

На правах рукописи

Нечаев Станислав Юрьевич

**Концепция абдуктивного мышления Ч.С. Пирса и
философия искусственного интеллекта**

09.00.03 История философии по философским наукам

Автореферат диссертации на соискание ученой степени
кандидата философских наук

Саратов – 2013

Работа выполнена в ФГБОУ ВПО «Саратовский государственный университет имени Н.Г. Чернышевского»

Научный руководитель доктор философских наук, профессор Беляев Евгений Иванович

Официальные оппоненты:

Гасилин Владимир Николаевич, доктор философских наук, профессор, Поволжский институт управления имени П.А. Столыпина – филиал РАНХиГС при Президенте РФ, профессор

Нестеров Александр Юрьевич, доктор философских наук, профессор, Самарский государственный аэрокосмический университет имени академика С.П. Королева, профессор

Ведущая организация Поволжский государственный технологический университет

Защита состоится «3» апреля 2013 г. в 14 часов на заседании диссертационного совета Д 212.243.09 при Саратовском государственном университете имени Н.Г. Чернышевского по адресу: 410012, г. Саратов, ул. Астраханская 83, XII корпус, ауд. 203.

С диссертацией можно ознакомиться в Зональной научной библиотеке Саратовского государственного университета имени Н.Г. Чернышевского.

Автореферат разослан «___» _____ 2013 года.

Ученый секретарь
диссертационного совета

Листвина Е.В.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования

На сегодняшний день в философской логике существует ряд сложных, пока неразрешимых проблем, касающихся различных типов вывода: в какой степени мы должны утверждать, что многочисленные формальные экспликации индукции и дедукции смогут воспроизвести единую теорию научного метода, или ответить на вопрос, как возможна содержательная логика. В этой суматохе обычно игнорируется идея об *абдукции*. Пирсовское прочтение контекста абдукции относит такой вывод к «расширяющим» (амплиативным – термин Ч.С. Пирса) аргументам, в отличие от дедуктивных экспликативных, – т.е. таким, какие позволяют получать *новое знание*, качественно дополняя дедуктивно-индуктивное познание, учитывая творческий элемент. Это говорит скорее о том, что абдукция, возможно, остается единственным общим критерием когнитивных способностей человека/«ментальности» искусственного интеллекта¹. Концепция абдуктивного мышления Пирса имплицитно соотносится его работами по математической логике, *научной метафизике* и философскими исследованиями, в развитии целостного подхода к проблеме, – в то время как его категориальное учение многими интерпретаторами опрометчиво исключается, хотя именно систему категорий Пирс напрямую соотносит с типами логического вывода, равно как и с общими принципами всей философии.

Историко-философский подход в диссертации движим прояснением философских принципов («прагматизма»), анализом логической аргументации и выявлением возможных следствий учений Ч.С. Пирса, в попытках применения полученных концептов в философии искусственного интеллекта, в стремлении репрезентативного ответа на то, как могут выглядеть возможности создания релевантных интеллектуальных систем. Трудности подобных изысканий, в особенности философских, очевидны – по-прежнему не прояснены основания для создания «сильного ИИ», а, следовательно, не решена и «трудная проблема сознания», ведь с философской точки зрения есть повод допускать, что эти две задачи тесно взаимосвязаны. Поэтому в работе проводится анализ «критериев ментального процесса» Г. Бейтсона и рассматривается «прогностическая теория интеллекта» Дж. Хокинса (активно дискутируемая в данном ключе), как примыкающие к концепции Пирса, что позволяет говорить о методологических пролегоменах к философии искусственного интеллекта в онтологических и логико-эпистемологических понятиях. Такие вводные замечания (и другие, которые, несомненно, также могут обсуждаться) предполагаются важным условием для дальнейших разработок.

Актуальность данной работы заключается в том, что предложено специальное историко-философское обоснование общей методологии ИИ, рекомендованы конкретные пути преодоления препятствий, продемонстрированы определенные перспективы. Сложности, какие могут возникнуть, сформируются уже в процессе последующих изысканий, но как призывал сам Ч.С. Пирс известной фразой против самореференции: *не преграждайте пути исследованию!*

¹ Далее принято предметное, техническое обозначение – ИИ (по аналогии с англ. artificial intelligence – AI).

Степень научной разработанности проблемы

Историко-философское исследование проще всего выстраивать в соответствии с правилами Р.Дж. Коллингвуда: 1) никогда не принимай на веру критику любого автора до тех пор, пока не убедишься, что эта критика действительно имеет к нему отношение; 2) никогда не считай, что ты понимаешь утверждение любого философа, пока ты не решил с максимально возможной точностью, на какой вопрос это утверждение должно было служить ответом.

Философские работы Ч.С. Пирса испытали заметное влияние спора об универсалиях реалистов и номиналистов, учений Дунса Скота и У. Оккама, *res cogitans* и *res extensa* Р. Декарта, одновременно и «коперниканского переворота» И. Канта, различных идей Д. Юма и представителей философии «здорового смысла», позитивизма Дж.Ст. Милля и эволюционной теории – в оригинальном виде подходя к решению многих из такого рода амбивалентных интенций в науке. Возникает некоторая сложность систематического концептуального понимания всего наследия Пирса, когда необходимо выстраивать эту систему вокруг опять-таки сложных понятий таких, как *непрерывность*, *привычка* или *фанерон*. Именно поэтому, сформулировав однажды прагматистские основы для философствования, Пирс в дальнейшем намеренно дистанцируется от большинства современников и последователей-приверженцев, ввиду их простоты-утилитарности признания того, как должен выглядеть прагматизм. После того, как У. Джемс ввел термин «прагматизм» в публичной лекции 1898г., с 1905г. в статьях Пирса появляется новое название для общих принципов философии: *прагматицизм* (логический прагматизм) – наименование, которое, по мнению Пирса, *звучит достаточно безобразно для тех, кто смог бы его присвоить*².

Многие биографы и исследователи философии Ч.С. Пирса (А. Беркс/А. Burks, Дж. Brent/J. Brent, Ю. Бюхлер/J. Buchler, К.Л. Кетнер/K.L. Ketner, Дж. Рэнделл/J. Ransdell, М. Фиш/М.Н. Fisch) курировали различные сборники сочинений Пирса, публиковали оригинальные статьи и монографии, в том числе в рамках Transactions of the Charles S. Peirce Society – ежеквартального журнала американской философии, выпускающегося с 1965г. по сей день. В отечественной научной литературе, кроме единственной биографической книги и работы того же автора (В.В. Кирющенко), посвященной рассмотрению природы языка и знака в контексте пирсовского учения о категориях, написана одна обширная монография о Пирсе (Ю.К. Мельвиль), где основательно изложены его важнейшие философские взгляды (в довольно критическом духе) и где практически не затронута логическое учение.

Вслед за Ч.С. Пирсом, в попытке сведения чистых математических понятий к понятиям логики, основываясь на постулатах булевой алгебры, и нахождения комплексного отношения математической логики и обыденного языка в принципе, к данным проблемам обращались и главные представители направления логицизма – Г. Фреге, Б. Рассел, Л. Витгенштейн и Ф. Рамсей; лидеры Венского и Берлинского кружков – Р. Карнап, К. Гедель и Г. Рейхенбах; логик А. Тарский и лингвист Н. Хомский – в попытках выстраивания формальных

² Далее принят термин «прагматицизм» как наиболее точный.

систем; проблемы широко обсуждались в контексте математического интуиционизма. Больше того, известный популяризатор математики М. Гарднер в своем безупречном стиле рассуждает о пирсовском фанероне – структурной частице феноменального мира.

Влияние идей Ч.С. Пирса ощутимо в работах К.И. Льюиса, автора концептуального прагматизма, предложившего одно из ключевых понятий современной аналитической философии сознания – *квалиа*; «аппроксимация» квалиативного сознания – актуальная тема для современных дискуссий. Историко-философский подход позволяет обнаруживать «интерферентные» и такого рода смешанные интерпретации многих из научных терминов, в своей общности формируя понимание «смутных» пирсовских понятий *тихизм*, *синехизм*, *агапизм*, именованные которых, в частности, навеяны греческой культурой. Как замечает Пирс, *автор научного понятия имеет право первым назвать его, и это название должно быть принято, если против него нет существенных возражений*. В современной философской мысли может потребоваться не только определение теоретических принципов демаркации таких сложных понятий как «разум», «сознание», «интеллект», но и философский статус самого понятия «мышление», или логическое определение того, что называется «принципом континуальности мышления». О роли Пирса в подобной «трансформации философии» рассказал уже К.-О. Апель. Не случаен и интерес к Пирсу со стороны ведущих философов 20в. – У.в.О. Куайна, Х. Патнема, Д. Дэвидсона – в той или иной степени, представителей так называемого неопрагматизма.

Считается, наиболее существенным вкладом Ч.С. Пирса в современную науку, эпистемологию и логику научного открытия, является тезис о погрешности научного знания в форме *фаллибилизма* – где знание есть лишь промежуточная интерпретация истины, и доминантой которого выступает контекст *абдукции*. Методологическое сходство теоретических положений Пирса со многими из недавних фундаментальных логико-эпистемологических картин мира, разрабатываемых философами науки – К.Р. Поппером, И. Лакатосом, Т. Куном, П. Фейерабендом, Н.Р. Хэнсоном, П. Липтоном – исходит из общего для них принципа индетерминизма. В том числе, в различных расширениях диспута и проблематики относительно фальсифицируемости эмпирических теорий и их верификации. Параллельное непротиворечивое развитие подобных идей можно найти в генетической эпистемологии Ж. Пиаже и в трудах одного из основателей этологии К. Лоренца. Некоторые позиции критического рационализма Поппера подвергнуты собственно критике уже в 70-х гг. прошлого века, за вычетом своих же учеников, в открытой полемике с другими оппонентами (Ю. Фрименом, Г. Сколимовским, Н. Решером, Я. Хинтиккой) именно в связи с «поворотом к Пирсу» с точки зрения того, как сможет выглядеть эволюционная эпистемология. К этим темам обращались в своих работах В.А. Лекторский, В.Н. Садовский, Д.Г. Лахути, Е.И. Беляев, С.Ф. Мартынович, В.Н. Гасилин.

Контекст абдукции называют исключительно сложной проблемой не только в эпистемологии (А. Алиседа/A. Aliseda, Дж. Нубиола/J. Nubiola, С. Паавола/S. Paavola, Дж. Фетцер/J.H. Fetzer), еще Ч.С. Пирс признал, что амплиативные типы вывода с трудом подвержены логической формализации.

Больше того, каждый из известных методов мышления – наряду с абдукцией (часто неверно принятая за ретродукцию), индукция и дедукция, традукция, или аналогия, редукция – одно из актуальных следствий учения о категориях *фанероско́пии* (У. Розензон/W.L. Rosensohn, Ч. Хартсхорн/C. Hartshorne, Дж. Эспозито/J.L. Esposito, С. Ика/S. Ika); теория знаков (формальная семиотика) является тем же следствием в классификации; равно как и все метафизические концепты Пирса (Р. Майорга/R.M.P.-T. Mayorga, Э. Рейнолдс/A. Reynolds) – и нормативные науки (этика, эстетика, и снова логика) – образуют *philosophia prima*. По его замыслу, три универсальные категории *Первость*, *Вторость*, *Третьесть*³ в качестве триадического со-отношения, приобретая характер универсалии, в смысле реализма схоластики, обоснованные математически (в принципе поддающиеся дескрипции на языке математики) – краеугольный камень прагматизма, или «эволюционного прагматического реализма».

Проблема абдукции подкрепляется доводами теории автоматизированных правдоподобных рассуждений Д.А. Бочвара, Н.А. Васильева, В.А. Смирнова, А.С. Есенина-Вольпина, В.К. Финна, О.М. Аншакова, А.С. Бобровой; категориальные отношения – в теории систем, математическом описании иерархических многоуровневых систем М. Месаровича. Абдукция становится предметом внимания *educology* (термин используется для обозначения специальной дисциплины, содержащей весь фонд знаний о процессе образования и методике преподавания). Разработчики курса (Дж. Маччия/G.S. Maccia, К. Томпсон/K.R. Thompson, Ф. Чиассон/P. Chiasson, Э. Штейнер/E. Steiner) занимаются методологией построения научной теории.

Историко-философский корпус дополняется работами «маргинальных» ученых – лидеров так называемой «ультракультуры»: Г. Бейтсон, Дж. Лилли, С. Гроф, Ф. Капра. Инспирированные концепцией абдуктивного мышления Ч.С. Пирса, многие идеи Бейтсона значительно расширяют контекст абдукции, позволяя дать законченное определение «ментальному процессу». Известно, насколько многогранны были его научные интересы: наблюдения за коммуникацией дельфинов, занятия культурной антропологией, поиски способов лечения шизофрении, кибернетическая теория алкоголизма, нейролингвистическое программирование – сыграли существенную роль в формировании общих взглядов Бейтсона на эпистемологию (М.К. Бейтсон/M.C. Bateson, Л. Бэйл/L.S. Vale, П. Хэррис-Джонс/P. Harries-Jones, Н. Чарльтон/N.G. Charlton). В отечественной традиции интересны работы В.В. Налимова, В.А. Фриауфа, Е.М. Иванова.

Социобиологическая парадигма философии сознания, вслед за *социолингвистической* и *биологицистской*, рассматривает ментальный процесс в двоякой перспективе искусственного и естественного. Дж.Р. Серл, Д. Деннет, Д. Чалмерс – ведущие специалисты по аналитической философии, затянутые в ин-

³ Далее приняты такие варианты перевода терминов *Firstness*, *Secondness*, *Thirdness*, т.к. «основная мысль Пирса при употреблении этих терминов не в том, что *Secondness* производна от *Firstness*, а *Thirdness* – от *Secondness* (что предполагается терминами Первичный, Вторичный, Третичный), а в том, что категория *Firstness* связана с числом 1, категория *Secondness* – с числом 2, а категория *Thirdness* – с числом 3 (имеется в виду количество объектов, предполагаемое каждым из этих терминов)» (Пирс Ч.С. Рассуждение и логика вещей: Лекции для Кембриджских конференций 1898 года / Науч. ред. Д.Г. Лахути и В.К. Финна. М.: РГГУ, 2005. С.348).

тенсивное обсуждение «трудной проблемы сознания» в контрверзе к проблеме «легкой», разрешение так называемых «аргументов знания» («комната Мэри», «Китайская комната», «философский Зомби» и проч.). Научная литература на русском языке примечательна спором Э.В. Ильенкова и Д.И. Дубровского о природе психики, исследованиями В.В. Васильева, Д.Б. Волкова, А.Ф. Грязнова, Д.В. Иванова, М.В. Лебедева, Н.В. Мальчуковой, А.Ю. Нестерова, Е.А. Никитиной, Н.С. Юлиной.

Проектирование ИИ начинается и развивается в работах А. Тьюринга, А. Ньюэлла, Г. Саймона, Дж. Маккарти, М. Минского, Х. Дрейфуса, Р.А. Брукса, Х. Моравека, Дж. Хокинса – с точки зрения различных технических сторон проблемы. Несмотря на то, что некоторые оптимисты продолжают верить в создание уже к 2020г. адекватного «сильного ИИ», способного решать трудные задачи уровня человека, в отличие от «слабого ИИ», – ни один из компьютеров пока не сумел пройти даже *теста Тьюринга*.

Объектом исследования выступает концепция абдуктивного мышления Ч.С. Пирса, иначе общие принципы прагматизма (логики абдукции).

Предметом исследования являются возможности применения абдукции и примыкающих концептов в философии искусственного интеллекта.

Основная цель исследования заключается в аналитическом рассмотрении оснований теории абдукции Ч.С. Пирса, в том числе через историко-философский анализ различных концептов его учений, и прежде всего, в системе категорий, повсеместного влияния философского наследия Пирса на проблемы в аналитической традиции (компаративистский анализ), во-первых; и, во-вторых, – в определении репрезентативных и релевантных перспектив создания ИИ (контрфактический анализ), через актуализацию абдукции в эпистемологии Г. Бейтсона, в теории интеллекта Дж. Хокинса, в выявлении методологических и философских оснований для ИИ в целом. Достижение поставленной цели определило логику решения следующих **задач**:

- проанализировать и представить авторскую интерпретацию учения о категориях Ч.С. Пирса
- исследовать историко-философские материалы и проблемное поле философии Ч.С. Пирса (архитектоники наук, фанероскопии, фаллибилизма, прагматизма, научной метафизики – тихизма, синехизма, агапизма)
- провести логико-философский анализ *теории абдукции* Ч.С. Пирса
- осуществить анализ «критериев ментального процесса» и выявить контекст абдукции в «рекурсивной эпистемологии» Г. Бейтсона
- определить методологические пролегомены к философии искусственного интеллекта (конфронтация «сильного»/«слабого» ИИ, «аргументы знания», шахматные программы); раскрыть контекст абдукции в «прогностической теории» Дж. Хокинса; представить авторскую концепцию репрезентации ИИ

Методологические основания исследования

Основным методологическим принципом работы избирается метод прогнозирования (моделирование, экстраполяция, предвидение), в купе с методом экспликации онтологических и гносеологических целей в форме нейтрального

(или аномального) монизма (Д. Дэвидсон), либо философской эклектики (столкновения идей и концепций), плюрализма, а также герменевтическим методом. В основания диссертации заложены принципиальные позиции холизма (У.в.О. Куайн, Х. Патнем, Н. Решер), нередуктивного физикализма, функционализма (в виде коннективизма), системного подхода (эмерджентизм), кибернетического и информационного подходов. Историко-философский подход отражает ретроспективный анализ различных интенций науки – в архитектонике наук Ч.С. Пирса этим обусловлена *philosophia ultima*.

Научная новизна диссертации заключается, прежде всего, в том, что концепция абдуктивного мышления Ч.С. Пирса рассматривается сразу в нескольких аспектах научного знания – в контексте философской логики, в сфере метафизического, в онтологическом и эпистемологическом контекстах, в логике научного исследования, философских и методологических основаниях ИИ. В соответствии с этим в диссертации получены следующие новые результаты:

- конкретизированы спорные моменты категориального учения (фанероскопии) Ч.С. Пирса; проведен компаративистский анализ фанероскопии Ч.С. Пирса в перспективе феноменологии Э. Гуссерля и отмечено, что Ч.С. Пирс не является феноменологом; категории рассмотрены в качестве рекурсивной иерархической системы; проведен концептуальный анализ фанерона (суммы категорий) Ч.С. Пирса/«теории 3-х миров» К.Р. Поппера и обнаружено, что «миры» Ч.С. Пирса – вариант лучшего объяснения
- уточнены основания теории абдукции в прагматизме и фаллибилизме, метафизических концептах Ч.С. Пирса; обсуждены точки соприкосновения концепции Ч.С. Пирса и взглядов других философов (Р. Декарта, Д. Юма, И. Канта, Г.В.Ф. Гегеля, Л. Витгенштейна, Н.Р. Хэнсона); установленный принцип континуальности мышления Ч.С. Пирса, представлен в связи с концептуальным принятием «аномального монизма» Д. Дэвидсона; отражена специфика терминологии (триадомания и синонимия) Ч.С. Пирса
- продемонстрировано проблемное поле *теории абдукции* в различных аспектах философского знания, ввиду ее амбивалентной трактовки самим Ч.С. Пирсом; представлены аргументы в пользу различения абдуктивной и ретродуктивной способностей интеллекта
- впервые соотнесены эпистемологические взгляды Ч.С. Пирса и Г. Бейтсона – «рекурсивная эпистемология» Г. Бейтсона раскрыта в контексте абдукции; представлен компактный анализ «критериев ментального процесса» Г. Бейтсона и их возможное расширение
- методологические и философские основания проблемы ИИ определены логико-эпистемологическими принципами науки; предложено обоснование теории интеллекта Дж. Хокинса через теорию абдукции и эпистемологию Г. Бейтсона; репрезентация ИИ дана по форме пирсовского фанерона в связи с авторской концепцией «другого интеллекта»

Положения, выносимые на защиту:

1. В фанероскопии Ч.С. Пирса первость соответствует абдукции, вторость – дедукции, третьность – индукции; каждая категория, в отношении к числу, за

исключением первости, является концептом множества (вторость содержит первость, третьесть содержит первость и вторость, третьесть репрезентируется в первость; в дедукции первость становится второстью через медиацию третьести, в индукции первость через вторость поднимается до третьести); первость есть универсалия (универсальность будет атрибутом вторости и третьести, поскольку категории не могут быть диссоциированы друг от друга). Категориальные отношения *непрерывно верифицируются* как рекурсивная иерархическая система, поэтому пирсовский фанерон (совокупность категорий) можно назвать *паттерном* (моделью, способом) общих отношений.

2. Прагматизм (или эволюционный прагматико-семиотический реализм) Ч.С. Пирса, объективируясь в качестве универсального принципа непрерывности в различных формах философской рефлексии (закреплении верования, привычке), логике научного открытия и тезисе о погрешимости любого знания (фаллибилизме), метафизических концептах (тихизме, синехизме, агапизме), есть не что иное, как *логика абдукции* (манифестируемая прагматистской максимой). Эта логика детерминирует континуальное мышление, где первость соответствует тихизму, вторость – синехизму, третьесть – прагматизму.

3. *Контекст абдукции* рассматривается в качестве доминанты фанерона. Фанероскопия проясняет абдуктивность и показывает, насколько рассуждениям свойственен смешанный характер, в общем виде – посредством трех элементарных типов вывода (абдукции, дедукции, индукции), где именно недедуктивные фрагменты (абдукция и, частично, индукция) создают возможности «креативного» мышления. Теория абдукции Ч.С. Пирса определяется тезисами *выводимости* (процесс вывода в эмпирическом и рациональном планах) и *автономности* (процесс получения нового знания), составляющими качественное различие абдукции и интуиции. Экстрагирована и допущена вероятность демаркации ретродуктивной и абдуктивной способностей интеллекта (ввиду амбивалентной трактовки Ч.С. Пирсом собственных терминов). Следствием абдуктивного рассуждения является *гипотеза*.

4. Анализ «критериев ментального процесса» Г. Бейтсона показывает, насколько его «рекурсивная эпистемология» задействует *контекст абдукции*. Предложенное расширение критериев, с одной стороны, является их уточнением (критерия 5 Бейтсона – «в ходе ментального процесса результаты воздействия различий следует рассматривать как трансформанты (т.е. закодированные версии) предшествующих событий»): *вся* действенная (эффективная) информация должна быть *сохранена*. С другой стороны, обеспечивает бейтсоновскую «экологию разума» онтологическим базисом концепции Ч.С. Пирса (критерия 6 Бейтсона – «описание и классификация этих процессов трансформация вскрывает иерархию логических типов, внутренне присущую феномену»): нарушения этой иерархии инициируются элементарными типами логического вывода (абдукцией, дедукцией, индукцией) в системе категорий (фанероном) Ч.С. Пирса.

5. Методологическими пролегоменами к философии ИИ предполагаются возможности реализации сложных процессов *познания, понимания и обучения* (в частности, используя концепции Ч.С. Пирса и Г. Бейтсона). «Прогностическая теория» ИИ Дж. Хокинса излагается перспективно, выстраивается через

логическую теорию абдукции и методологию, основанную на последних достижениях нейронаук и кибернетики. Репрезентацией ИИ предложена авторская интерпретация следствия пирсовской системы категорий в виде «другого интеллекта (another intelligence)», поскольку проектированием ИИ не предполагается моделирование «сознательной машины».

Теоретическое и практическое значение исследования состоит в расширении историко-философского материала по философии Ч.С. Пирса, эпистемологии Г. Бейтсона и возможном цельном изложении их сложных систем в виде монографии. С другой стороны, научная работа может быть продолжена в конкретных аспектах формализации абдуктивного мышления, уточнения методологических принципов философии искусственного интеллекта. Особое значение исследование имеет для дальнейшей разработки интеллектуальных систем и объяснения предикативных свойств ментального процесса.

Апробация работы

Результаты и выводы научной работы апробированы на заседаниях кафедры теоретической и социальной философии СГУ имени Н.Г. Чернышевского и аспирантском спецсеминаре в 2009-12 гг.

Основные положения диссертационного исследования представлены на всероссийских и международных научных конференциях, круглых столах: в VI Всероссийских Аскинских чтениях (Саратов, СГУ, 18 октября 2011 г.); в V Всероссийской конференции студентов, аспирантов и молодых ученых «Искусственный интеллект: философия, методология, инновации» (Москва, МГТУ МИРЭА, 9-11 ноября 2011 г.) – участие отмечено дипломом лауреата конференции; в 1-й Всероссийской конференции с международным участием «Аналитическая философия: проблемы и перспективы в России» (Санкт-Петербург, СПбГУ, 29-31 мая 2012 г.) с заочным участием; в VI Всероссийской конференции студентов, аспирантов и молодых ученых «Искусственный интеллект: философия, методология, инновации» (Москва, МГТУ МИРЭА, 29-30 ноября 2012 г.).

Структура работы

Диссертация состоит из введения, двух глав, по два параграфа в каждой, библиографического списка использованной литературы.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе «Принципы философии Ч.С. Пирса и теория абдукции» анализируется концепция абдуктивного мышления и выявляется универсальный принцип прагматизма Ч.С. Пирса.

В первом параграфе «Логическая аргументация Ч.С. Пирса» проводится логико-философский анализ теории абдукции Пирса через рассмотрение его учения о категориях (фанероскопии). Аргументация выстраивается Пирсом на следствиях математики и фанероскопии, но поскольку концепция абдуктивного мышления Пирса не может быть репрезентирована вне предмета нормативных наук (в понимании Пирса, прежде всего, – семиотики/логики) и научной метафизики, – философская мысль Пирса видится поразительно междисциплинарной. В связи с этим изложена манифестация трансдисциплинарности:

Концепция *трансдисциплинарности* показывает, что исследование выходит за рамки своих дисциплинарных пределов. Следует понимать, что эти дисциплины сформировались в ходе истории науки, и границы их применения не в объектах и не в теориях, но также исторические. В то же время их историческая идентичность формирует определенные объекты исследования, теории, методы и цели, которые часто не когерентны дисциплинарности по определению. Существует некая асимметрия между развитием проблемы и дисциплиной, занимающейся этой проблемой. Замысел трансдисциплинарности как новой формы холизма означает научный принцип, или научную ориентацию, в которой проблемы могут быть решены в полном объеме.

Историко-философский анализ классификации наук (архитектоники) Пирса показывает, как им определены задачи фанероскопии, логики и научной метафизики – через триадические отношения. Трудно сказать, что Пирс придерживался позитивистских взглядов, по простой причине того, что философия располагается сразу после математики. Несмотря на то, что Пирс изначально характеризует логику как «общую теорию знаков», впоследствии в ней выделены три самостоятельных раздела: *grammatica speculativa* (семиотика) изучает взаимоотношения «знаков»; *logic proper* (чистая логика) анализирует и синтезирует различные условия истинности (собственно логика); *rhetorica speculativa* (чистая риторика), задача которой состоит в установлении законов объективной логики. Метафизика постепенно приобретает характер логики в архитектонике наук Пирса, логика в большей степени преобразовывается в математику.

Через историю развития логики известно ставшее классическим бинарное отношение средневековой схоластики – *logica utens* (как некоторое общее представление о том, что есть наилучшее рассуждение; такая теория предшествует систематическому изучению предмета в качестве особого инстинктивного убеждения, как следует решать реальные проблемы) и *logica docens* (когда необходимо рассматривать нестандартные проблемы; т.е. научная, или критическая логика). Пересмотренное Пирсом, данное «парадоксальное отношение» можно свести к простым монотонным и контрарным немонотонным типам рассуждения, где первое (*logica docens*, «логика изучаемая») должно быть выводимо из второго (*logica utens*, «логика употребляемая»).

Фанероскопия (изначально феноменология) – фундаментальное учение о *Первости, Вторости, Третьести*, – по-видимому, исходит из признания того, что сознание, которое воспринимает феномены, есть само по себе *феномен*, или сумма всего «контента», содержащегося в сознании. Поэтому нет смысла редуцировать феноменальное сознание методами Э. Гуссерля, когда следует репрезентировать его посредством *паттерна* (модели, способа) общих категориальных отношений. В связи с тем, что логические запросы занимают значительные позиции в учениях Пирса и Гуссерля, возникает необходимость сравнения и противопоставления их представлений о логике. Компаративистский анализ устанавливает ключевые моменты *различия* и некоторые сходства (Г. Шпигельберг, К.-О.Апель), в связи с чем называть Пирса феноменологом *некорректно*.

Последовательно системную составляющую фанероскопии Пирса можно определить в основание формообразующего паттерна «*фанерон*». Три катего-

рии Пирса конституируют собственную комплексную структуру в *каждом* феномене. Поскольку, по Пирсу, *знак* (фанерон в узком смысле) в процессе (непрерывного) семиозиса переходит в другой знак, – каждая мысль сменяется другой мыслью. Иными словами, вторость содержит первость; третьность содержит первость/вторость; третьность *репрезентируется* в первость и так до бесконечности (или какого-то предела).

Первость (в своем монадическом аспекте) можно прочувствовать, но невозможно оформить, как если взять «мысль» саму по себе, – такая качественно определенная вещь выносится за границы познания, находится в возможности. Вторость есть «грубый» факт действительности, субъект-объектное отношение, коррелят. Третьность – закон и необходимость, репрезентамен. Вторость и третьность – концепты множества (в своем полиадическом аспекте).

Некоторые исследователи утверждают, что Пирс изначально определяет универсалией первую категорию, затем универсалией становится вторая категория (реализм из двух категорий), в конечном итоге, третья категория есть универсалия (реализм из трех категорий). Но это спорный вопрос, возникший, вероятно, из допущения того, каким значением Пирс наделяет третьность, (т.е. символичный знак, в отличие от иконического и индексного знаков), что верно, но совсем в другом контексте. Автором показано, что Пирс определяет универсалией только первость, но системная верификация категорий позволяет допускать, что универсальность является атрибутом вторости и третьности, поскольку категории не могут быть диссоциированы друг от друга.

Неизвестно, что в точности сыграло роль в создании категорий в том виде, в каком они показаны Пирсом, поскольку, по его осторожному замечанию, в принципе их трудно назвать понятиями, а значит, язык имеет к ним косвенное отношение. Математик скажет, что число само по себе ничего не обозначает. Масса примеров, приводимых Пирсом из истории философии и науки, обнаруживает триадические отношения не только в философии и логике, а в психологии, физиологии, биологии и физике. Возможно, предпосылкой явилось отношение категории к числу. Как предположено: именно числовые зависимости устанавливают, каким образом Пирс приходит к идее, что есть подлинная вторость и ее вырожденная форма, есть подлинная третьность и есть две ее вырожденные формы. И поскольку вторая категория соответствует *дедукции* – есть вырожденная форма дедукции; третья категория соответствует *индукции* – есть две вырожденные формы индукции; первая категория соответствует *абдукции*.

Посредством фанероскопии рассматривается прагматический критерий значения, и демонстрируется, чем обусловлены категории (Г. Райл), чем является реальность с точки зрения категорий, чем обусловлена связь категорий Пирса с типами логического вывода. Отмечено, что многие исследователи неправильно интерпретируют отношение индукции и дедукции к категориям, несмотря на последовательное соотнесение абдукции и первости. В итоге, дедукция соответствует вторости, индукция – третьности, но (т.к. вторость и третьность являются концептами множества) соотношение категорий с типами вывода нельзя переводить в простые связи вторости/дедукции, третьности/индукции: в

дедукции *первость* становится второстью через медиацию третьести; в индукции *первость* через вторость поднимается до третьести.

Таким образом, *контекст абдукции* рассматривается в качестве доминанты фанерона, абдуктивное мышление (спонтанность первости) – формой различия, абдукция предполагается амплиативным (расширяющим) типом логического вывода и самоочевидной когнитивной способностью. Если единственное (но довольно емкое) оправдание абдуктивного рассуждения заключается в том, что оно дает объяснение фактов, – можно снабдить теорию абдукции двумя тезисами: *выводимости* (процесс вывода в эмпирическом и рациональном планах) и *автономности* (процесс получения нового знания).

Далее интерпретируется формулировка абдукции, данная Пирсом:

наблюдение неожиданного факта *C*

если бы было верно *A*, то следствие *C* – естественно

∴ следовательно, есть основания предположить, что *A* – истинно.

Появление такого «неожиданного факта» есть, согласно Пирсу, единственная возможность для начала любого исследования – человек не может автономно и безотносительно к чему-либо заставлять себя углубляться в решение проблем, пока не встретит некоторые, до сих пор неизвестные факты или не окажется в состоянии получения нового опыта, когда абдуктивное рассуждение инициируется контекстом фактического. Вторая посылка предоставляет правдоподобное объяснение. Откуда это *A*? Объяснение, как правило, исходит из различения в *C* такого элемента, который до того не соответствовал другому элементу аналогичного (*C*) факта, воспринятого раньше, – в соответствии с правилами генерализации; и таким образом, единственная трудность заключена в определении того, как именно *A* оказывается тем самым необходимым истинным предположением (гипотезой) – по какой причине все положение дел состоит в принятии больше истинных гипотез, чем негодных, чем обусловлена тенденция прироста знания вообще.

Через постановку тезисов, логическое обоснование, примеры применения абдукции и обсуждение аспектов определения того, какой, по сути, должна быть гипотеза – по Пирсу, *объясняющей, проверяемой, экономичной*, – отмечается, что абдукцию не следует соотносить с интуицией. Фанероскопия проясняет абдуктивность и показывает, насколько рассуждениям свойственен смешанный характер, в общем виде – посредством трех элементарных типов вывода (абдукции, дедукции, индукции), где именно недедуктивные фрагменты (абдукция и, частично, индукция) создают возможности «креативного» мышления.

В качестве заключения параграфа проводится концептуальный анализ категориального учения Пирса и «теории 3-х миров», метафизической программы К.Р. Поппера, где установлено, что возможные «миры» Пирса содержат более сильные онтологические и эпистемологические основания.

Во втором параграфе «Континуальность мышления» демонстрируется как универсальный принцип непрерывности (принцип Ч.С. Пирса в прагматизме) актуализируется в различных формах философской рефлексии (закреплении верования, привычке), логике научного открытия (форме закона) и тезисе о погрешимости любого знания (фаллибилизме), метафизических концеп-

тах (тихизме, синехизме, агапизме), – по своей сути, через полиадические отношения категорий.

Излагается позиция по существенным философским вопросам (в которых Пирс выступает противником, прежде всего, Р. Декарта и И. Канта) и принимаются возражения относительно «четырех способностей, приписываемых человеку» в таких простых формулировках: (1) *у нас нет способности интуиции*; (2) *у нас нет способности интроспекции* (Пирс понимает под интроспекцией непосредственное восприятие внутреннего мира, но совсем не обязательно восприятие его *как* внутреннего); (3) *у нас нет концепции абсолютно непознаваемого*; (4) *у нас нет способности мыслить без помощи знаков* – в частности, через рассмотрение понятия *знака* (в форме непосредственной, динамической и окончательной интерпретанты) и проблематики лингвистической философии (Л. Витгенштейн, У.в.О. Куайн).

Поскольку эпистемология Пирса сосредоточена на проблеме «закрепления верования», выявляется следующая «диалектическая» зависимость эпистемологического процесса: *презумпция* → *мнение* → *привычка* → *подлинное сомнение* (или исследование) → *«привычка верить»* → *верование* → *презумпция* ←, где понятие «презумпции» выступает синонимом абдуктивности.

Далее анализируются пирсовские методы закрепления верования, каждый из которых обладает своими достоинствами: *метод упорства* в какой-то степени способствует формированию позиции индивидуума; *метод авторитета*, прежде всего, *пропорционально* более велик, чем метод индивидуума, но, как и последний, упускает диалектическую ступень сомнения, – и тот и другой, однако, возникают сдвигом *презумпции*; единственная и важнейшая отличительная ступень *априорного метода* создается состоянием подлинного сомнения; *метод науки* в своей структуре аппроксимирован прагматистской максимой (фаллибилизмом), которая служит правилом достижения высокой степени ясности идей. Метод науки, как только вера закреплена, подтверждает амплиативный статус презумпции, которая подвергается проверке при помощи индукции, в частных случаях, являясь предпосылкой какого-то другого мнения. В этом заключен основной смысл логического прагматизма по форме фаллибилизма, в котором любая идея изоморфна следующей, поскольку обе как таковые погрешимы. В этой связи для *максимы фаллибилизма* формулируются такие посылки: (1) истинное определение является своего рода инстинктивным притяжением реальных фактов; (2) знание и разум изоморфны; (3) научное знание (и не только научное) погрешимо. Доминантой фаллибилизма является контекст абдукции (Н.Р. Хэнсон, П. Липтон).

В конечном итоге, вне логики развития науки, рассматриваются пирсовские определения *качества* (простой идеи), *факта*, *закона* и, наконец, *привычки*. Данный анализ позволяет уточнить репрезентацию фанерона через другие понятия. Разобрано «опровержение» Пирсом известного аргумента Д. Юма «против чудес» и показано, почему Пирс определяет *универсалией* лишь *первость*, вводятся ограничения для понимания понятий «закона» и «привычки».

Для этого анализируются метафизические доктрины Пирса – *тихизм* (учение о случайности), *синехизм* (учение о непрерывности) и *агапизм* (эволю-

ционная космология). Основное внимание уделяется тому моменту, что синехизм тесно связан с пирсовской интерпретацией фаллибилизма. Фаллибилист обязуется принимать любые гипотезы, которые проясняют феномены. Таким же образом он принимает условия не только синехизма, но и тихизма. Тихизм ссылается на элемент беспричинной случайности, происходящей в мире. Тихистские события можно назвать отклонениями от законов природы. Аргумент гипотетического отбора в синехизме из-за своей приверженности фаллибилизму позволяет *не преграждать пути исследованию*. Необходимо выбирать гипотезу, объясняющую явление наиболее успешно, под страхом блокирования исследовательской возможности. Учение синехизма позволяет лучше объяснять многообразие законов природы, в отличие от детерминистских моделей.

Образование привычки, согласно Пирсу, есть индукция, которая затрагивает в уме различные уровни внимания или абстрагирования. Волевые акты происходят от ощущений, вызванных привычками, так же как инстинктивные действия происходят от врожденных форм в нашей природе. Привычка есть нахождение собственно закона (принципа непрерывности) в ментальной сфере. Но привычка свойственна и структурам, находящимся вне ментального. Видимо, Пирс полагает, что привычка служит подобием формирования какого-либо «постоянства» закона, и поскольку наш разум не объясняется такими видами законов в полной мере, привычка становится формой закона в ментальном. Таким же образом Пирс объясняет эволюционное развитие (и в частности, космологический агапизм). Континуальное мышление (как следствие рекурсивной системы категорий) детерминировано концептами научной метафизики Пирса.

В конце параграфа отмечается, что принцип континуальности мышления связан с концептуальным принятием аномального монизма Д. Дэвидсона, в решении исключать «трудную проблему сознания» из списка проблематики философии искусственного интеллекта. Поскольку абдуктивность есть абстракция (форма различия), одинаково представленная как в чувственном, так и в разумном, во всех субъектах и объективных законах, – абдукция, которая в системе Пирса соотносится с идеей первости, в качестве логического типа вывода применима в проектировании ИИ.

Во второй главе «Репрезентация искусственного интеллекта» анализируются «рекурсивная эпистемология» Г. Бейтсона и «прогностическая теория» Дж. Хокинса, вводятся методологические замечания к философии искусственного интеллекта – в связи с концепцией абдуктивного мышления.

В первом параграфе «Критерии ментального процесса Г. Бейтсона» представлено, как Бейтсон формулирует необходимые предпосылки для смещения устаревшей парадигмы понимания того, что можно назвать «интеллектуальным». Критерии выстраиваются в соответствии с принципами теории информации и коммуникации, кибернетики и общей теории систем, соприкасаются и с философским дискурсом. Анализ показывает, насколько «рекурсивная эпистемология» Бейтсона задействует контекст абдукции; критерии *могут* использоваться при создании разумных систем, предоставляя оригинальные пути решения эпистемологических проблем ИИ. Проводится анализ исчерпывающе-

го списка критериев, нацеленных на описание общих аспектов «феномена разумности» живых систем:

В соответствии с холистическим подходом, Бейтсон формулирует критерий 1 (*разум есть совокупность взаимодействующих частей, или компонентов*), в котором утверждает, что разум не возникает каким-то мистическим образом и не является неразделенной, даже неделимой/единой субстанцией; разум состоит из эмерджентных саморегулирующихся и самоорганизующихся частей, в приближении к которым мы не обнаруживаем ничего, что можно назвать «разумным», но, тем не менее, можем признать, насколько части находятся друг с другом в *отношениях*. С точки зрения Бейтсона, если мы намерены выстроить общую модель разума, мы должны быть в состоянии объяснить такую модель фундаментальными свойствами, в том числе ее неаддитивных элементов.

Эпистемология Бейтсона впервые встречается с контекстом абдукции в критерии 2 (*взаимодействие между частями разума инициируется различием*). Определяется, что Бейтсон предложил собственные версии обоснования абдукции: в частности, (1) термин «тавтология» понимается как совокупность связанных утверждений, правомерность *связей* между которыми не подлежит сомнению (от простого утверждения вида «если *S* истинно, то *S* истинно», чуть более сложного «если из *S* следует *Q*, то из *S* следует *Q*» до геометрии Евклида и формальной логики); (2) тавтология становится необходимым медиатором *описания* реальности и *объяснения* реальности (соотношения текста/контекста); (3) функция абдукции определяется нахождением *новых* релевантных феноменов, которые подходят под такие же правила вывода и могут быть отражены на ту же тавтологию. Если выносить тавтологию Бейтсона за пределы формальности (витгенштейновской тавтологии), (эмпирические) принципы тавтологии (здесь – аналогии), выстроенные на *естественном различии*, обусловлены амплиативными свойствами. Различие, таким образом, обосновывает предварительный статус одной из посылок в логической схеме абдуктивного вывода.

Анализируется проблема соотношения текста/контекста, исходя из положений Р.Дж. Коллингвуда (исторический контекст), А. Коржибски (общая семантика), К. Юнга (плерома и креатура), самого Бейтсона (культурная антропология, теория коммуникации), Пирса и Дж. Джастроу (экспериментальная психология). Анализ позволяет сделать вывод, что коммуникация не происходит «в вакууме», но содержится в *контексте* («веровании» Пирса), который образован презумпцией и вопрошанием («сомнением» Пирса). В соответствии с этим актуальная информация определяется Бейтсоном как «различие, которое производит различие», «действенное различие» (т.е. единица информации).

Критерий 3 (*ментальный процесс требует сопутствующей энергии*) является следствием критерия 2: т.к. взаимодействие по триггерному принципу инициируется различием, или информацией, энергия не является ключевой переменной. Метафора «энергии» вводит в заблуждение, когда понимается двусмысленно: как источник и как причина. В критерии 3 Бейтсон подчеркивает основной принцип теории систем, где ментальный процесс должен быть открытой системой по отношению к «энергии», иначе процесс неизменно завершится увеличением энтропии. Однако, уменьшение энтропии обусловлено не посту-

пающей энергией, а негэнтропией («отрицательной энтропией» Э. Шредингера). Критерий 3 косвенно подтверждает тезис Пирса о том, что у нас нет способности к интуиции и мы не обладаем непосредственным знанием, предваряющим опосредствованное – такие концепты делают систему замкнутой.

Критерий 4 (*ментальный процесс требует циркулярных – или еще более сложных – детерминирующих (каузальных) цепей*) постулируется центральным тезисом кибернетических систем в качестве рекурсивности. Разум определяется имманентным комбинированному взаимодействию дифференциальных частей системы, т.е. части системы связаны и взаимодействуют в рамках замкнутой цепи. Т.к. система рекурсивна, события могут влиять на всю систему в целом, либо изменять свою начальную точку. Ментальный процесс необходимо должен ссылаться на самого себя, создавая самореференцию. Это и есть прямое подтверждение принципа континуальности мышления.

Критерий 5 (*в ходе ментального процесса результаты воздействия различных следует рассматривать как трансформанты (т.е. закодированные версии) предшествующих событий*) признает, что обратная связь неизбежно кодируется в некотором роде. Это согласовано с тем, как познание включает в себе ментальное отображение внутреннего мира (иначе «фанерона» Пирса). Таким образом, критерий 5 постулирует структуру идеи или «смысла», как процесса, в котором информация кодируется в новую форму и только затем становится доступной другим уровням системной организации.

Ключом к пониманию критерия 6 (*описание и классификация этих процессов трансформации вскрывает иерархию логических типов, внутренне присущую феномену*) становится теория простых типов Рассела. Применяя теорию Рассела, Бейтсон использует обобщение, согласно которому объекты (в данном случае сообщения) различных логических уровней следует четко отделять друг от друга: изменения, происходящие на низших уровнях, почти никогда не затрагивают высшие; изменения, происходящие на высших уровнях, всегда вызывают целый комплекс изменений на низших уровнях. Согласно Бейтсону, нужно различать сообщение и мета-сообщения, текст/контекст и контекст контекстов – как следствия простой теории типов Рассела. Коротко говоря, аналоговая часть коммуникативного акта (вербальная) занимает мета-позицию по отношению к содержательной части (невербальной). Отмечается, что Бейтсон интерпретирует теорию Рассела как наглядный инструмент для разбора формальных паттернов общения (коммуникации) в контексте основного действия эпистемологии – создания различий. «Различие», в свою очередь, есть синоним абдукции.

Предполагается расширение критерия 5 в том, что ментальный процесс сохраняет *практически всю информацию*. На основании феномена гипертимезии, трансперсональной психологии (С. Грофа) и экспериментального доказательства японскими нейропсихологами (Я. Мияшита/У. Miyashita, М. Камеяма/М. Kameyama, И. Хасегава/И. Hasegawa, Т. Фукушима/Т. Fukushima) делается вывод, что на бессознательном уровне возможна «системная память» обо всех эффективных различиях (фактах).

Независимо от Бейтсона, к идее метапаттерна пришли У. Матурана и Ф. Варела в «теории Сантьяго» (аутопоэзиса); однако, первое существенное расхождение между концепциями встречается в критерии 5, и теория Бейтсона, по некоторому мнению, лишена онтологического базиса. Свои критерии Бейтсон не заканчивает, но останавливается, как и Пирс, в сфере логического. В частности, не ясно, каким образом разум «вскрывает» иерархию логических типов через «описание и классификацию процессов трансформации». Поскольку Бейтсон придает большое значение абдукции, предполагается, что критерий 6 Бейтсона можно снабдить онтологическим базисом фанероскопии Ч.С. Пирса – *метапаттерном* категориальных отношений (фанероном), как синтезом познавательных процедур (абдукции, дедукции, индукции, традукции (аналогии), редукции, «ретродукции»).

Во втором параграфе «Философия искусственного интеллекта: методологические пролегомены» определены такие особенности возможного создания ИИ, какие поддаются объяснению логико-эпистемологическими принципами науки. «Прогностическая теория» ИИ Дж. Хокинса излагается перспективно, выстраивается через логическую теорию абдукции и методологию, основанную на последних достижениях нейронаук и эпистемологии. Репрезентация ИИ предложена по форме пирсовского фанерона, в связи с авторской концепцией «другого интеллекта» (another intelligence (AI)).

Отмечается, насколько Пирс был у истоков создания автоматизированных систем. «Пирсовский» тезис 1 о том, что электрические цепи моделируют логику, можно считать первым методологическим «оправданием» создания «сильного ИИ», ввиду того, что, вероятно, необходим специальный метаязык для коммуникативности машины – таким «языком» может быть только логика, способная наделить компьютер общими правилами решения определенных задач. Для этого, впрочем, не обязательно решение психофизической проблемы в полной форме, а достаточно анализа того, как происходят сложные процессы *познания, понимания и обучения*.

Анализируется аргумент *Китайская комната* (Дж.Р. Серла), направленный против «сильного ИИ», и отмечается, что аргумент не в состоянии отвергнуть системного подхода. Как только бихевиоризм потерял свои позиции в психологии, *тест Тьюринга* перестал казаться надежным, поскольку отдельное «поведение» не будет следствием интеллектуальности (и это показано критериями Бейтсона). «Ответ от систем» исходит из тезиса о том, что понимание не верифицируемо. И каждая такая верификация упирается в семантический экстернализм, где значение понятия определено внешними факторами по отношению к потребителю понятия. Тезисом 2 утверждается, что тест Тьюринга обходит стороной главные особенности интеллекта – познавательные процедуры и обучение, поэтому не будет гарантом тестирования «сильного ИИ». В рамках концепции о приросте научного знания, критерий оценки машины Тьюринга должен исходить из того, какие новые идеи машина способна производить на основе того опыта, который ей получен.

Приводится пример шахматной программы, которая, используя такие методы, как перебор (или метод «грубой силы») и эвристический алгоритм (опе-

рирующий базами данных дебютов/эндшпилей), по большому счету, есть реализация человека из Китайской комнаты. Эти базы данных, т.е. (частичная) формализация того «языка», который находится в голове гроссмейстера, не формируют в машине способности «разговаривать» на этом языке, поскольку его формальная экспликация чужда ментальному процессу (не достаточна для него).

В связи с этим анализируется концепция «обучения», которую Бейтсон рассматривает как коммуникативный феномен и выделяет в нем несколько уровней: *нулевое обучение, обучение-I, обучение-II, обучение-III, обучение-IV*. Если потребуется установить, может ли машина в принципе чему-то научиться, возникнет необходимость определить какой из данных уровней она достигла.

Прогностическая теория Дж. Хокинса рассматривается как наиболее перспективная, поскольку может быть соотнесена с теорией абдукции. Ключевой недостаток современного состояния исследований ИИ, по Хокинсу, – отсутствие зоны, отвечающей за «понимание». На данный момент наука не располагает рабочей «теорией интеллекта». Анализ «иерархической темпоральной памяти» Хокинса показывает, каким образом реализуется такая «зона», а «кортикальные алгоритмы обучения» моделируют способность системы к «предсказанию».

Еще одну проблему шахматных программ составляют возможности ретроспективного анализа. «Движением назад» от отдельно взятой позиции анализируются возможные игровые комбинации, которые могли привести к исходной позиции. При помощи ретроспективного анализа придумываются нестандартные шахматные задачи, которые понятным образом отличны от стандартных тем же «движением назад». Идея, что ретроспективный анализ предполагает нахождение неочевидного логического решения задачи «n-шагов назад» (Р. Смаллиан), допускает предположение способности, схожей с абдуктивной, но в определенном смысле обратной ей, – ретродуктивной способности. К тому же такое предположение основано на том, как Пирс амбивалентно интерпретирует само понятие абдукции.

Герменевтический анализ термина «абдукция» экстрагирует и допускает вероятность демаркации ретродуктивной и абдуктивной способностей. В текстах Пирса это различие напрашивается, к тому же, идея обсуждается в последнее время (К. Томпсон/К.Р. Thompson, Ф. Чиассон/Р. Chiasson). Но многими, вслед за отождествлением понятий в первом Гарвардском собрании сочинений Пирса, между терминами ставится равенство. Абдукцию и ретродукцию сложно назвать синонимами, так же как дедукцию и индукцию, к тому же у латинского *abductio* – нет субъекта. Есть вероятность, что редакторы решили вопрос лексически, тогда как семантически у Пирса абдукция и ретродукция – различны. Можно утверждать, что абдукция, в том смысле, который ей принадлежит, – придумана Пирсом. По аналогии с ретроспективным анализом, ретродукция, возможно, имеет дело исключительно с нестандартным анализом готовых гипотез, тогда как абдукция на основе получаемого опыта предполагает выдвижение новых гипотез, объясняющих его. Таким образом, гипотеза есть прямое следствие абдуктивного процесса.

Предлагается репрезентация «другого интеллекта» – бессознательной машины, способной решать сложные задачи уровня человека. Пока трудно сказать будет ли «сильный ИИ» обладать сознанием, или сознание сможет возникнуть в качестве «системного эффекта», возможно в результате «ошибки» программирования. По крайней мере, на данном этапе необходима работа над посылами способности «быть интеллектуальной», а не создания уникальной личности машины.

В заключении диссертационной работы излагаются итоги выполненного исследования, рекомендации по структурному отображению диссертации, анализ выводов и положений, перспективы дальнейшей разработки темы.

Основные положения диссертации отражены в следующих научных работах автора (в объеме 1,7 п.л.):

Публикации в изданиях, входящих в Перечень российских рецензируемых научных журналов, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней доктора и кандидата наук:

1. Нечаев, С.Ю. Китайская комната Дж.Р. Серля в контексте проблем философии искусственного интеллекта / С.Ю. Нечаев // Известия Саратовского университета. Новая серия. Серия Философия. Психология. Педагогика. – 2010. – Т.10, вып.4. – С.19-23. 0,6 п.л.

2. Нечаев, С.Ю. Фанероскопия Ч.С. Пирса versus феноменология Э. Гуссерля / С.Ю. Нечаев // Известия Саратовского университета. Новая серия. Серия Философия. Психология. Педагогика. – 2012. – Т.12, вып.1. – С.25-28. 0,5 п.л.

Публикации в других изданиях:

3. Нечаев, С.Ю. Стохастические системы и критерии ментального процесса Г. Бейтсона / С.Ю. Нечаев // Искусственный интеллект: философия, методология, инновации / Под ред. Д.И. Дубровского и Е.А. Никитиной. – М.: Радио и связь, 2011. – С.54-55. 0,1 п.л.

4. Нечаев, С.Ю. Трансдисциплинарность и постнеклассический идеал рациональности / С.Ю. Нечаев // Ценности, риски, коммуникации в изменяющемся мире: Сборник научных трудов. – Саратов: Изд-во «КУБиК», 2012. – С.317-320. 0,2 п.л.

5. Нечаев, С.Ю. Континуальность мышления – аномальный монизм Д. Дэвидсона / С.Ю. Нечаев // Аналитическая философия: проблемы и перспективы развития в России. – СПб.: Изд-во. СПбГУ, 2012. 0,1 п.л.

6. Нечаев, С.Ю. Джефф Хокинс об интеллекте / С.Ю. Нечаев // Искусственный интеллект: философия, методология, инновации / Под ред. Д.И. Дубровского и Е.А. Никитиной. Часть I. – М.: МГТУ МИРЭА, 2012. – С.27-29. 0,2 п.л.