: William Kaufman, 1989.
4. Blum M., Ahn von L. A., Langford J., —The CAPTCHA Project, Completely Automatic Public Turing Test to Tell Computers and Humans Apart, 11 Nov. 2000
5. Cassirer E., *An Essay on Man*., New York., 1956
6. Cole D., "The Chinese Room Argument" // *The Stanford Encyclopedia of Philosophy* (Winter 2020 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2020/entries/chinese-room/>>.
7. Confalonieri Robe, Ludovik Çoba, Benedikt Wagner and Tarek R. Besold. "A historical perspective of explainable Artificial Intelligence." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 11 (2021): n. pag.
8. Dreyfus H., Why Heideggerian AI failed and how fixing it would require making it more Heideggerian // *Artificial Intelligence*, Volume 171, Issue 18, 2007, pp 1137-1160, <https://doi.org/10.1016/j.artint.2007.10.012>.
9. Gugerty L., *Newell and Simon's Logic Theorist: Historical Background and Impact on Cognitive Modeling*// *Human Factors and Ergonomics Society Annual Meeting Proceedings*, October 2006
10. Guilford J.P. Three faces of intellect. *American Psychologist*. 1959. № 14.
11. Fjelland, R. Why general artificial intelligence will not be realized. *Humanit Soc Sci Commun* 7, 10 (2020). <https://doi.org/10.1057/s41599-020-0494-4>
12. Fodor, J, *Methodological Solipsism Considered as a Research Strategy in Cognitive Science*, "Behavioral and Brain Sciences, 1980, vol 3, p.63-73.

13. Freeman W., The physiology of perception, Scientific American, Vol. 242, February 1991
14. Freeman, W. Neurodynamic Models of Brain in Psychiatry. Neuropsychopharmacol 28, S54–S63 (2003). <https://doi.org/10.1038/sj.npp.1300147>
15. Harnard S., Minds, machines and Searle // Journal of Experimental and Theoretical Artificial Intelligence, 1989, Vol 1, p 5-25
16. Haugeland J., Artificial Intelligence: The Very Idea (MIT Press, Cambridge, MA, 1985)
17. Koza J.R. Genetic Programming. Cambridge: MA:MIT Press, 1992.
18. Lighthill R., Artificial Intelligence: A General Survey // Artificial Intelligence: a paper symposium, Science Research Council, 1973
19. Minsky M., ACM Turing Lecture. Form and Content in Computer Science. Massachusetts Institute of Technology, Cambridge, Massachusetts, 1970
20. McDowell J.H., Mind and World. Camb., 1994
21. McCartney S., [ENIAC: The Triumphs and Tragedies of the World's First Computer.](#), Walker & Co., 1999
22. McCarthy J., What is artificial intelligence? 2007, <http://www-formal.stanford.edu/jmc/whatisai/>
23. McCulloch W., Pitts W., A logical calculus of the ideas immanent in nervous activity // Bulletin of Mathematical Biology. — New York: Springer New York, 1943. — T. 5, № 4. — C. 115—133.
24. Miller G. A., The cognitive revolution: A historical perspective. TRENDS in Cognitive Sciences, 7(3), 2003, 141- 144.
25. Minsky M., Why A.I. Is Brain-Dead//Wired, issue 01.08.2003
26. Newell A. and Simon H., Computer science as empirical inquiry: Symbols and search. Communications of the ACM 19:113-126, 1976.
27. Newell, A. & Simon, H., The logic theory machine: A complex information processing system. IRE Transactions on Information Theory, 1956, 2, 61-79.

28. Shannon C. E., Weaver W. The Mathematical Theory of Communication. Urbana. Univ. of Illinois Press., 1949.
29. Shannon C., Chess playing machine // Scientific American, 1950, vol. 2., p. 15
30. Scheer Joseph K (ed.), The McDowell-Dreyfus Debate. Routledge, 2013
31. Searle, J. Minds, brains, and programs // Behavioral and brain sciences., 1980., T. 3, № 3 (September), P. 417-424.
32. Stove D. Idealism: A Victorian Horror Story (Part Two) // The Plato Cult and Other Philosophical Follies/D. Stove (ed.). Oxford: Blackwell, 1991.
33. Parsing the Turing Test, Springer Science + Business Media B.V. 2008
34. Popper K. Indeterminism in Quantum Physics and in Classical Physics. // The British Journal for the Philosophy of Science. Edinburgh — London, 1950. Vol. I, № 2, pp. 117–133;
35. Popper K, The Logic of Scientific Discovery Basic Books, New York, Hutchmson & Co, Ltd, London, 1980.
36. Ryle, G., Concept of mind., London ; New York : Hutchinson's University Library, 1949
37. Stiegler B. La société automatique. Fayard, 2015. T. 1. L'avenir du travail.
38. Stiegler B., Technics and Time, 1: The Fault of Epimetheus., Stanford: Stanford University Press, 2016
39. Turing A., Computing machinery and intelligence, Mind 1950; LIX (236): 433-460 doi:10.1093/mind/LIX.236.433
40. Turing A. M. On Computable Numbers, with an Application to the Entscheidungsproblem. // Proceedings of the London Mathematical Society — [London Mathematical Society](#), 1938. — Vol. s2-43, Iss. 6. — P. 544—546. — ISSN [0024-6115](#)
41. Quach K., Meta trains data2vec neural network to understand speech, images, text so it can 'understand the world'//The Register, Fri 21 Jan 2022 // 23:44 UTC https://www.theregister.com/2022/01/21/meta_data2vec_algorithm/
42. Quillian R. Semantic memory, In: M. Minsky (ed.) Semantic information processing

43. Quine W. V. From a Logical Point of View. New York: Harper, 1963, pp. 20—46.
44. Wheeler M., *Reconstructing the Cognitive World: The Next Step* (Cambridge, MA: A Bradford Book, The MIT Press, 2007)
45. Wright J. The Turing Bombe Victory and the first naval Enigma decrypts, *Cryptologia*, 2017
46. Алексеев А. Ю. Машина Корсакова (1832 г.) как прототип мультиагентного суперкомпьютерного автомата // *Искусственные общества*. – 2019. – Т. 14. – Выпуск 1.
47. Анохин К.В., Когнитом: в поисках фундаментальной нейронаучной теории сознания // *Журнал высшей нервной деятельности им. И.П. Павлова*, 2021, Т. 71, № 1, стр. 39-71
48. Апель К.-О., *Трансформация философии*, М.: «Логос», 2001
49. Асмус В.Ф. *Лекции по истории логики: Авиценна, Бэкон, Гоббс, Декарт, Паскаль*, М.: Издательство ЛКИ, 2019
50. Бостром Н., *Искусственный интеллект. Этапы. Угрозы. Стратегии*. М.: Манн, Иванов и Фербер, 2016
51. Бергсон А. «Творческая эволюция». Академический проект: Трикса, 2019
52. Бергсон А., *Мысль и движущееся: Статьи и выступления*, М.: Центр гуманитарных инициатив, 2019
53. Брасье Р., *Понятия и объекты* // *Философско-литературный журнал «Логос»*. 2017. №3 (118).
54. Бриллюэн Л., *Научная неопределенность и информация*, М.: Книжный дом «ЛИБРОКОМ», 2010
55. Васюков В.Л., *Формальная онтология и искусственный интеллект* // *Искусственный интеллект: междисциплинарный подход*, М.: ИИнтелЛ, 2006
56. Витгенштейн Л., *Логико-философский трактат*, М.: «Канон+» РООИ «Реабилитация», 2011

57. Витгенштейн Л., *Философские исследования*, М.:Издательство АСТ, 2018
58. Величковский Б.М., *Когнитивная наука. Основы психологии познания. В 2 томах., Т. 2.*, М.:Издательство Юрайт, 2022
59. Венда В.Ф., *Системы гибридного интеллекта: Эволюция, психология, информатика.*, М.:ЛЕНАНД, 2020
60. Вертгеймер М., *Продуктивное мышление.*, М.:Прогресс, 1987
61. Выготский Л.С., *Собрание сочинений, том 6*, М.1982
62. Габриель М., *Я не есть мозг: Философия духа для XX века.*, М.:УРСС: ЛЕНАНД, 2020
63. Гаджиев А. М., *Применение дсм-метода машинного обучения для класса задач экономического прогнозирования // Инновации и инвестиции. 2015. №5.*
64. Гиренок Ф. И., *Подступы к философии человека // Научные труды Московского гуманитарного университета. 2018. №2.*
65. Германтова А.Д. *Логика: учебник для бакалавров.*, М.:Издательство «Омега-Л», 2015
66. Гридин В.Н., Солодовников В.И., *Совместное использование нейросетевых технологий и деревьев решений для поиска логических закономерностей в данных. / В. Н. Гридин, В. И. Солодовников // Информационные технологии и нанотехнологии. 2017. С. 1756–1762.*
67. Гоббс Т. *Сочинения в 2 т., т.2*, М.: Мысль, 1989
68. Горохов В.Г., *Петр Клементьевич Энгельмейер. Инженер-механик и философ науки.*, М.: Наука, 1997
69. Горохов В.Г., *Философия и история науки.*, Дубна: Издательство Объединённого института ядерных исследований, 2012.
70. Гребенюк О.С., Гребенюк Т.Б., *Теория обучения: Учеб. для студ. высш. учеб. заведений.*, М.: Изд-во ВЛАДОС-ПРЕСС, 2003.
71. Гумбольдт В., *Язык и философия культуры*, М.:Прогресс, 1985

72. Гуссерль Э. Идеи к чистой феноменологии и феноменологической философии. Книга первая., М.: Академический Проект, 2009.
73. Доманов О.А., Ален Бадью между формализмом и интуиционизмом // Вестн. Том. гос. ун-та. Философия. Социология. Политология. 2012. №2 (18).
74. Домингос П., Верховный алгоритм: как машинное обучение изменит наш мир. М.: Манн, Иванов и Фербер, 2016.
75. Дотцлер Б. Технонотация: Бэббидж и могущество цифр // Философско-литературный журнал «Логос». 2010. №1 (74).
76. Дрейфус Х., Чего не могут вычислительные машины: Критика искусственного разума., М.: ЛИБРОКОМ, 2010.
77. Делез Ж. Логика смысла, М.: Академический проект, 2015
78. Делез Ж., Различие. повторение, ТОО ТК «Петрополис», 1998
79. Дессауэр Ф., К философии техники. Что есть техника? - термин и сущность // Онтология проектирования. 2016., том 6, №3 (21).
80. Деннет Д., Сладкие грезы: Чем философия мешает науке о сознании., М.: УРСС: ЛЕНАНД, 2017
81. Журавлева А.С. Универсальная характеристика в метафизическом проекте Г. В. Лейбница // Вестник Пермского университета. Философия. Психология. Социология. 2012. №2.
82. Иванов Д.В., Критика метафизики в контексте дискуссий о реализме и антиреализме // Философский журнал. 2017. №1.
83. Иванов Н.В. О двух подходах в современной семиотике (к философским основаниям построения кумулятивной модели семиогенеза) / Н.В. Иванов // Язык в пространстве коммуникации и культуры. Материалы VI Международной научной конференции по актуальным проблемам теории языка и коммуникации. 29 июня 2012 года. - М.: ЗАО "Книга и бизнес", 2012. - С. 27-37
84. Кант Э. Собрание сочинений в восьми томах, Том 3, М.: Чоро, 1994

- 85.Капп Э., Гунов Г., Нуарье Л., Эспинас А. Роль орудия в развитии человека. - Л. - 1925.
- 86.Климова, С.Г., Михеенкова, М.А. и Финн, В.К. 2016. ДСМ-метод в качественном социологическом исследовании: основные принципы и опыт использования. Социологический журнал. 22, 2 (июн. 2016), С. 8-30.
- 87.Клини С. К. Математическая логика. — М.: Мир, 1973.
- 88.Койре А. От замкнутого мира к бесконечной вселенной., М.: Издательство «Логос», 2001
- 89.Коплстон Ф., История философии. Средние века., М.: ЗАО Центрполиграф, 2003
- 90.Корниенко М. А. Лингвофилософия Ноама Хомского: от картезианской традиции к генеративной грамматике // Вестн. Том. гос. ун-та. Философия. Социология. Политология. 2018. №43.
- 91.Кнут Д. Э., Информатика и ее связь с математикой. / Современные проблемы математики. Сб. статей. М.: Знание, 1977. С. 4–32.
- 92.Кнут Д. Э., Искусство программирования, том 1. Основные алгоритмы., М.: издательский дом «Вильямс», 2001
- 93.Кудряшев А.Ф., Елхова О.И. Процесс творчества в системах с искусственным интеллектом // Вестник Башкирск. ун-та. 2016. №4. URL: <https://cyberleninka.ru/article/n/protsess-tvorchestva-v-sistemah-s-iskusstvennym-intellektom> (дата обращения: 30.01.2022).
- 94.Краевский В.В., Методология педагогики: новый этап: учеб. Пособие для студ. высш. учеб. раведений / В.В. Краевский, Е.В. Бережнова, М.: Издательский дом «Академия», 2006
- 95.Кузнецов В.Г., Парадигмальный анализ теории искусственного интеллекта //Философия искусственного Интеллект. Труды всероссийской междисциплинарной конференции, посвященной шестидесятилетию исследований искусственного интеллекта, 17-18 марта 2016 г, М.:ИИнтелл, 2017.

96. Культурология. XX век: Антология, М.: Юрист, 1995
97. Кучинов Е. В., Закаблуковский Е.В. Техника будущего (Симондон, Гваттари, брайант, Стиглер) // Манускрипт. 2016. №8 (70).
98. Лапшин И.И., Философия изобретения и изобретение в философии. Введение в историю философии, М., «Республика», 1999 г., с. 312-313.
99. Латур Б., Пересборка социального: введение в акторно-сетевую теорию., М.:Изд. дом Высшей школы экономики, 2020
100. Лорьер Ж.-Л., Системы искусственного интеллекта., М.:Мир, 1991
101. Луман Н. Л., Общество как социальная система. Пер. с нем./ А. Антоновский. М: Издательство "Логос". 2004.
102. Лысак И.В., Информация как общенаучное и философское понятие: основные подходы к определению // Философские проблемы информационных технологий и киберпространства. 2015. №2.
103. Лурия А.Р., Язык и сознание.,СПб.: Питер, 2021
104. Лотман Ю.М., Внутри мыслящих миров, СПб.: Азбука, Азбука-Атикус, 2014
105. Люгер Дж.Ф., Искусственный интеллект: стратегии и методы решения сложных проблем., М.: Издательский дом «Вильямс», 2003.
106. Маркова Л.А. О метафизических основаниях аналитической философии (К. -О. Апель и его оппоненты) // Философия науки и техники. 2009. №1. URL: <https://cyberleninka.ru/article/n/o-metafizicheskikh-osnovaniyah-analiticheskoy-filosofii-k-o-apel-i-ego-opponenty> (дата обращения: 27.01.2022).
107. Ляшенко Д. Н. Семантика и структурная онтология / Д. Н. Ляшенко // Известия Уральского государственного университета. Сер. 3, Общественные науки. — 2011. — № 4 (97). — С. 3-16.
108. Мелик-Гайказян И.В., Семиотическая диагностика аттракторов конвергентных технологий // Философия искусственного Интеллекта. Труды всероссийской междисциплинарной конференции, посвященной

- шестидесятилетию исследований искусственного интеллекта, 17-18 марта 2016 г, М.:ИИнтелл, 2017.
109. Мид Дж Гр, Разум, я и общество (главы из книги) //Социальные и гуманитарные науки. Социология, Москва, 1994, № 4,
110. Мейясу К. Время без становления//Доклад в Университете Мидлсекса в рамках семинара Центра исследования современной европейской философии, 08.05.2008
111. Мейясу К., После конечности: Эссе о необходимости контингентности., М.: Кабинетный ученый, 2015
112. Минский М., Сообщество разума., М.: Издательство АСТ, 2018
113. Ортега-и-Гассет Х. Избранные труды., М.:Издательство «Весь мир», 1997
114. Милль Дж. Ст., О некоторых важных вопросах познания., М.:Книжный дом «ЛИБРОКОМ», 2019
115. Милль Дж Ст., Система логики силлогистической и индуктивной: изложение принципов и доказательства в связи с методами научного исследования., М.:ЛЕНАНД, 2011
116. Миллнер Д., Галантер Ю., Прибрам К., Планы и структура поведения., М.:Прогресс, 1965
117. Никитина С.Е., Семантический анализ языка науки на материале лингвистики., М.:УРСС: Книжный дом «ЛИБРОКОМ», 2016
118. Невважай И.Д., Проблема региональных онтологий в современном естествознании // Философия науки. Вып. 14: Онтология науки. М.: ИФ РАН, 2009.
119. Непейвода Н.Н., Вызовы логики и математики XX века и «ответ» на них цивилизации // Вопросы философии., 2005, No 8, стр. 118-128.
120. Новик И.Б. Философские идеи Ленина и кибернетика. М.: Знание, 1969. – 48 с. (Novik I.B. Lenin's philosophical ideas and cybernetics. М.: Znanie, 1969. – 48 p.)

121. Пенроуз Р., Новый ум короля. О компьютерах, мышлении и законах физики, М.:Ленанд, 2020
122. Петренко А. К., Петренко О. Л. Машина Беббиджа и возникновение программирования // Историко-математические исследования.,1979., т. 24., с. 340.
123. Петрунин, Ю. Ю. Искусственный интеллект / Ю. Ю. Петрунин // Новая философская энциклопедия: в 4 т. Т. 2 / Ин-т философии РАН, Нац. общ.-науч. фонд; науч.-ред. совет: пред. В. С. Степин, зам. пред.: А. А. Гусейнов, Г. Ю. Семигин, уч. секр. А. П. Огурцов. – М.: Мысль, 2010. – С. 159–160.
124. Пирс Ч.С. Из работы «Элементы логики». *Grammatica speculativa*. – В кн.: Семиотика. М., 1983;
125. Платон, Собрание сочинений в 4 т. Т.2., М.: Издательство социально-экономической литературы «Мысль», 1970
126. Поппер К., Объективное знание. Эволюционный подход. М.: Эдиториалл УРСС, 2002
127. Поспелов Д.А., Десять «горячих точек» в исследованиях по искусственному интеллекту // Искусственный интеллект и принятие решений, М., 2019., № 4.
128. Поспелов Д.А., Фантазия или наука: на пути к искусственному интеллекту., М.:Наука. Главная редакция физико-математической литературы., 1982
129. Поспелов Д.А., Ситуационное управление: Теория и практика. М.:ЛЕНАНД, 2021.
130. Пуанкаре А., Кутюа Л., Математика и логика. М.: Издательство ЛКИ, 2018.
131. Рассел Б., История западной философии., М.:Издательство АСТ, 2018
132. Рассел С. Норвиг П., Искусственный интеллект: современный подход., М.: ООО «И.Д. Вильямс», 2016

133. Родзин С.И. Генетическое программирование и проблемы синтеза программ // Известия ЮФУ. Технические науки. 2002. №3.
134. Рубашкин В.Ш., Лахути Д.Г. Семантический (концептуальный) словарь для информационных технологий. Ч. 1 // НТИ. Сер. 2.—1998.— № 1.— С.19-24.
135. Саймон Г., Науки об искусственном., М.: Книга по требованию, 2013
136. Самойлова О.А., Гносеологическая случайность // НАУ. 2015. №9-3 (14).
137. Свасьян К.А., Философия символических форм Э. Кассирера: Критический анализ, М.:Академический Проект; Альма Матер, 2010
138. Силен Д., Мейсман А., Основы Data Science и Big Data. Python и наука о данных., М.: Издательство Питер, 2017.
139. Симондон Ж., О способе существования технических объектов / сокращ. пер. и коммент. М. Куртова [Электронный ресурс] // Транслит. – 2011. – № 9. – URL: <http://www.translit.info/materialy/9-vypuski/zhilber-simondon-o-sposobe-sushhestvovaniya-tehnicheskikh-obektov> (дата обращения: 13.03.2017).
140. Седых О.М., Хаменков М.А., Органопроекция: русский контекст // Философия науки и техники, 2016. Т. 21. № 1. С. 132–151
141. Сериков В.В., Станет ли педагогика наукой? Размышления о методологии В.В.Краевского // Отечественная и зарубежная педагогика. 2016. №5 (32).
142. Сёрль Дж. Рациональность в действии., М.:Прогресс-Традиция, 2004
143. Семенов С. П., Славский В. В., Татаринцев П. Б. Системы компьютерной математики. Учебное пособие для студентов математического факультета АГУ Барнаул: Изд-во Алт. ун-та, 2004.

144. Степин В.С. Научные революции как точки бифуркации в развитии знания. // Научные революции в динамике культуры. — Минск, 1987. С. 38–76.
145. Смирнов В.А., О взаимоотношении символической логики и философии // Философия в современном мире. Философия и логика. М., 1974. С. 5-35.
146. Смирнов В.А., О перспективах анализа учения И.Канта о праве и морали средствами современной логики // Кантовский сборник. Вып. 15. Калининград, 1990. С. 68-73.
147. Смирнов В.А., Символическая логика и теория познания // Философия и социология науки и техники, Ежегодник 1983, М., 1985., с. 45-60
148. Смирнов В.А., Уровни знания и этапы процесса познания // Проблемы логики и научного знания., М., 1964. с. 23-52
149. Смирнова Е.Д., Логическая семантика и философские основания логики., М.:ЛЕНАНД, 2021
150. Спиридонов В.Ф., Задачи, эвристики, инсайт и другие непонятные вещи // Философско-литературный журнал «Логос». 2014. №1 (97).
151. Степин В.С., Теоретическое знание. М.: «Прогресс-Традиция», 2000
152. Степин В.С. Философия и методология науки., М.: Академический проект; Альма Матер, 2015
153. Серр М., Девочка с пальчик, М.:Ад Маргинем Пресс, 2016
154. Твардовский К., Логико-философское и психологическое исследование. М.:РОССПЭН, 1997
155. Соссюр Ф. де., Курс общей лингвистики. Труды по языкознанию. М.: Прогресс., 1977
156. Терешкун О.Ф. Антропологическое осмысление техники (Э. Капп и П. Флоренский) // Приволжский научный вестник. 2013. №9 (25).

157. Урсул А.Д. Отражение и информация. М.: Мысль, 1973. – 231 с.
(Ursul A.D. Reflection and Information. M.: Mysl', 1973. – 231 p.)
158. Уотерман Д., Руководство по экспертным системам. М.: Мир, 1989
159. Философия искусственного интеллекта. Труды Всероссийской междисциплинарной Конференции, посвященной шестидесятилетию исследований искусственного интеллекта, 17-18 марта 2016 г., Философский факультет МГУ имени М.В. Ломоносова / Под ред. В.А. Лекторского, Д.И. Дубровского, А.Ю. Алексеева. М.:ИИнтел, 2017
160. Финн В.К. Индуктивные методы Д.С. Милля в системах искусственного интеллекта. Часть I // Искусственный интеллект и принятие решений. №3, 2010, стр.3-21.
161. Финн В.К. Искусственный интеллект: Методология, применения, философия., М.:КРАСАНД, 2011
162. Фомин И. В., Семиотический фронтир: сквозь глубины веков и границы дисциплин // МЕТОД: Московский ежегодник трудов из обществоведческих дисциплин. 2017. №7.
163. Фуко М., Воля к истине: по ту сторону знания, власти и сексуальности., М.:Касталь, 1996
164. Хайдеггер М., Время и бытие., М.:Республика, 1993
165. Хайдеггер М., Основные проблемы феноменологии, СПб: Высшая религиозно-философская школа, 2001
166. Хайдеггер М., Бытие и время., М.:Академический проект, 2015
167. Хайдеггер М., Кант и проблема метафизики. Пер. с нем. / Перевод, послесловие О. В. Никифорова. М.: Издательство “ Логос” 1997.
168. Хайдеггер М., Кант и проблема метафизики. Пер. с нем. / Перевод, послесловие О. В. Никифорова. М.: Издательство “ Логос” 1997.
169. Хайдеггер М., Лекции по метафизики., М.: Издательский Дом ЯСК: Языки славянской культуры, 2016
170. Харари Ю.Н., Номо Deus, Краткая история будущего., М.:Синдбад, 2021

171. Хейес-Рот Ф., Уотерман Д., Ленат Д., Построение экспертных систем, М.: Мир, 1987
172. Харуэй Д., Манифест киборгов: наука, технология и социальный феминизм., М.:Ад Маргинем, 2017
173. Харман Г., Объектно-ориентированная онтология: новая «теория всего»., М.: Ад Маргинем Пресс, 2021
174. Хомский Н., Картезианская лингвистика. Глава из истории рациональной мысли. М.: КомКнига, 2005
175. Хуэй Ю., Рекурсивность и контингентность., М.: V-A-C Press, 2020
176. Хьюбнер К., Критика научного разума, М., 1994
177. Цофнас А.Ю. Структурная онтология, математика и системный подход / А.Ю. Цофнас // Философия математики: актуальные проблемы. – М.: МАКС Пресс, 2009.
178. Чалмерс Д., Сознательный Ум: В поисках фундаментальной теории., М.:УРСС: Книжный дом «ЛИБРОКОМ», 2015
179. Черкасов Л. Н., О понятии “структура системы” // Ярославский педагогический вестник. 1998. №3.
180. Шачин С.В. Коммуникативная теория разума Юргена Хабермаса: Опыт диалектической реконструкции теории коммуникативного действия. М.:ЛЕНАНД, 2019
181. Шапиро С.И., От алгоритмов – к суждениям (Эксперименты по обучению элементам математического мышления), М.: Сов. Радио. 1973
182. Шестова Евгения Александровна Проблема трансцендентального языка в VI Картезианской медитации О. Финка // Вестник РГГУ. Серия «Философия. Социология. Искусствоведение». 2015. №5 (148).
183. Шеффер Ж.-М., Конец человеческой Исключительности., М.:Новое литературное обозрение, 2010

184. Шиян А.А. Феноменологическое обоснование логики // Ученые записки Крымского федерального университета имени В. И. Вернадского. Философия. Политология. Культурология. 2015. №1.
185. Шумский С.А., машинный интеллект. Очерки по теории машинного обучения и искусственного интеллекта., М.:РИОР, 2019
186. Щур Г.С., Теории поля в лингвистике., М.: Книжный дом «ЛИБРОКОМ», 2009.
187. Энциклопедия философии и методологии науки, М.: «Кантон+» РООИ «Реабилитация», 2009
188. Эпштейн М., О будущем искусственного интеллекта. По поводу выступления Митио Каку - 2018 [Электронный ресурс] https://snob.ru/profile/27356/blog/138377#_ftn1 (Дата обращения 08.10.2021)
189. Ясперс К., Смысл и назначение истории. М.: Политиздат, 1991