

*На правах рукописи*

**БАРЫШНИКОВ Павел Николаевич**

**МЕТОДОЛОГИЧЕСКИЕ ВОЗМОЖНОСТИ И ГРА-  
НИЦЫ ВЫЧИСЛИТЕЛЬНЫХ МОДЕЛЕЙ СОЗНАНИЯ**

Специальность 09.00.01 – онтология и теория познания

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени  
доктора философских наук

**Москва – 2018**

Работа выполнена на кафедре исторических социально-философских дисциплин востоковедения и теологии Федерального государственного бюджетного образовательного учреждения высшего образования «Пятигорский государственный университет».

**Научный консультант** доктор философских наук, профессор, профессор кафедры онтологии и теории познания факультета гуманитарных и социальных наук ФГАОУ ВО «Российский университет дружбы народов»  
**Найдыш Вячеслав Михайлович**

**Официальные оппоненты** доктор философских наук, профессор, заведующий кафедрой, кафедра социальной философии философского факультета, ФГАОУ ВО «Новосибирский национальный исследовательский государственный университет»  
**Целищев Виталий Валентинович**

доктор философских наук, доцент, профессор, кафедра философии, социологии и политологии института инновационных технологий и государственного управления, ФГБОУ ВО «МИРЭА – Российский технологический университет»  
**Никитина Елена Александровна**

доктор философских наук, доцент, заведующий кафедрой, кафедра философии, ФГАОУ ВО «Самарский национальный исследовательский университет имени академика С.П. Королева»  
**Нестеров Александр Юрьевич**

**Ведущая организация** **ФГБУН Институт философии Российской академии наук (Институт философии РАН)**

Защита состоится «26» декабря 2018 года в «\_\_\_» часов на заседании диссертационного совета Д 212.203.02 при Федеральном государственном автономном образовательном учреждении высшего образования «Российский университет дружбы народов»: 117198, г. Москва, ул. Миклухо-Маклая, д. 10/2, ауд. 415.

С диссертацией можно ознакомиться в Учебно-научном информационном библиотечном центре РУДН (УНИБЦ РУДН) и на сайте РУДН <http://dissovet.rudn.ru>

Автореферат разослан «\_\_\_» \_\_\_\_\_ 2018 года.

**Ученый секретарь**  
диссертационного совета Д 212.203.02

**О.Ю.Бондарь**

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность исследования.** Второе десятилетие XXI века характеризуется технологическими прорывами, формирующими предпосылки для зарождения шестого технологического уклада, среди ключевых направлений которого особо выделяются нанотехнологии, биотехнологии, информационные технологии, когнитивные науки, социогуманитарные технологии, системы искусственного интеллекта и др. Становление шестого технологического уклада по сути является технологической революцией, характерной особенностью которой является нацеленность на совершенствование природно-биологических и социально-культурных черт способностей человека как главной производительной силы общества. В этих условиях искусственный интеллект из наивно-романтического образа, сложившегося в середине XX в. в связи с осмыслением перспектив развития таких зарождавшихся тогда научных направлений как кибернетика, информатика, когнитивные науки, постепенно приобретает черты реалистической научно-технологической программы.

Информационно-компьютерная реализация интеллектуальных и познавательных процессов, становится ключевой задачей научно-технологического развития, решение которой даст человечеству эффективные инструменты постижения природной действительности и управления социальной реальностью, формирования новых социально-экономических, политических и гуманитарных практик. На этом пути возникает и целый комплекс вопросов философского уровня, связанных с проблемой сознания, концептуальных систем компьютерных наук и информационных технологий, в том числе вычислительных моделей интеллектуальных и когнитивных функций, интерпретацией инженерно-технических подходов в реконструкции онтологических и эпистемологических свойств разумной деятельности.

Проблема методологических оснований вычислительных моделей сознания в современной философии когнитивных наук является приоритетной

и обладает особым статусом и значимостью. Это связано, с одной стороны, с тем, что природа сознания и его онтологические свойства в полной мере не определены, с другой стороны, с тем, что революционные прорывы компьютерных подходов в современных когнитивных науках позволяют говорить о вычислимых свойствах разума, сводя сложнейшие компоненты ощущений, восприятия, репрезентаций, смыслопорождения к глубинным алгоритмическим функциям ментальных процессов.

Историческая связь вычислительных теорий с философией сознания обладает сложным междисциплинарным характером. Наиболее ярким примером сближения компьютерных моделей с теоретическим предметом аналитической философии сознания является трансформация репрезентационализма в так называемую Вычислительную теорию сознания (Computational Theory of Mind – СТМ) с последующим превращением в информационно-функциональную парадигму когнитивных наук.

Согласно Вычислительной теории сознания, все когнитивные процессы представлены в виде сложных вычислительных систем: вера, мышление, эмоции, мотивы, желания – это различные виды информации, которые обрабатываются агентом для достижения некоторых целей. В основе отношений человека и реальности лежат не цифровые вычисления, а своеобразные формы механической рациональности: каждое сознательное действие агента имеет под собой ментальную причину, вызванную, в свою очередь, алгоритмической (эволюционной) обработкой статистических данных для достижения адаптивных целей.

При такой постановке вопроса в задачи сторонников вычислительного направления входит расшифровка и анализ глубинного «программного кода» деятельности сознания, поверхностный уровень которого реализуется через структуры репрезентационального содержания.

Информационные свойства сознания, согласно репрезентационализму, состоят из интенциональных состояний, представленных в особых символических формах. В данном ракурсе ментальный язык представляется как

разновидность вероятностного автомата, работающего по предзаданным алгоритмам. Усиливают эту позицию данные эволюционной психологии, которые сводились к пониманию человеческого сознания как вычислительного устройства, унаследованного от биологических предков и предназначенного для адаптивных функций организма в физическом и социальных мирах. Однако вычислительные модели сознания (как философские, так и прикладные) не лишены методологических противоречий, которые заслуживают отдельного концептуального анализа.

В современных теоретических направлениях, таких как философия сознания, философия когнитивных наук, философия искусственного интеллекта, существует ряд фундаментальных проблем:

1) не ясны способы перехода физического в ментальное на уровне когнитивных процессов;

2) не определена онтологическая природа символической обработки информации когнитивными механизмами;

3) существуют противоречия в вопросах локализации когнитивных процессов на уровне нейрофизиологии;

4) несмотря на популярность информационных моделей, не доказаны вычислительные свойства индивидуального феноменального опыта концептуализации и ментальных состояний;

5) не описаны методологические границы вычислительной парадигмы в контексте проблем феноменального сознания.

В Вычислительной теории сознания особое место занимают исследования семантической концептуализации. Проблема онтологического статуса семантических процессов на сегодняшний день остается ключевой. Особую роль в вычислительных моделях и методологии искусственного интеллекта занимает проблема содержания функциональных состояний системы, которая не до конца изучена: имеются значительные противоречия между теориями, гипотезами и выводами.

Данное исследование основывается на положении о том, что граница

между информацией и значением, телесным и ментальным, вычислимым и невычислимым проходит на уровне особых свойств семантической информации. В таком ракурсе проблема сознания, насколько нам известно, не исследовалась.

Актуальность описанных проблем обеспечивается также необходимостью рассмотрения с иных позиций фундаментальных вопросов о природе человека и способах познания мира, философского анализа этапов формирования компьютерных метафор «мозг-компьютер», «разум-программа», в рамках которых раскрываются исторические и методологические предпосылки зарождения радикального компьютеризационизма.

Результаты философского решения проблемы вычислимости ментальных процессов могут применяться как в фундаментальных теоретических областях, так и в прикладных системах по компьютерному моделированию интеллектуальных и когнитивных процессов.

### **Степень разработанности проблемы**

Проблема сознания в философии относится к разряду вечных проблем. Отказ от картезианского дуализма со стороны британской философии заложил основы эмпирических исследований ментальных процессов. Можно утверждать, что антикартезианский поворот, осуществленный Д. Юмом в исследовании сознания, нашел свое продолжение в аналитической философской традиции и в эмпирических науках о сознании. Главная новация этого поворота – отказ от субстанциализма и первая попытка представить сознание вне традиционных философских категорий. Тем самым обозначается связь между традиционной философией и современными научными исследованиями сознания.

Идеи философского конструктивизма кантианского толка получили новую интерпретацию в современных когнитивных науках. На первый исследовательский план выходит механизм преобразования физических сигналов в ментальную значимую информацию. В ходе когнитивного поворота объектом эмпирического познания становятся преобразующие ментальные

механизмы. Реалистские установки науки вкупе с конструктивистскими основаниями когнитивизма послужили методологическими основаниями для зарождения вычислительных моделей сознания.

Проблема вычислительных моделей сознания стала особенно остро обсуждаться во второй половине XX в., когда благодаря открытиям в кибернетике и когнитивной психологии в гуманитарные науки стали проникать системно-информационные методы исследований. Традиционные идеи находят свое продолжение в исследованиях по современной отечественной когнитивной психологии в трудах Б.М. Величковского, М.Ф. Фаликман, Ж.И. Резниковой, В.Д. Соловьева, А.Ю. Агафонова, Д.В. Ушакова, А.Д. Кошелева, Э.А. Костандова. В российской философской традиции первые успешные попытки сформировать антиредукционистское направление в рамках философии сознания и философии когнитивных наук реализуются в работах А.Ю. Алексева, И.А. Бесковой, В.В. Васильева, Д.Б. Волкова, Д.И. Дубровского, Д.В. Иванова, В.А. Лекторского, Н.С. Юлиной, Т.В. Черниговской, К.В. Анохина, В.М. Найдыша и др.

Получающая в последние десятилетия все большее признание в отечественной философии аналитическая традиция исследования сознания находит свое отражение в трудах А.Ф. Грязнова, В.В. Целищева, В.А. Суровцева, В.А. Ладова, Д.В. Винника, С.Ф. Нагумановой, М.В. Лебедева, А.З. Черняка, С.В. Никоненко, И.Г. Гаспарова, Д.Н. Разеева, М.А. Секацкой, Д.Э. Гаспарян, С.М. Левина. Философские и лингвофилософские аспекты искусственного интеллекта раскрыты в работах российских исследователей О.П. Кузнецова, А.А. Жданова, В.В. Лазарева, Г.Н. Рапопорта, Г.А. Герца, А.В. Напалкова, Л.Л. Прагиной, В.В. Петрова, В.Г. Редько, И.Ю. Алексеевой, Е.А. Никитиной, Н.А. Ястреб.

За рубежом в XX в. вычислительные модели формировались в основном в области аналитической философии, философии сознания и философии искусственного интеллекта – направлений, развившихся из различных вариаций логического позитивизма и бихевиоризма. К предтечам компьютерно-

лизм в аналитической философии сознания с определённой долей условности можно отнести Дж. Райла, Д. Льюиса, Л. Витгенштейна, Б. Рассела, Дж. Остина, У. Куайна, Г. Фреге, Р. Карнапа и др.

Наиболее влиятельными представителями различных направлений вычислительной философии сознания и философии искусственного интеллекта по сей день являются американские и британские исследователи Б. Баарс, Р. Джекендофф, Д. Дэннет, Н. Блок, Дж. Фодор, З. Пилишин, П. Черчлэнд, Р. Рорти, Х. Патнэм, Д. Хофштадтер, Д. Дэвидсон, Дж. Серл, К. МакГинн, Т. Нагель, Н. Винер, С. Крипке, Р. Пенроуз, О. Фланаган, Д. Чалмерс, Дж. Лукас, Ф. Дретске, С. Хорст, Дж. Пичинини, М. Рескорла, Л. Флориди и мн. др. На западе направления Philosophy of Mind и Philosophy of AI & Cognitive Science имеют самый высокий рейтинг среди междисциплинарных исследовательских областей.

На сегодняшний день проблемой семантических процессов сознания в междисциплинарном аспекте в Российской Федерации занимается сразу несколько ведущих научных школ: Межрегиональная ассоциация когнитивных исследований (рук. Т.В. Черниговская), Центр когнитивистики РГГУ (рук. В.И. Заботкина), отдел нейронаук НИЦ «Курчатовский институт» (рук. К.В. Анохин), Российская ассоциация лингвистов-когнитологов (пред. В.А. Виноградов, Н.Н. Болдырев), Московский центр исследования сознания при МГУ им. Ломоносова (рук. В.В. Васильев), Совет РАН по методологии искусственного интеллекта (пред. В.Л. Макаров, Д.И. Дубровский).

За рубежом существует большое количество научных центров, исследующих языковые процессы сознания с междисциплинарных позиций: The Institute of Philosophy of Mind and Cognition at National Yang Ming University, Association for the Scientific Study of Consciousness, The Society for Philosophy and Psychology, Consciousness Studies at University of Arizona, MIT Department of Brain and Cognitive Sciences (Department at the Massachusetts Institute of Technology), Center for the Study of Language and Information (Stanford University, SRI International, and Xerox PARC).



Даже при неполном списке лабораторий и научных центров, исследующих проблемы сознания, очевидно, что вычислительная методология является ведущей и представляется как одно из наиболее актуальных современных направлений, способных достичь новых результатов относительно процессов познания, понимания, коммуникации.

Особое место в западной традиции занимает малоизученная в России Вычислительная теория сознания (Computational Theory of Mind – СТМ). Данное направление зародилось в конце XX века на волне роста когнитивных исследований и популярности метафоры «Мозг-Компьютер». Основная задача данного направления состоит в выявлении вычислимых аспектов человеческого мышления и познавательных процессов.

С помощью СТМ были сформулированы важные для понимания работы сознания теоретические идеи:

- Вычисления могут быть определены через конечное число символов и правил, комбинирующих порядок этих символов.
- Вычисления могут быть приведены к алгоритму и пошаговым инструкциям, доступным для машинного исполнения.
- Вычисления могут быть обобщены логико-математическими методами.
- Вычислительные архитектуры позволяют моделировать сложные когнитивные процессы, минуя феноменальные компоненты отношений организма и среды.

Существует критика данного направления, основные тезисы которой сводятся к следующему:

- Свойства семантики не всегда вытекают из свойств синтаксиса.
- Интенциональные состояния сознания с трудом поддаются моделированию из-за многообразия символических выражений, «привязанных» к одному имплицатурному содержанию.
- При всех успехах современной нейрофизиологии остаются неясными способы кодировки мозгом ментальных состояний.

- Существуют методологические ограничения формальных систем и теории множеств, которые остро ставят вопрос о невычислимости когнитивных процессов, для которых необходимо понимание системой гёделевского предложения.
- Известны также антимеханистические и антиредукционистские аргументы, основанные на различных положениях теории информации, теории вычислимости, теории квантового поля.
- В инженерном плане не решена проблема минимального функционального коррелята сознания.

Лингвистическим ответвлением СТМ стала гипотеза «Языка мысли» Дж. Фодора (Language of Thought – LOT), которая на сегодняшний день представляется весьма авторитетной, но не лишенной изъянов. По мнению некоторых специалистов (например, С. Шнайдер), LOT может стать основанием для прагматического атомизма, открывая новые перспективы взаимодействия философии сознания с нейронауками. Основная идея синтеза LOT и СТМ сводится к тому, что теория символов нуждается в кооперации с вычислительными моделями нейронаук, для того чтобы приобрести черты строгой натуралистической теории и избежать накопившихся философских противоречий.

**Объектом исследования** является проблема сознания, представленная в вычислительных моделях, сформированных на различных этапах становления компьютеризационистской парадигмы.

**Предметом исследования** является семантика ментальных процессов в контексте вычислительной теории сознания и ее междисциплинарных разновидностей.

**Проблема исследования** состоит в прояснении объяснительных возможностей вычислительных моделей сознания как для инженерной реализации, так и для философского онтологического измерения. Сложность проблемы состоит в том, что онтологически противоречивый вычислительный подход к сознанию продолжает использоваться в прикладных когнитивных

исследованиях как эвристическая метафора, с помощью которой достигаются весомые результаты в компьютерной инженерии, теории искусственного интеллекта, современных нейробиологических исследованиях. Физические корреляты сознания в языке науки описываются с помощью компьютерной терминологии и концептуального аппарата теории информации. При этом не учитывается, что научное понятие сознания фокусирует в себе все многообразие онтологических и эпистемологических проблем, связанных сознанием как философской категорией. Вопрос о соотношении значений термина «сознание» как эмерджентного свойства материи мозга, как вычислительных функций информационных процессов, как носителя феноменальных состояний, как генератора абстрактно-символьного языка описания является ядром проблемы.

### **Цель исследования**

*Цель* исследования состоит в обосновании методологических границ вычислительных подходов к онтологии сознания одновременно с обоснованием методологических перспектив компьютерно-параллелизма в вычислительных инженерных моделях когнитивных процессов.

### **Задачи исследования**

В соответствии с поставленной целью в данном диссертационном исследовании решается следующий ряд *задач*:

1. Выявить общие основания в научных и философских программах исследования сознания с указанием исторических предпосылок для зарождения эмпирических методов в изучении ментальных процессов.
2. Проанализировать этапы становления вычислительной парадигмы в современных теориях сознания с указанием зависимости философских представлений от роста научно-технического знания в области компьютерных наук.
3. Раскрыть эвристические функции компьютерной метафоры в

становлении информационно-системной методологии когнитивных наук.

4. Провести концептуальный анализ методологических оснований критики радикального компьютеризационизма, в результате которого построить новую классификацию антивычислительных философских подходов.
5. Обосновать несовместимость онтологических свойств информации с формальными аспектами проблемы машинной вычислимости.
6. Объяснить через связь теории концепта с вычислительной теорией сознания специфику семантических моделей ментальных процессов.
7. Раскрыть противоположности в семантических принципах представления знаний в искусственных и биологических интеллектуальных системах на основании положений информационной теории репрезентации и телесно-ориентированного подхода.

### **Теоретико-методологические основания исследования**

Теоретико-методологические основания диссертационного исследования обусловлены несколькими фундаментальными исследовательскими программами: философской вычислительной теорией сознания, компьютерными моделями ментальных процессов в когнитивных науках и концептуально-семантической теорией ментального содержания. Под концептуальным анализом в данной работе понимается аналитическая процедура прояснения значения выражений, лежащих в основании той или иной теории, а также раскрытие содержания языковых структур теории. При формулировании положений критики радикального компьютеризационизма используется аксиоматико-дедуктивный метод, в базовых положениях которого лежит не-

сводимость процедур вывода ментального содержания к алгоритмической вычислимости.

Для обоснования указанных положений используется феноменологический метод, включающий последовательное дескриптивное представление ментальных процессов. Метод сравнительного анализа используется при сопоставлении различных этапов становления как вычислительных, так и антивычислительных подходов. При исследовании логической представимости инженерной реализации перспективы от первого лица используется метод мысленного эксперимента.

В методологических основаниях данного исследования лежат следующие посылки и теоретические принципы:

1. Принцип имитационного моделирования ментальных процессов в рамках функционалистских аналогий.
2. Принцип каузальной значимости содержательных компонентов ментальных состояний.
3. Принцип репрезентативной релевантности, необходимый для реализации функционального значения.
4. Положение о наличии концептуальных связей, возбуждаемых предметной соотнесенностью языкового выражения.
5. Натуралистическая установка: семантическая информация, обладая особым онтологическим статусом, является результатом суммирования физических процессов.

Представленные положения позволяют построить критическую программу вычислительных моделей сознания и вместе с тем обосновать значимость семантического измерения ментальных процессов и возможность реализации принципа семантической каузальности в сложных информационных процессах биологических физических систем.

**Новизна диссертационного исследования**

Новизна данного исследования состоит в обосновании новых аналитических решений, способных систематизировать и сопоставить данные когнитивных наук (нейробиология, когнитивная психология, когнитивная лингвистика, искусственный интеллект, теория информации) с аналитической философией сознания и представить их в системном ключе. Комплексное философское исследование вычислимости ментальных процессов является перспективным и способно принести новые теоретические данные для решения проблемы методологической обоснованности компьютеризации и проблемы интеграции западных информационных подходов с отечественной традицией когнитивных наук.

**Новизна диссертационного исследования** заключается в следующем:

1. Обоснована методологическая взаимосвязь между компьютерным поворотом в когнитивных науках и эволюцией антикартезианских философских учений о сознании. Описаны историко-методологические предпосылки кризиса когнитивного конструктивизма. Пересечение понятий информации и значения, приводящее к появлению противоречий в рамках эмпирических объяснений феномена сознания, представлено в качестве «слабого звена» в этапах развития компьютерной метафоры.
2. Впервые раскрыты основания проблемы символа в Классической вычислительной теории сознания, суть которой выражается в противоречивых свойствах символьных операций; в результате становится невозможным выявление точных критериев соотношения операциональных абстракций с частными физическими состояниями мозга.
3. Доказано, что коннекционистские модели как альтернатива последовательным вычислениям укрепили методологические позиции компьютерной метафоры, при этом исследовательский

фокус сместился с проблемы ментального содержания в сторону архитектуры мозговых связей.

4. Показаны противоречия компьютерного функционализма, в котором не дифференцированы с методологических позиций элементы теории данных и элементы теории информации физических систем. Установлено, что семантика данных при любом уровне вычислительной сложности редуцируется к реализации численной меры (информации) в физических состояниях аппаратных узлов.
5. Выявлены недостатки машинной имитации интеллектуальных процессов: критерий когнитивно-коммуникативной релевантности машинного имитационного интеллекта остаётся размытым и зависит от философских трактовок сознания и выбора минимальных условий разумности, в которых, как правило, не находится места феноменальным аспектам.
6. Аргументирована значимость семантического ракурса в исследованиях ментальных процессов для современного этапа развития компьютерационалистской парадигмы в когнитивных науках и философских теориях сознания.
7. Обнаружены новые объяснительные возможности когнитивно-семантической теории концепта, которые применимы для реабилитации положений компьютерационализма в философских теориях сознания.
8. Обосновано, что феноменологическая онтология сознания требует от информационного подхода отказаться от принципов вычислимости. В этом случае вместо «прямого» компьютерационально-функционалистского решения требуется акцент на проблеме ментальной причинности.

## **Положения, выносимые на защиту:**

1. Вычислительные модели сознания, разработанные в рамках философских теорий и прикладных когнитивных исследований, были сформированы в результате развития неоднородных и конкурирующих идей, вследствие чего компьютерационалистский подход к проблеме сознания обладает рядом противоречий. Одновременное соблюдение вычислительной достаточности, эмпирической адекватности и физической реализуемости становится невозможным ни в одной из моделей вычислимости. Вычислительные абстракции меняют свои характеристики в зависимости от принципов компьютерной реализации; таким образом, в вопросе о сознании понятие вычисления используется в нестрогом значении.
2. Компьютерная метафора, в основании которой лежит аналогия между вычислительными архитектурами мозга и компьютерными системами, обладающими схожими функциональными компонентами, обладает эвристическим потенциалом, который в полной мере реализован в системно-информационном представлении ментальных процессов в когнитивных науках. Вместе с тем через указанную аналогию в эмпирические исследования сознания проникают теоретические постулаты, ослабляющие позиции компьютерационалистской парадигмы, в том числе принцип минимальной корреляции, формальные ограничения теории алгоритмов и принципов эффективной вычислимости, проблема ментальной каузации, положение о соотношении символа и функционального состояния.
3. Успешное прохождение теста Тьюринга, реализованного в эффективных вычислениях на машине Тьюринга, оказывается недостаточным для объяснения сознания, т.к. в терминах теории



вычислимости эти машины не справляются с содержанием гёделевского предложения. Радикальный компьютеризм предъявляет вычислительным моделям сознания взаимоисключающие требования: 1) наличие детерминированной алгоритмической разрешимости; 2) способность к осознанию гёделевского предложения.

4. Концептуальный атомизм, обладая рядом слабых мест, тем не менее ценен тем, что остро ставит вопрос об онтологии содержания ментальной репрезентации. Семантика ментальных состояний требует особой информационной реализации, структурно-функциональные свойства которой должны осуществлять строгую каузальную связь между содержанием сознания и когнитивными процессами, обеспечивающими это содержание. Концептуальный атомизм позволяет сформулировать монистические положения информационного реализма, понимая ментальные процессы как информационные, но при этом избегая редуцирования информационных процессов к состояниям нейрофизиологической системы.
5. Концептуализация телесного опыта является содержательной основой для формирования феноменальных суждений. Телесность, фиксирующая в дорепрезентативном языке опыт обладания сознанием, является тем «видом» сознания, которое элиминируется из компьютеризма. Направление телеологической семантики включает в себя преимущества информационной трактовки репрезентации с учетом специфики теории воплощенного значения. Ключевую роль вновь начинает играть дорепрезентативный язык базовых абстракций, сформированный за счет организменно-средовой системы отношений. Смысловая трактовка сознания наделяет каузальными атрибутами се-

мантическую информацию, не имеющую однозначных физических коррелятов. Ментальное содержание, помимо динамики нейрональных процессов, усложняется различными видами контекстов: телесность, языковые игры, социально-исторический опыт, ситуативная обусловленность.

### **Теоретическая значимость исследования**

Научная значимость диссертационного исследования состоит в том, что в нем сформулированы новые теоретические положения:

- Процессы сознания требуют концептуального прояснения в контексте проблемы перехода ментального в физическое и физического в ментальное на уровне информационно-вычислительных моделей. Результаты исследования дают широкий концептуальный анализ указанных проблем в дополнительном «срезе» онтологии языкового знака и телесно-ориентированной парадигме.
- Раскрытые принципы информационной корреляции и их связь с проблемой ментального содержания открывают новые перспективы для развития когнитивных технологий создания новых кибертехнических систем, способных к адаптивным формам функционирования в изменяющейся среде, к принятию творческих решений, к формированию ассоциативных концептуальных связей.
- Семантический аспект вычислимости языковых процессов сознания позволяет расширить представления о связи информации и значения на уровне онтологии. Ментальное содержание возможно описать в терминах субсимвольного абстрактного языка, который реализуется в физических информационных процессах, но не представим в вычислительных моделях.
- Решение вопроса о границах информационной составляющей

феноменального опыта позволяет разработать новые модели семантической концептуализации, применимые при конструировании искусственных когнитивных агентов, способных к обработке информации от interoцептивных ощущений.

- Результаты исследования позволяют раскрыть методологические перспективы и пределы вычислительных моделей сознания в рамках как философских, так и инженерно-прикладных исследований.

### **Практическая ценность исследования**

Прикладная значимость ожидаемых результатов диссертационного исследования связана с их применением как в фундаментальных теоретических областях (онтология, теория познания, семиотика, когнитивная лингвистика, методология искусственного интеллекта), так и в прикладных системах анализа семантики естественного языка (автоматизированный перевод, машинный концептуально-семантический анализ, нейросетевое моделирование языковых процессов), а также в робототехнической инженерии. Исследование феномена вычислимости с позиций когнитивной семантики позволяет развить и усовершенствовать языковые человеко-компьютерные интерфейсы и компьютерные способы представления семантических процессов сознания. Философская интерпретация вычислительной теории сознания привносит ценный методологический материал для теории искусственного интеллекта в аспекте концептуально-семантического анализа при распознавании различных видов interoцептивных состояний.

Исследовательские материалы, представленные в диссертации, могут быть использованы в практике преподавания учебных курсов для бакалавров, магистров и аспирантов философских, технических и педагогических направлений подготовки и специальностей. В ФГБОУ ВО «Пятигорский государственный университет» результаты, полученные в диссертации, были внедрены в следующие учебные курсы: «Философия» для бакалавров, обучающихся по направлению подготовки «Бизнес-информатика»; «История и

методология науки» для магистрантов отделения прикладной лингвистики; «Методы и методология научного исследования» для магистрантов педагогических направлений, а также в программу профессиональной переподготовки «Управление научным контентом» для профессорско-преподавательского состава ФГБОУ ВО «Пятигорский государственный университет». Полученные результаты также отражены в Основных профессиональных образовательных программах высшего образования для аспирантуры по направлению 09.00.01 – Онтология и теория познания.

### **Степень достоверности результатов исследования**

Степень достоверности диссертационного исследования определяется верифицируемостью и репрезентативностью применяемых эмпирических теоретических данных, методологической обоснованностью использования актуальной научно-теоретической литературы. В работе обоснованно используются общенаучные и философские методы и аналитические подходы: теоретико-методологический инструментарий эпистемологии, аналитической философии сознания, концептуально-когнитивной семантики, применимый при концептуальном анализе конкретно-научных и философских вычислительных моделей сознания. Выбор данных аналитических шагов обусловлен целями, задачами, а также спецификой предмета исследования – семантических аспектов ментальных процессов в контексте вычислительной теории сознания.

### **Апробация работы**

Результаты диссертационного исследования отражены в 2 монографиях, 22 статьях в журналах из Перечня ведущих рецензируемых научных журналов и изданий, в которых должны быть представлены основные научные результаты диссертации на соискание ученой степени доктора наук, 2 статьях в журналах, включенных в международные базы данных (Scopus, ERIN PLUS) и 26 статьях в других изданиях.

Результаты диссертации были представлены в научно-теоретических докладах на международных и всероссийских конференциях:

Международная конференция «Д. Юм и современная философия» (Москва, ИФ РАН, 15-17 ноября 2011 г.); Всероссийская научно-практическая конференция «Проблема сознания в междисциплинарном измерении» (Москва, ИФ РАН, 20-22 марта 2012 г.); Десятый всемирный онтологический конгресс. X International ontology congress «From elementary particles to human nature» (Испания, Сан-Себастьян, 1-6 октября, 2012 г. / Барселона 8-9 октября, 2012 г.); Всероссийская конференция молодых ученых «Философия, Язык, Культура» (Москва, НИУ ВШЭ, апрель 2013 г.); Всероссийская научная конференция студентов, аспирантов и молодых ученых «Искусственный интеллект: философия, методология, инновации» (Москва, МИРЭА, 2012, 2013 гг.); Международная конференция «Следование правилу: рассуждение, разум, рациональность» (Москва, НИУ ВШЭ, 21-23 октября 2014 г.); Вторая всероссийская конференция молодых ученых «Человек в технической среде: конвергентные технологии, глобальные сети, Интернет вещей» (Вологда, ВоГУ, 2014, 2015 гг.); Международный конгресс по когнитивной лингвистике «Язык и сознание в междисциплинарной парадигме исследования» (Санкт-Петербург, СГЭУ, 30 сентября-2 октября 2015 г.); Международная научная конференция «Современная аналитическая философия» (Томск, Национальный исследовательский Томский государственный университет, октябрь 2015 г.); Международная конференция по транслатологии «Komunikacja międzykulturowa w świetle współczesnej translatologii» (Польша, Ольштын, Варминско-мазурский университет, 12-15 апреля, 2015 г.); Международная научно-техническая конференция «Перспективные информационные технологии» (Самара, СГАУ, 2015, 2017 гг.); Международный конгресс по когнитивной лингвистике «Антропоцентрический подход в когнитивной лингвистике» (Москва, ИЯ РАН, 8-9 ноября, 2016 г.); Всероссийская конференция «Философия науки и техники в России: вызовы информационных технологий» (Вологда, ВоГУ, 1-2 июня, 2017); Школа молодых ученых «Субъект, сознание и познание в контексте современной философии когнитивных наук» (Москва, ИФ РАН, 8 ноября 2017 г.).

Основные положения работы обсуждались в секторе теории познания Института философии Российской академии наук, на кафедре онтологии и теории познания Российского университета дружбы народов, а также на кафедре исторических, социально-философских дисциплин, востоковедения и теологии Пятигорского государственного университета.

Теоретические положения работы были апробированы в ходе реализации проекта, поддержанного ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009-2013 годы по теме «Развитие теоретико-методологических исследований по философии языкового сознания» (Соглашение № 14.В37.21.0510).

Диссертация состоит из введения, 5 глав (17 параграфов), заключения и списка литературы, включающего 461 источник, 241 из которых на иностранных языках. Общий объем работы составляет 407 страниц.

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обоснована актуальность исследования, охарактеризована степень теоретической разработанности темы, определен объект, предмет, цели и задачи работы, представлена теоретико-методологическая база исследования, сформулированы выносимые на защиту положения, подтверждена их новизна, охарактеризована теоретическая и практическая значимость результатов исследования, описана его апробация.

**В первой главе «Методологические трансформации в современной теории познания: от философских категорий к объектам науки»** рассматривается процесс трансформации традиционных эпистемологических категорий в объекты эмпирической науки. Выявляются методологические предпосылки и следствия научной «экспансии» предметной области философии сознания.

*Первый параграф «Проблема Д. Юма и антикартезианская революция в аналитической философии сознания»* посвящен выявлению и обоснованию влияния философских идей Д. Юма на развитие эмпирической психологии, компьютеризации и аналитической философии сознания.

На основании работ А.Ф. Грязнова, В.В. Васильева, А.Е. Питсона, Дж. Фодора, Дж. Биро, Т.М. Бетчер, А. Папа рассматриваются причины, по которым в начале XXI в. интерес к творчеству Д. Юма проявляется как со стороны аналитической философии, так и со стороны континентальной традиции. Идеи, высказанные шотландским философом, нашли свое воплощение в новых теоретических проблемах современной эпистемологии, философии сознания и философии когнитивных наук. Особенность моносубъектного эмпиризма Д. Юма состоит в том, что в этом подходе используется фундаментальный принцип формирования содержания сознания вследствие перехода от ощущений (впечатлений) к идеям (репрезентациям). Этот принцип заложен во многие формы современного философствования, которые осно-

вываются на наукоцентричных формах эмпиризма.

По мнению Д. Юма, работа сознания направлена на искусственное поддержание целостности воспринимаемого мира, но при этом разум не способен раскрывать реальную связь между предметами. В контексте философии сознания эти тезисы открывают глобальную проблему: интенциональность сознания не согласуется с его рефлексивностью. То есть содержанием сознания не могут стать механизмы управления сознанием.

Делается вывод о том, что антикартезианская революция, реализованная в трактовке сознания Д. Юма, находит свое продолжение в разнообразных философских течениях последующих веков. Юмовский отказ от субстанциального дуализма и от традиционных философских категорий в изучении познавательных процессов становится предпосылкой для теоретической возможности представить когнитивные процессы в технологизированном механистическом виде, что открыло в дальнейшем перспективы для конкретно-научных методов исследования.

Рассмотрены положения, доказывающие, что юмовская концепция послужила причиной для очищения проблемы сознания от философских спекуляций и своеобразной подготовкой сообщества философов и ученых к объективации классической философской категории. Д. Юм, выступая с критикой картезианского дуализма и анализируя содержание суждений о сознании, закладывает основы натуралистической философии, но при этом указывает на тщетность попыток совместить перспективы «от первого и от третьего лица». Такая трактовка сознания подводила к тому, чтобы либо отказаться от эмпирического объяснения сознания, либо вообще исключить понятие «личность» из исследований сознания. Обосновывается, что подобная трактовка нашла свое плодотворное продолжение в разнообразии течений научного физикализма: в необихевиоризме Б. Скиннера, логическом бихевиоризме Г. Райла, физикализме Р. Карнапа, теории приватного языка Л. Витгенштейна, элиминативизме Дж. Серла и Т. Нагеля. Перечисленные философские концепции послужат методологическим основанием для развития нейронаук и теории



искусственного интеллекта. Элиминируя личностный субъект из категории сознания, можно представлять сознание как технический объект.

Во *втором параграфе «Познавательные процессы как объект философских и конкретно-научных теорий на рубеже XX-XXI вв.»* изложены методологические основания кантианской теории познания в свете когнитивной революции; также рассмотрены причины, по которым радикальный эпистемологический конструктивизм стал неотъемлемой чертой не только нейронаук, нейрокибернетики, когнитивной психологии, но и социально-гуманитарной области знаний.

На основании работ В.А. Лекторского, П.Д. Шашкевич, В.А. Бажанова, У. Селларса, Р. Рорти, Р. Брэндома, К. Вестфаля показано, каким образом идеи кантианского конструктивизма по-новому переосмысляются в современных когнитивных исследованиях и нейронауках, выводя на первый план деятельность субъекта, создающую онтологические предпосылки для формирования конструирующих «фильтров». Когнитивные науки с успехом применяют кантианский принцип демаркации эмпирического знания от философско-умозрительного. «Когнитивная революция» объектом эмпирического познания полагает преобразующие конструктивные механизмы в деятельности человека и общества. В контексте данного исследования является существенным то, что многомерный феномен информации становится основным «строительным материалом», участвующим как в извлечении данных из среды, так и в построении каналов обратной связи в коммуникативной деятельности субъекта. Конструктивизм присущ когнитивным исследованиям несмотря на то, что эмпирические исследования сознания исходят из реалистических онтологических установок. В таком методологическом контексте созрели предпосылки для создания Вычислительной теории сознания.

Рассмотрены положения эмпирических теорий сознания, таких как Теория информационной интеграции (Дж.М. Эдельмана, Дж. Тонони), Теория рекуррентной обработки (В. Ламме), Теория микросознания (С. Зеки) и др. Указанные теории направлены на поиск нейрональных коррелятов. Обос-

новано, что проблема обнаружения нейрональных коррелятов сознания в эмпирических подходах вытекает из следующих исследовательских установок:

- поиск автономных модулей, последовательно обрабатывающих входящую информацию;
- поиск принципов интеграции информации из разных областей с целью формирования единой репрезентации (например, рекуррентная обработка);
- поиск физической локализации глобальной единой системы, суммирующей активность всех областей коры;
- обнаружение взаимосвязи феноменального сознания с высшими когнитивными процессами – вербальной рефлексией, осведомленностью.

Делается вывод, что в эмпирических теориях сознания информация тесно связана с физической организацией мозга и понимается в терминах качественного подхода к информации. Остается открытым вопрос, можно ли представить феноменальное сознание как информационный код высшего качественного состояния материи. Обосновывается тезис о том, что радикальный эпистемологический конструктивизм нейронаук обладает четкими методологическими границами. С одной стороны, кантианская программа заложила основы для исследований конструирующих функций сознания и его материальных коррелятов, с другой – эмпирический реализм ставит под сомнение возможность научного объяснения феноменального сознания.

*Третий параграф «Общие основания научных и философских программ исследования сознания»* посвящен анализу исторических и методологических предпосылок для формирования компьютерно-информационного подхода в философии и науках о сознании. На основании работ А. Сломана, Ф. Брэя, Дж. Соракера, Л. Флориди, Н.А. Ястреб, К.К. Колина, А.И. Ракитова, А.Д. Урсула, В.Б. Гухмана в параграфе рассматриваются такие понятия, как «вычисление», «алгоритм», «вычислимость», «вычислительная сложность», способы их миграции из теоретического поля математики и информатики в философский дискурс с последующим приобретением специфического иноска-

зательного содержания.

Обосновано, что фундаментальные категориальные планы философских исследований сознания на современном этапе развития когнитивных наук имеют свое воплощение в информационно-компьютерном измерении. Анализ объяснительных возможностей и методологических границ данной парадигмы относительно природы сознания представляет собой центральную ось данного исследования. Со стороны науки проблема сознания «распадается» на ряд предметных областей: биологический генезис сознания, языковая деятельность сознания, генетические свойства сознания и т.д. Также указаны причины терминологических неточностей информационной парадигмы в философии. В контексте проблемы сознания понятие информации теряет четкие границы объема, включая в себя свойства физических процессов, функциональные типы связей, синтаксис символических элементов, семантические свойства знания и т.д. Ясность теряется в результате «грубого» влияния так называемой «компьютерной метафоры», которая предписывает некорректные аналогии цифрового устройства при исследовании познавательных процессов человека.

В результате анализа работ К.В. Анохина, Г.Н. Рапопорта, А.Г. Герца, Дж. Тарди, П.О. Хайконена, Т. Маудлина, М.А. Арбиба показано, что информационная природа ментальных репрезентаций может рассматриваться как сфера для обоюдного интереса и со стороны философии, и со стороны науки. Для философии здесь, возможно, кроется ответ на традиционный вопрос о ментальной каузальности и на вопрос об универсальной информационной онтологии; для науки – возможность эмпирически обосновать корреляции между нейросетевыми ансамблями и ментальным содержанием. На значимость нейросетевых моделей в понимании когнитивных процессов указывается как в работах исследователей-когнитивистов, так и в работах инженеров-кибернетиков, т.к. для последних возможность проектирования «машинного» феноменального опыта через информационные архитектуры открывает широкие технологические перспективы.

Таким образом, вычислительный поворот в философии и революционные сдвиги в когнитивных науках привели не только к новым методологическим принципам понимания природы ментальных процессов, но и создали мировоззренческие предпосылки для формирования новой функционально-компьютерной антропологии, в соответствии с которой любой процесс человеческой деятельности (физический, поведенческий, психологический, феноменальный, социально-коммуникативный) возможно редуцировать к некоей вычислимой модели. Важной чертой философской Вычислительной теории сознания (особенно в ее позднейших теоретических надстройках) является ее установка на объяснение феноменального опыта и на информационное объяснение содержательной природы ментальных репрезентаций.

Во второй главе **«Этапы становления вычислительной парадигмы в современных теориях сознания»** подробно рассматриваются этапы становления компьютерной метафоры в когнитивных науках. Определяется влияние категориального аппарата компьютерных наук на аналитическую философию сознания второй половины XX в.

В первом параграфе **«Роль компьютерной метафоры в развитии когнитивных наук»** отражены методологические контексты, в рамках которых в философские теории сознания проникли представления из компьютерных наук и теории информации.

В параграфе обосновывается, что с самых своих ранних форм компьютеризация развивается в нескольких направлениях. И каждое направление содержит собственный объект, к которому применяются аналогии вычислительно-алгоритмических процедур.

- Во-первых, абстрактная модель машины Тьюринга применима к исследованию формальных и каузальных свойств функций нервной системы.
- Во-вторых, свойства структурных компонентов физической памяти компьютера переносятся на структурную организацию материи мозга.

- В-третьих, параллельно развивается «обратная метафора» – вычислительные процедуры живой материи переносятся на архитектуры искусственных вычислительных устройств.

Такого рода антропологический механицизм стал возможен благодаря тому, что в середине XX в. все еще были прочны позиции психологического бихевиоризма, рассматривающего природу человека через призму автоматизированных поведенческих реакций на разного рода стимульные вызовы природной и социальной среды. Модель «стимул-реакция» хорошо согласовывалась с количественными методами учета операций над символами в ленте памяти вычислительного устройства и с их представлением в виде машинных таблиц.

Таким образом, доказана плодотворность моделей машинного функционализма для дальнейшего развития вычислительной теории сознания. Однако впоследствии сторонники идеи машинных состояний были вынуждены признать несовместимость функционального определения параллельных (одновременных) ментальных состояний с последовательными дескрипциями машинной таблицы.

Однако такого рода методологические ограничения никак не повлияли на темпы роста популярности так называемой IP (information processing)-парадигмы в когнитивных исследованиях сознания и мозговых процессов. Сознание в упрощенном виде представляется как набор программных команд, исполняемых на биологических вычислениях мозга. Все многообразие подходов, которые выражают значимые эпистемологические последствия компьютерной метафоры, можно выразить в нескольких положениях:

- Сознание (разум) может рассматриваться как система, реализующая символические вычисления.
- Информация репрезентируется в сознании символически.
- Компьютерная программа и сознание реализуются в ходе сериального выполнения процедурных шагов. Когнитивные процессы в этом ракурсе приобретают свойства прерывной последовательности изменяемых со-

стояний.

- Анализ обрабатываемой информации проводится за счет регистрации символов вычисления (в терминах машины Тьюринга – печать символов на ленте памяти) и редукции ментальных операций к этим символам. Ментальные состояния в этом случае понимаются как результат обработки разных видов информации.

- Обработка информации организована по принципу стадий, распределённых по модулям всей системы.

- Когнитивные процессы, как и компьютерная обработка информации, протекает во времени. Длительность и хронологическая последовательность каждого процесса могут объяснить природу этого процесса и принципы его информационной организации.

- Сознание (как и любая вычислительная система) представляет собой систему с ограниченным набором функциональных возможностей.

Очевидно, что развитие вычислительных подходов к человеческим процедурам познания оказало серьезное влияние на становление инженерных методов искусственного интеллекта. Энтузиазм, свойственный исследованиям 60-х гг. по поводу интеграции методологии искусственного интеллекта и когнитивной психологии в единое знание о «механике» интеллектуальных процессов человека, постепенно сменился более взвешенной позицией. Такая тенденция связана с тем, что интенциональные свойства сознания оказались непомерно шире репрезентативных функциональных процедур. К этому добавлялась непрояснённая семантическая природа самих вычисляемых символов и не потерявшая актуальности традиционная психофизическая проблема сознания.

Во *втором параграфе «Функциональные модели сознания и репрезентационализм»* раскрываются методологические характеристики машинного функционализма и анализируется его влияние на развитие вычислительных теорий сознания.

В данной части работы подробно рассматривается линия аргументов

Х. Патнэма в защиту машинного функционализма и ряд методологических противоречий, которые из этой защиты вытекают.

Слабые стороны вычислительного функционализма обнаруживаются вследствие противоречий, вытекающих из понятия вычисления, и из того, что функционалистам не удастся избежать семантического аспекта при оперировании аргументами лингвистического типа. В машинном интеллекте невозможно представить выражение «1 есть один» или «0 есть ноль», т.к. синтаксис любой сложности будет редуцироваться к электромагнитным состояниям памяти. Тождество значений требует тождества экстенционалов, но в машинном представлении «1» и «один» не тождественны. То есть в терминах функционализма экстенционал не определяет содержание.

В диаде «мозг-компьютер» возникает противоречие: с одной стороны, ментальное содержание – это результат функционально-вычислительных состояний некоей материальной системы, с другой – ментальные содержания определяют экстенционалы. Если мозг (согласно функционализму) – это функционально-биологический вычислитель, существует возможность двух тождественных функциональных состояний у двух организмов с различным ментальным содержанием.

Защита аргумента множественной реализации впоследствии привела к созданию философской программы расширенной «теории» референции, в которой экстенционал не зависит от ментальных состояний. В терминах компьютеризации этот тезис формулируется так: сознание может быть сравнимо с вычислительной машиной, если тождественные состояния машин Тьюринга способны реализовывать нетождественные ментальные состояния.

Таким образом, обоснована глубинная методологическая взаимосвязь между ранними формами машинного функционализма и когнитивными науками. Функционалисты путем концептуального анализа доказывают, что ментальные состояния суть вычислительные состояния мозга. Становится очевидным, что аргументы об экстенциональной нетождественности никак не затрагивают методологические основания когнитивных наук. Когнитивные

исследования могут помочь философии сознания в вопросе о вычислительной природе механизмов репрезентаций, извлекаемых, например, из визуального опыта (изоморфная кластеризация образов, формирование трехмерной модели из двумерных ретинальных образов и т.п.), но они остаются малоинформативными в вопросе об онтологии ментального содержания.

Анализ этапов становления компьютерного функционализма приводит к выводу о том, что радикальные формы данного направления сводятся либо к материализму, либо к бихевиоризму уровня «сигнал-ответ». Описания вычислительных механизмов ментальных репрезентаций никак не объясняют онтологию ментального содержания. Символы как материал вычислений обладают рядом качеств, способных присваивать физическим системам репрезентативные свойства. Однако физические процессы накладывают существенные ограничения на корреляционные возможности репрезентативных систем.

*Третий параграф «Нейросетевое представление процессов сознания»* посвящен следующему витку развития компьютерной метафоры, связанному с успехами в области нейрофизиологии. Главным исследовательским объектом становится мозг как нейронная сеть, состоящая из узлов входящей / исходящей информации и различных типов соединений. Если в предыдущей версии машинного функционализма мозг понимался как биологический эквивалент бесконечной ленты машинной памяти, то в теории нейросетей, наоборот, – принципы организации биосистем были приняты за основания для конструирования искусственных нейронов.

В параграфе на основании работ Дж. Хопфилда, У. МакКуллака, О'Рэйли, Ю. Мунакаты, Р. Рохаса, Л. Черняка, В.Г. Редько, Ф. Розенблатта, М. Мински рассматриваются «технологические» группы моделей вычислимости. Проанализирован ряд методологических упущений при логическом моделировании функций биологического нейрона:

- *Во-первых*, логические элементы нейронных вычислителей работают по принципу универсальной машины Тьюринга и не способны избежать



последовательности в обработке входящих весов.

- *Во-вторых*, принцип «все или ничего» не является определяющим для биологических нейронов. В состоянии покоя живой нейрон генерирует низкочастотные импульсы и способен использовать вероятностный компонент возбудимости, когда один и тот же стимул приводит к различным результатам.

- *В-третьих*, слабой стороной нейросетевой аналогии является отсутствие в мозге точной синхронизации с помощью постоянной тактовой частоты, которая при этом необходима для обеспечения корректной работы полупроводниковых процессоров.

- *В-четвертых*, в биологических нейронах информационные процессы носят аналоговый характер в силу разнообразия химических составов нейромедиаторов (аминокислоты, пептиды, моноамины), функциональные свойства которых невозможно воспроизвести в цифровой логике искусственного вычислителя.

Раскрыты полемические аспекты классификации коннекционизма. Доказано, что вычислительные абстракции в нейросетевых моделях меняют свои атрибуты в зависимости от принципов реализации. Символьные структуры, активность паттернов или отдельных узлов обрабатываются правилами распределения весов групп соединений или весов индивидуальных соединений соответственно. Таким образом, вычисление (понимаемое как реализация вычислительной абстракции) является неоднородным сложным понятием, теоретико-методологические трактовки которого порождают противоречащие друг другу вычислительные гипотезы о природе сознания.

***В четвертом параграфе «Перспективы и пределы нейрофилософии в решении проблемы сознания»*** проанализированы и сопоставлены достоинства и недостатки нейросетевого подхода в инженерном и философском измерениях.

Рассмотрен основной тезис коннекционистской версии вычислительной теории сознания, который сводится к тому, что объяснение ментальных

свойств сознания необходимо начинать с нейросетевой структуры материи мозга. То есть коннекционизм становится неотъемлемой частью биологического реализма. При этом остаются значимыми компоненты классического функционализма, утверждающего, что полноценные интеллектуальные функции разума реализуемы на любом носителе, поддерживающем вычислительную архитектуру биологического мозга.

Сильные стороны коннекционистской модели (мягкие ограничения, амортизация отказов системы, ассоциативная память, нечеткая логика, способность к обучению) широко используются в инженерных реализациях интеллектуальных процессов. Тем самым, с одной стороны, укрепляются методологические позиции компьютерной метафоры, с другой – исследовательский фокус смещается с проблемы ментального содержания в сторону архитектуры мозговых связей.

Делается вывод о том, что нейрофилософия как направление, использующее достижения инженерного коннекционизма к проблеме сознания / мозга, приобретая ряд преимуществ, неизбежно наследует все слабые стороны коннекционизма и физикализма:

- a. биологические вычисления несводимы к цифровым архитектурам;
- b. понятие биологического вычисления не доопределено в терминах теории информации;
- c. моделирование вычислительных свойств живой материи ничего не сообщает о каузальной связи ментального и физического.

Стремление эмпирических наук о сознании к компьютерной реализации когнитивных функций приводит к философским проблемам, связанным с основаниями физикализма. Семантические свойства физической информации не являются структурными элементами самой информации. Содержание опыта (знания) извлекается из косвенных инференций, которые не имеют од-

нозначных физических коррелятов.

В третьей главе **«Методологические основания критики радикального компьютеризма»** речь идет о методологических основаниях критических философских программ, направленных на ослабление методологической значимости компьютеризма. Логико-математическая и кибернетическая трактовки философских вопросов о формах познания, природе сознания, свободе воле породили целое направление философии ИИ, которое оказало прямое влияние на исследовательские программы в философии сознания.

В первом параграфе *«Роль мысленных экспериментов при построении вычислительных моделей сознания»* исследуется эвристический потенциал теста Тьюринга и его вариаций в рамках мысленных экспериментов для аналитической философии сознания.

Раскрывается проблема верификации интеллектуальных функций. Из полемики наиболее известных философских программ «интеллект как исчисление понятий», «интеллект как восприятие» и «интеллект как рефлексия» вытекает, что в окончательном виде определить интеллект невозможно. Существует так называемый «интервал Тьюринга» – совокупность машинных задач, решение которых позволяет признавать за машиной разумность в рамках какой-либо философской программы. Следовательно, машинная имитация воспроизводит лишь тот набор свойств сознательной деятельности, который в той или иной философской парадигме представляется достаточным для прохождения теста. Из этого следует, что прохождение машинным интеллектом теста Тьюринга верифицирует лишь те методологические установки, в рамках которых описываются вычислимые свойства интеллекта.

Проанализированы методологические основания функционализма и проблемы логической представимости феноменальных свойств машинного интеллекта. Проблема логической представимости вычислительного устройства, функционально тождественного вычислительным архитектурам мозга, усложняется тем, что функции каузальных отношений мозга и сознания в со-

временных нейронауках не доопределены. Это позволяет предполагаемому разработчику подобного устройства называть свое устройство сознательным, т.к. даже отсутствие квалиа остается в рамках непреодолимой приватности. Для оценки интеллектуальности и сознательности в распоряжении наблюдателя имеются лишь различные формы компетентности системы, достаточной для решения инженерных задач. Таким образом, проведена строгая демаркационная линия между машинным функционализмом и эссенциализмом. Для Дж. Серла и авторов мысленных экспериментов типа «Китайская комната» важны онтологические сущностные черты обладания разумом, для инженеров достаточно симуляции деятельности мозга или получения непротиворечивого результата на выходе.

Подтверждена значимость инженерного принципа минимальной корреляции. Проблема минимальной корреляции (поиск достаточных условий для реализации интеллектуальных функций) сводится к воспроизведению в технических системах нескольких принципов:

- a. согласование входных данных с содержимым эпизодической памяти;
- b. самоорганизация правил управления конкуренцией между модулярными процессами;
- c. соблюдение принципа агентоцентричности.

При этом, если допустить физическую представимость и техническую реализуемость машинного «зомби», возникают затруднения метафизического плана, выраженные следующим списком вопросов:

- a. Что является денотатом машинных феноменальных суждений?
- b. В чем состоит критерий истинности выражений машинных разноразрядных языков?
- c. Может ли искусственная вычислительная физическая система обладать эмерджентными нерегистрируемыми со стороны раз-

работчика состояниями?

Таким образом, с одной стороны, инженерная представимость машинного разума с рядом допущений отвечает на вопрос о логической возможности адекватной машинной имитации интеллектуальных функций, с другой стороны, компьютерная модель разумного поведения не дает ответов на фундаментальные философские вопросы о природе сознания.

***Второй параграф «Онтология информации и проблема вычислимости»*** демонстрирует ряд формальных ограничений вычислительных моделей сознания, вытекающих из онтологических свойств данных и информации.

Прежде всего, речь идет о критике компьютерационалистских программ со стороны менталистов. Ментализм гёделевского типа ставит дилемму: либо существуют абсолютно неразрешимые математические утверждения, либо человеческий ум превосходит конечную машину. Если природа сознания в полной мере не объяснима компьютерационалистскими моделями, то невозможно и построение вычислительного устройства (конечного автомата), адекватно воспроизводящего когнитивные отношения «входа» и «выхода», реализованные в функциях сознания.

В области определения эффективной вычислимой функции разворачивается основная полемика между менталистами и механицистами, в которой первые полагают, что гёделёво предложение непреодолимо для конечных алгоритмических систем, вторые – что базовые свойства человеческого мышления могут быть формализованы либо на уровне компьютерационных моделей работы мозга, либо на уровне синтаксических архитектур мыслительных процессов. При этом, с позиций механицизма, формальные ограничения арифметических процедур – недостаточное условие, чтобы отказать машинному интеллекту в возможности реализации человеческих интеллектуальных функций или в объяснении онтологии сознания, т.к. существуют иные формы (кроме тьюринговской) представления вычислимости.

Исследованы условия, при которых успешное прохождение теста

Тьюринга, реализованного в эффективных вычислениях на машине Тьюринга, оказывается недостаточным для объяснения сознания, т.к. в терминах теории вычислимости эти машины не справляются с содержанием гёделевского предложения. Радикальный компьютеризм предъявляет сознанию как объекту взаимоисключающие требования: 1) наличие детерминированной алгоритмической разрешимости; 2) способность к осознанию гёделевского предложения. С инженерной точки зрения критика компьютеризма выглядит необоснованной, т.к., если представить ИИ как класс объектов, способных решать определённый класс задач вычислительными методами, то противоречия снимаются. ИИ не способен решить задачи, доступные для естественного разума, т.к. этот класс задач невозможно привести к строгому непротиворечивому описанию.

Доказано, что компьютеризмская парадигма вместила в себя элементы теории данных и элементы теории информации физических систем без достаточной проработки на предмет методологической согласованности этих двух подходов. К трудным проблемам вычислительных моделей сознания отнесены следующие:

- проблема генератора / исполнителя инструкций вкупе с вопросом о символьной кодировке;
- проблема аппроксимации в цифровых моделях биологических процессов;
- проблема алгоритмического описания эмерджентных свойств системы.

Обоснованы заключения о том, что информационные вычислительные системы не имеют дела с репрезентативным содержанием. Семантика данных при любом уровне вычислительной сложности редуцируется к реализации численной меры (информации) в физических состояниях аппаратных узлов. Этот принцип тождества не оправдывает себя в философии сознания.

*В третьем параграфе «Классификация критических программ*

*компьютационализма»* проанализированы основания классификации анти-компьютационалистских подходов, предложены новые принципы разграничения аналитических и онтологических программ. Обоснована необходимость семантического ракурса построения критики радикального компьютеризма.

К типичному списку критических программ Вычислительной теории сознания, которые можно условно обобщить названием «онтологическая критика» (Аргумент от синтаксической комбинаторики – «Китайская комната»; Аргумент от тривиальности; Аргумент от гёделева предложения; Аргумент от темпоральности; Аргумент от телесности), целесообразно добавить блок аналитической критики. Специфика данной линии аргументов состоит в том, что критический фокус смещается на проблему значения.

Семантический аспект Вычислительной теории сознания в рамках аналитической критики имеет две непересекающиеся интерпретации: в терминах ментальных состояний и в терминах символьных операций. В итоге из любой интерпретации следует заключение, что, не обладая феноменальным содержанием, физические информационные системы накладывают жесткие синтаксические ограничения на принципы согласованности и технической реализуемости. Специфика семантических свойств символьных систем состоит в том, что денотатом некоторого предложения является эффективно вычисляемое аппаратно реализуемое выражение, а не некое положение вещей «внеязыкового мира». Следовательно, основная функция символьной семантики – это согласованность синтаксиса программы с кодами операций, режимами адресации и числовыми эквивалентами, которые будут реализовываться в физических состояниях машины. Несмотря на то, что данное определение пересекается с функционалистской трактовкой сознания / мозга, необходимо признать следующее: если содержательная сторона феноменального опыта является ключевой для объяснения сознания, то компьютеризаторское понимание семантики в этом вопросе выглядит бесполезным.

**Четвертая глава «Проблемы семантики в вычислительных моде-**

лях сознания» посвящена семантическим аспектам вычислительной философии сознания и обоснованию связи семантических программ с фундаментальными вопросами вычислительных подходов.

В первом параграфе «Семантика и онтология: когнитивные аспекты проблемы» уточняется ряд терминологических трактовок, проводится анализ эпистемологических аспектов когнитивной семантики и выявляется связь вычислительной семантики с философскими теориями сознания. Аргументируется тезис о центральной роли проблемы значения в вычислительных моделях сознания.

Для Вычислительной теории сознания (как и для теории ИИ) фундаментальным является ответ на вопрос: каковы функциональные механизмы построения ментальных моделей на основании пропозициональной репрезентации?

Результаты исследования в области когнитивной семантики оттеснили хомскианскую идею следования правилу в рамках формально-алгоритмического описания языковой системы. Принципы семантической организации языкового выражения впервые были выведены на основе телесности и метафорических конвертаций физического и социального опыта человека. Для сторонников Вычислительной теории сознания стало необходимым построить систему доказательств, обосновывающих, что содержание ментальных репрезентаций есть результат функционально-информационных процессов, которые, помимо физических каузальных связей, способны к порождению феноменального опыта, вербальной осведомленности, построению отчетов. Семантика естественного языка оказалась связующим звеном между ментальными и физическими процессами.

Развитие вычислительных подходов к проблеме содержания языкового выражения отразилось на понимании природы ментальных процессов. В теории ментальных моделей последовательно доказывается витгенштейнианский тезис о том, что для понимания необходимо построение ментальной модели дискурса. То есть свойства ментального содержания выводятся из



свойств вычислительной референциальной системы. Итогом этих размышлений становится утверждение, которое является основанием для атомистической компьютерной теории концепта, в свою очередь ставшей фундаментом для Вычислительной теории сознания: *содержание пропозициональной репрезентации, обладающее условиями истинности, позволяет строить непротиворечивые модели дискурса на своеобразном абстрактном языке за счет семантических процедур сопоставления и вывода*. В этом смысле понимание – автоматический процесс. Так деятельность сознания стало возможно редуцировать к процессам автоматизированного вывода, основанного на пропозициональном содержании.

Во *втором параграфе «Связь теории концепта с вычислительной теорией сознания»* речь идет о связи когнитивных наук и философских теорий сознания с теоретическими и практическими подходами в изучении феномена концепта и когнитивного процесса концептуализации.

Проанализированы этапы становления теории концепта с рассмотрением слабых и сильных сторон каждого из подходов, начиная от классической дефинитивной теории, заканчивая концептуальным атомизмом.

Обосновывается, что идеи Концептуального атомизма обладают наибольшей экспликативной силой, т.к. «обходят» большинство традиционных для Классической теории и Теории прототипов проблем. При этом требуются прояснения вопросы об онтологическом статусе концептов (являются ли концепты репрезентациями, и обладают ли они информационной природой) и о вычислимости ментального содержания. Особое значение имеют следующие постулаты, обобщенные на основании различных теорий концептов и не противоречащие атомистическому подходу:

- прототипы не образуют какой-либо строгой модели для обработки данных в терминах процедурной семантики (это размытая система семантических связей);
- представление категорий в виде конечной совокупности признаков не способно охватить многообразие опыта репрезентаций;

- концепт является содержательным ментальным состоянием, а не интуитивной абстракцией;
- проблема значения имеет прямое отношение к содержанию ментальных состояний сознания;
- концептуализация реализуется через оперирование атомарными элементами знания.

Из представленных постулатов делается заключение: если ментальные процессы сознания представить как процессы концептуальной семантизации опыта, то вычислительные модели сознания релевантны лишь на структурном и операциональном уровнях, но не на содержательном.

*Третий параграф «“Язык мысли” и Вычислительная теория сознания»* посвящен проблеме онтологии концепта в контексте соотношения атомизма Гипотезы языка мысли (LOTH) и Вычислительной теории сознания (СТМ) Дж. Фодора.

Проанализирован ряд фундаментальных противоречий модулярного подхода. Важнейшим элементом воспроизводства как частной (ситуативно и контекстуально обусловленной), так и общей (фоновые знания) ментальной модели является блок памяти. Память осуществляет также разделение концептов от сценариев, фокуса от рамки, индивидуального от социального. Компьютационализм способен объяснить принципы порождения и интерпретации имплицитных и эксплицитных содержаний сознания. Но в этом случае утверждение о предзнаковой (доязыковой) природе ментальных моделей и абстрактного LOT, в терминах которого осуществляется концептуализация модулярных данных, явно противоречиво.

Это противоречие вытекает из природы информации и информационных процессов, участвующих в языковых процессах. Информация как содержание концептуальных контейнеров, возможно, каузально никак не связана с физической информацией о состояниях мозга или общей когнитивной системы субъекта. То есть базовый онтологический вопрос всей программы компьютеризации сводится к следующему: что вычисляется? С одной сторо-

ны, языковая информация связана с проблемой референции и значением языкового выражения, с другой – языковая информация интегрирована в когнитивную деятельность (принятие решений, понимание, действие, планирование, предсказание, выбор целей и т.д.). Для теории ментальных моделей остается нерешенным, какой тип информации концептуализируется, семантический или когнитивный.

Концептуальный атомизм, обладая рядом слабых мест, тем не менее ценен тем, что остро ставит вопрос об онтологии ментальной репрезентации. Семантика ментальных состояний требует некоей информационной реализации, структурно-функциональные свойства которой должны осуществлять строгую каузальную связь между содержанием сознания и когнитивными процессами, обеспечивающими это содержание. Если концептуализация формируется за счет вычислительной комбинаторики атомами знаний, то необходимо, чтобы эти атомы были представлены не только на уровне абстракций, но и на уровне материальных процессов. Вычисление требует аппаратной составляющей. При этом важно, что многообразие репрезентативных форм превосходит возможности процедурной семантики, применимой к атомарным знаниям.

Концептуальный атомизм в контексте проблемы сознания позволяет занять умеренную методологическую позицию:

- с одной стороны, заменить ментальные процессы информационными;
- с другой стороны, не сводить информационные процессы к состояниям нейрофизиологической системы.

Таким образом, постулируется особый онтологический статус информационных процессов сознания. Философскую программу, обосновывающую этот статус, можно рассматривать как форму умеренного информационного реализма.

Между ментальным содержанием и когнитивными процессами лежит

символ-посредник, природа которого эмпирически не прояснена. Наборы правил, управляющие принципами формирования атомов знания, не обеспечивают возможность их вычислительной реализации (даже на абстрактном уровне). Необходимо признать наличие некоего досодержательного уровня, который заполняет зазор между физическим состоянием системы и семантикой языковых выражений (даже если речь идет о семантике программного «языка мышления»), т.к. информация (как содержание концептуальных «контейнеров») каузально не связана с физической информацией в когнитивной системе субъекта.

**Пятая глава «Модели представления знаний в естественных и искусственных интеллектуальных системах»** посвящена обоснованию, с опорой на когнитивно-биологическую концепцию знака, положения о методологическом пределе когнитивного подхода к единству символического вычисления и порождения знаковых систем в процессах сознания.

В *первом параграфе «Семантические аспекты интенционализма»* рассматривается когнитивно-биологическая модель языкового знака в онтологическом и эпистемологическом измерениях.

Онтологический статус языкового знака в когнитивном подходе опирается на традиционную репрезентациональную теорию сознания. В когнитивных науках репрезентацию принято понимать как некоторое содержание опыта, основная функция которого состоит в описании, представлении и сохранении результатов взаимодействия со средой.

Ключевой вопрос относительно онтологического статуса знака связан с проблемой причинности. Известна так называемая Каузальная теория концептуальной репрезентации. Суть ее сводится к онтологическому обоснованию базовых концептуальных структур, кодирующих чувственный и когнитивный виды опыта в систему значений. Р. Джекендофф выделяет несколько уровней репрезентации (фонологическую, синтаксическую и концептуальную) и полагает, что опредмечивание результатов чувственных данных и когнитивных процессов происходит в результате «кодирования значения» по-

средством концептуальных структур. Все многообразие опыта, таким образом, сводится к чистым концептуальным формам, которые позволяют манипулировать формальным значением.

Из предположения о существовании внутренних абстрактных механизмов кодирования опыта (механизмов, работающих в системе отношений организм-среда) строится теория адаптивной концептуальной семантики. Согласно данной теории, концептуальная картина мира есть интегрированная система усложненных прототипических концептов. «Эта система, с одной стороны, отражает обобщенный опыт взаимодействия со средой, а с другой стороны, она сама является объектом взаимодействия». «Осознанное» поведение организма в этом случае определяется как поведение, направленное на согласование концептуальных картин мира.

В концептуальной семантике присутствует противоборство двух подходов: эксперенциально-экологического и компьютерного. В первом случае языковая прототипическая категоризация основывается на уникальном опыте языкового субъекта, помещенного в культурную среду, и его «перспективы от первого лица». Во втором – категоризация и классификация объектов сводится к компьютерным семантическим онтологиям, в которых знание представлено как база данных для оперирования градуализированными метаданными.

Интенциональность как важнейший компонент эпистемологических свойств знака реализует такой элемент когнитивной системы, как значимость в рамках процесса взаимной каузации организма и среды. Речь и языковое поведение в данном случае выступают как мета-дескрипции некоторого положения вещей в мире, а интенциональность – как способ реализации значимости. Интенциональность как когнитивную функцию организма в теории живых систем можно понимать следующим образом: «функции, рассматриваемые как когнитивные, суть результат деятельности специализированной подсистемы организма, непрерывно воспроизводящей паттерны, функционально или референциально соотносящиеся с изменениями среды». Из дан-

ного положения вытекает, что знаковая (семантическая) природа сознания, используя субстанциальный ресурс биологической материи (единый для всех живых организмов), тем не менее обладает специфической структурной организацией поведенческого коммуникативного опыта.

Таким образом, изложенное приводит к затруднениям в натуралистическом объяснении единства семантики и онтологии. Эти затруднения сохраняются до того момента, как будет открыт физический компонент организма, строго регистрирующий результат семантизации отношений организма и среды. Хранилищем репрезентаций как в естественных, так и в искусственных интеллектуальных системах является физическая память. В первом случае этот отдел отвечает за накопление опыта, во втором – за каузальные свойства программных команд. Если говорить о функциональных элементах сознания, реализованных в ИИ, то цифровой машинный опыт верифицируется лишь в вычислительно разрешимой непротиворечивости, чего явно недостаточно для ориентирования в аналоговом мире природных процессов и метафорическом мире человеческих смыслов.

***Второй параграф «Информационная природа репрезентации»*** посвящен анализу информационных свойств репрезентации. Информационная природа репрезентативных структур и элементы автоматизированных действий, которые делают отчасти правомерными вычислительные подходы, напрямую зависят от структурной и процессной организации памяти.

Для нашего подхода, представленного в данном диссертационном исследовании, важно экспериментально подтвержденное участие семантических процессов в обеспечении концептуального единства репрезентативных систем. Теория двойного кодирования конкурирует с Гипотезой языка мысли Дж. Фодора и с Теорией ментальных пропозиций З. Пилишина, но при этом все теории исходят из одного положения: существуют некоторые виды знания, которые извлекаются в результате информационных процессов и которые преобразуются с помощью вычислительных процедур. Единицей вычислительной обработки может выступать концепт, репрезентация или логиче-

ская пропозиция.

В параграфе показано, что информационная природа ментальных репрезентаций обладает особыми интенциональными характеристиками, которые передают вычислительным процессам особые семантические свойства. При этом компьютеризация применима лишь в объяснении поведения и причин принятия решений, при условии, что ментальные репрезентации трактуются лишь как некие бессодержательные модальности. При этом наличие содержания ментальных репрезентаций интуитивно определяемо каждым носителем сознания. В вопросе о соотношении интенционализма и компьютеризации существуют традиционные версии ответов:

- Вычислительные процедуры мозга не достаточны для обеспечения интенциональных содержаний .
- Когнитивные процессы не задействуют интенциональные состояния сознания, но представляют собой лишь функциональную комбинаторику вызовов среды и «ответов» системы.
- Когнитивность есть вычисление, минуя содержание репрезентаций.

Если рассматривать репрезентации с позиций концептуальной семантики, то, с одной стороны, содержание высказывания полностью зависит от предзаданной концептуальной информации, а с другой – формирует и изменяет концептуальную систему. В данном случае концептуальная система может приравниваться к репрезентациям, т.к. ментальное воспроизведение знания обеспечивает интенциональное содержание высказывания. Информация, содержащаяся в концептуальной системе, служит как для восприятия (выделения и маркирования) реальных объектов, так и для обработки языка как особого объекта (нестрогая семантика). Концептуальная система соотносит довербальное знание с порождающими возможностями языковой системы и кодирует с помощью языка «фрагменты» концептов, релевантные для коммуникации (синтаксически устойчивые структуры).

Концептуальная система обладает двумя состояниями: 1) размытое

дорепрезентативное прототипичное состояние, не обладающее дескриптивными свойствами; 2) репрезентативное состояние, адаптивно представляющее концепт в языке. Очевидно, что речь идет о различных режимах работы системы. Таким образом, допустимо, что информация, применимая к репрезентативной «кодировке» концептуальных фрагментов, ничего не имеет общего с информацией, обеспечивающей прототипические модели.

Собственно, фодоровская гипотеза LOT претендует на то, чтобы присвоить имплицитной памяти некий скрытый синтаксис для фиксации и передачи результатов неосознаваемого опыта. По Дж. Фодору, «Язык мышления» вычислим, но недеklarативен – на наш взгляд, это явное противоречие. Исследования процессов памяти указывают на то, что семантические свойства репрезентаций обуславливаются информацией, не доступной для декларативных волевых процедур. При построении лексико-ассоциативных карт испытуемые отмечают «автоматическое» формирование образов. Образы (репрезентации) формируются как из фрагментов недавних впечатлений, воспоминаний и ассоциаций, так и из фрагментов содержания долговременной памяти.

На основании вышеизложенного можно утверждать, что в рамках когнитивных наук не решается однозначно вопрос об информационной природе репрезентаций, т.к. результаты экспериментов демонстрируют либо устойчивые состояния мозговых зон при однотипных процессах, либо устойчивые эксплицитные свидетельства со стороны самих испытуемых. То есть речь идет либо о физической информации (отчет о состоянии функциональной системы), либо о семантике высказывания. Таким образом, для натуралистического объяснения репрезентаций необходимы эмпирические доказательства строгой связи этих двух типов информации с ментальным содержанием.

*Третий параграф «Феноменальное и вычислимое в языковых процессах сознания»* посвящен проблеме информационного объяснения феноменальных качеств в контексте знакового подхода к сознанию.

Феномениалистическая онтология сознания является одним из ключе-



вых направлений в современной философии сознания и имеет свою традицию постановки вопросов. Если исключить проблему субъекта как каузального основания чувственного синтеза, то возможно машинно-функциональное представление познавательных процессов: информация подается на ввод, преобразуется в рамках конечных алгоритмов функциональных состояний системы и затем подается на вывод (это может быть любой процесс – печать страницы, банковская транзакция, поднятие роботизированной руки, словесное признание в любви и т.д.). Главное, что преобразователь четко регистрирует каузальные отношения между сигналами ввода и вывода. В вопросе с человеком каузальность в отношениях ментального и физического становится неопределимой. Согласно феноменализму, человеческая склонность переносить опыт с известного на неизвестное является важнейшей особенностью когнитивных способностей. В этом случае универсальная система концептуальных схем, накладываемая на опыт восприятия вещей, разрушает информационные каузальные связи.

Феноменальная онтология сознания пытается ответить на вопрос именно о тех свойствах сознания, для которых трудно найти информационные корреляты (если понимать информацию как свойство материи). Но если под информацией понимать субъективную ментальную модель, то эти подходы совместимы, но с тем условием, что ментальная модель чаще всего не алгоритмизируема и невычислима. Таким образом, мы приходим к парадоксальной мысли о возможности информационного объяснения сознания, но вне компьютеризации, предполагающего вычислимость и алгоритмируемость.

Стоит отметить, что философская дискуссия вокруг ментальной каузальности и проблемы физической замкнутости ведется специалистами не первый год, и в ней не используются последние достижения психолингвистики, нейролингвистики и био-эпистемологии. В некоторых междисциплинарных работах очевидны разнонаправленные векторы: в философии сознания феноменальный репрезентационализм продолжает успешно противосто-

ять критике, при том, что сознание, понимаемое как феноменальная репрезентация, ставит в тупик лингвистическую семантику. За формально-синтаксическими состояниями языковой системы стоит бесконечное семантическое творчество, основанное как на социальных языковых играх, так и на эволюционных адаптационных свойствах организма (организма носителя языка и сознания). Что-то в сознании и в языке не следует правилам репрезентации, а постоянно «пишет» и видоизменяет эти правила.

В параграфе рассмотрены принципы формирования концептуальных связей в семантике естественного языка и в искусственных интеллектуальных системах. Обосновано, что информационные процессы интеграции данных отчасти объясняют принципы семантических представлений знаний в различных контекстуальных системах. Ключевое отличие «слепого» вычисления релевантных значений от ментального процесса семиозиса состоит в том, что основанием для вывода в первом случае является иерархия классов, протоколы правил и применяемые к ним эвристики и стратегии, во втором случае основанием для формирования знания является качественный опыт и метафорическая концептуализация, реализуемая в коммуникативной практике. Компьютерные системы способны к непротиворечивому представлению таксономических данных на основании свойств и атрибутов классов.

Современные исследования по деривационным свойствам базовых концептов отчасти реабилитируют фодоровскую гипотезу о языке мышления (LOTH), классическую архитектуру нейронной базы и вычислительную теорию в целом. Комбинационная структура семантических примитивов позволяет выводить множество значений в различных контекстах, размывая и расширяя семантику ментальных представлений.

На основании фреймового анализа некоторые современные авторы предлагают компьютерные функциональные модели сознания. Всякий раз в подобных работах подчеркивается, что модель призвана имитировать биологические процессы обработки сенсорных данных и принятие решений автономными системами. Подобные модели, как правило, включают в себя сен-

сорно-моторные блоки, модули эпизодической памяти, самообучающиеся системы и мотивационный процессор. Отдельным блоком стоит семантическая память, которая связана с блоком эмоций и субкортикальными вычислениями (Emotion, rewards and sub-cortical processing), с блоком переключения внимания (Attention switching), с сенсорными процессорами (Sensor processors) и с эпизодической памятью (Episodic Memory).

Блок семантической памяти представляется как самое уязвимое место всей вычислительной модели сознания. Здесь проходит граница между инженерной и гуманитарной интерпретацией семантических процессов. Нечеткость машинных выводов никак не соотносится со спонтанностью и метафоричностью естественной семантики. Пока машина работает с данными и информацией, а не со значением и смыслом, противоречивая, но эффективная естественная семантика остается привилегией человека.

В заключительном *четвертом параграфе «Телесность и действие: методологические пределы вычислительных моделей»* рассмотрены преимущества и недостатки телесно-ориентированной парадигмы исследования сознания, обоснован объяснительный потенциал телеологической семантики, проанализирована роль interoцептивных ощущений в формировании феноменальных суждений.

Существует ряд затруднений, которые возникают при попытке синтеза вычислительных аспектов когнитивной деятельности с телесно-ориентированными и психосемантическими моделями. В первом случае неизбежно вновь встает вопрос о статусе информационной составляющей ментальных репрезентаций, во втором – о способах концептуализации социально-культурных смыслов, извлекаемых из коммуникативных процессов (основной вопрос – как конечные алгоритмизируемые вычисления порождают бесконечный свободный семиозис?).

Если информация предшествует некоторому содержательному знанию, то невозможно создать натуралистическую модель, редуцирующую информацию к репрезентации. В информационных моделях данные (data) и со-

держание (content) вне интерпретирующей человеческой деятельности онтологически несовместимы. При этом, следуя гипотезе о смысловой природе сознания, следует отметить, что в человеческой деятельности реализуются оба плана: с одной стороны, в перспективе от первого лица присутствует содержание ментальных репрезентаций (каузированных перцепцией, внутренними нарративами и даже сновидениями), с другой – в перспективе от третьего лица очевидны структурно-функциональные принципы кодирования в символических связях.

Отличительная черта человеческого преобразующего действия – метарепрезентативный «слой» языковых способностей. Учитывая, что коммуникативные вызовы – это естественные условия становления нормальной личности, то можно говорить о языковых потребностях, зарождающихся на границе тела и среды.

Делается вывод о том, что репрезентативный подход (несмотря на успехи телесно-ориентированных концепций) все еще остается перспективным направлением когнитивных исследований, т.к. закрепляет особый онтологический статус за содержательной (семантической) стороной феноменального сознания. В данном диссертационном исследовании было приведено достаточно аргументов в пользу семантического (не-информационного) репрезентационализма, т.к. репрезентации позволяют «конвертировать» данные физической системы в феноменальное содержание.

Главным здесь является то, что одним из ключевых языковых механизмов сознания является не знаковая коммуникация, а построение проекций субъективного опыта через семантику языковых выражений; и наоборот – «примеривание» опыта Другого к собственным телесному и ментальному видам опыта. Можно утверждать, что интенциональность метарепрезентаций направлена не на поддержание непротиворечивых связей семантики и онтологии, а на создание общего представления о прагматических и семантических функциях знака. Метарепрезентация вовлекает социально-коммуникативную среду в формирование концептуальных свойств сознания

по аналогии с телесным действием. В концептуально-семантической системе языка сохраняются следы аналогистической работы сознания.

Направление телеологической семантики включает в себя преимущества информационной трактовки репрезентации с учетом специфики теории воплощенного значения. Ключевую роль вновь начинает играть дорепрезентативный язык базовых абстракций, сформированный за счет организменно-средовой системы отношений.

Концептуализация телесного опыта является содержательной основой для формирования феноменальных суждений. Телесность, фиксирующая в дорепрезентативном языке опыт обладания сознанием, является тем «видом» сознания, которое элиминируется из компьютеризации. Вычислительные подходы в силу представленных методологических ограничений не способны воспроизвести функцию перехода от обработки средовой информации к построению интероцептивного «словаря», на котором воспроизводится феноменальный опыт. В формировании «словаря» участвуют как социальные контексты, так и способность сознания к построению личностной проекции.

В *заключении* подводятся итоги исследования, формулируются основные выводы, намечаются перспективы дальнейшей разработки темы диссертации.

### **Общие результаты исследования**

В результате проведенного исследования обоснована значимость семантического ракурса для вычислительной теории сознания. Семантический подход к проблеме сознания позволил сформулировать ряд важных положений:

1. Установлено, что в ходе концептуального формирования компьютерной метафоры центральным объектом вычислительной методологии становится содержание ментальных репрезентаций.
2. Обосновано, что особую роль играет малоисследованное соотношение таксономий абстрактных дескрипций с имплицитными формами ментального содержания.

3. Раскрыты семантические свойства ментальных репрезентаций, которые представимы в виде специфичных информационных структур, но эффективно вычислимая модель которых остается недостижимой.
4. Определена роль соматоцентричной парадигмы в антикомпьютерно-рационалистской традиции. Телесность как специфичное расширение ввода/вывода средовой физической информации одновременно является концептуализатором интероцептивных ощущений, составляющих основу для феноменального опыта и феноменальных суждений.
5. Описаны причины семантических смещений в ментальных репрезентациях. Высокая степень метафоричности концептуальных схем связана с уникальностью личностного имплицитного содержания; при этом очевидно наличие прототипичных концептов, позволяющих реализовать сознанию принцип Я-проекции.
6. Выявлены методологические возможности вычислительных моделей сознания:

- Вычислительные модели обладают недостаточной объяснительной силой для целостного непротиворечивого представления содержательных свойств сознания.

- Вычислительные модели когнитивных процессов обладают мощным эвристическим потенциалом для развития искусственных интеллектуальных систем, для инженерной реализации некоторых когнитивных функций человеческого разума.

Установлено, что онтологически противоречивая методология, которая явилась следствием особых условий становления компьютерной метафоры в когнитивных науках, способна привести к эффективному решению прикладных инженерных задач.

Важным этапом в формировании компьютероцентричной парадигмы стала попытка редуцировать физические корреляты сознания к абстрактно-символьной обработке данных, генерируемых в результате информационных процессов физической системе (мозга). Принцип минимальной корреляции (физической, функциональной или информационной) не оправдывает себя, т.к. в нем не обнаруживаются решения ряда проблем: онтология вычислимого инвариантного символа, проблема аппроксимации биологических процессов, проблема алгоритмического описания эмерджентных свойств системы.

Доказана закономерность переориентации компьютерационных исследовательских программ на семантические аспекты вычислительных компонентов когнитивных процессов и феноменального сознания. Специфика семантических свойств символьных вычислительных систем состоит в том, что значение выражения является эффективно вычислимым состоянием физической системы.

Теоретический раздел концептуальной семантики представлен в качестве объединяющего проблемного поля, в котором вычислительные модели сознания пересекаются с философскими теориями значения. В рамках теории концепта происходит своего рода реабилитация вычислительного атомизма. Ментальные процессы становятся объяснимыми на уровне информационных процедур, но при этом остаются несводимыми к физическим состояниям мозгового вычислителя. Для реализации такой модели приходится либо допустить разрыв в каузальной замкнутости физического мира, либо признать наличие особых онтологических свойств семантической информации как эмерджентных компонентов вычислительной сложности живого. При таком допущении непротиворечивой становится гипотеза о формировании базового концептуального содержания в репертуаре глубинного абстрактного языка.

Разработано направление аналитической критики компьютерационализма, в рамках которой обоснованы фундаментальные различия между семантикой свойств символов в вычислительной системе и семантикой символьных абстракций ментальных состояний. Доказана правомерность синтеза в

рамках единого философского подхода оснований телеологической семантики и вычислительной теории сознания.

Основные научные результаты диссертационного исследования отражены в следующих публикациях:

**монографии**

1. *Барышников П.Н.* Миф и метафора: лингвофилософский подход: Монография. – СПб. Изд-во: «Алетейя», 2010. – 216 с.
2. *Барышников П.Н.* Методологические пределы и перспективы вычислительных моделей сознания: Монография. – Пятигорск: Изд-во Пятигорского государственного университета, 2018. – 500 с.

**в журналах, входящих в перечень рецензируемых изданий ВАК РФ:**

1. *Барышников П.Н.* Архаический миф и современное мифотворчество // Вестник РУДН. Серия «Философия». – 2006. – № 1 (11). – С. 182–190.
2. *Барышников П.Н.* Становление лингвистики как самостоятельной науки Барышников // Вестник РУДН. Серия «Философия». – 2009. – № 3 (11). – С. 50–57.
3. *Барышников П.Н.* Эвристические и мифотворческие функции метафоры в научном дискурсе // Научный журнал «В мире научных открытий». – Красноярск, 2010. – №3 (09). – Часть 4. – С. 7–12.
4. *Барышников П.Н.* Философия сознания и лингвоэпистемология в теоретико-методологическом взаимодействии // Вестник ЧелГУ. – 2011. – вып. 22. – № 30 (245). – С. 134–141.
5. *Барышников П.Н.* От теорий логоцентризма к феноменологии дискурса // Вестник РГГУ. Сер. Философия. Социология. – 2011. – № 15 / 77. – С. 217–227.
6. *Барышников П.Н.* Метафора в науке и «реликты» языкового сознания // Когнитивные исследования языка; гл. ред. серии Е.С. Кубрякова, отв. ред. Н.Н. Болдырев. Вып. VIII. Проблемы языкового сознания: мат-лы Междунар. науч. конф. 15-17 сентября 2011 г. – Тамбов, 2011. – С. 385–388.
7. *Baryshnikov P.N.* Methodological problems of the linguistic philosophy of consciousness // Вестник РУДН. Серия «Философия». – 2011. – № 3 – С. 66–75.
8. *Барышников П.Н.* Проблемные вопросы номинативных процедур архаического языкового сознания // Вестник МГУ, Сер.7 Философия. – 2012. – № 4. – С. 61–69.
9. *Барышников П.Н.* Структурная диалектика в античной философии языка // Вестник РУДН. Серия «Философия». – 2012. – № 4. – С.55–63.



10. *Барышников П.Н.* Философские проблемы информационных технологий и киберпространства // Эпистемология и философия науки. – 2012. – Т.34. – № 4. – С. 230–236.
11. *Барышников П.Н.* Структуры стабильных значений на границах языка и сознания: философский квест // Эпистемология и философия науки. – 2014. – № 2. – Т.40. – С. 243–251.
12. *Барышников П.Н.* Семантические процессы сознания: от вычислительных моделей к языковому опыту // Эпистемология и философия науки. – 2014. – №3. – Т.41. – С. 96–114.
13. *Барышников П.Н.* Морфология технологической сказки: Интернет вещей и социальные дистанции // Социология власти. – 2015. – №1 (27). – С. 37–54.
14. *Барышников П.Н.* Эволюция вычислительного разума // Когнитивные исследования языка; гл. ред. серии Н.Н. Болдырев; Вып. XXI: Язык и сознание в междисциплинарной парадигме исследования: материалы Международного конгресса по когнитивной лингвистике. 30 сентября - 2 октября 2015 г. ; отв. ред. вып. Т.А. Клепикова. – 2015. – С. 32–34.
15. *Барышников П.Н.* Эхо диалектики в зеркале Вселенной // Свободная мысль. – 2016. – №1 (1655). – С. 207–214.
16. *Барышников П.Н.* Системы коммуникации животных и «проблема Рубикона» в теориях происхождения языка // Вестник РУДН. Серия «Философия». – 2016. – №1. – С. 81–90.
17. *Барышников П.Н.* Эмерджентные свойства языковых процессов сознания // Вестник Томского государственного университета. Философия. Социология. Политология. – 2016. №2 (34). – С. 21–31
18. *Барышников П.Н.* Между телом и значением: методологические преимущества телеологической семантики // Когнитивные исследования языка; гл. ред. серии Н.Н. Болдырев; Вып. XXVII: Антропоцентрический подход в когнитивной лингвистике: сб. науч. трудов; отв. ред. вып. В.З. Демьянков. – 2016. – С. 709–715.
19. *Барышников П.Н.* Когнитивная лингвистика и философия сознания: объяснительный разрыв в онтологии языкового знака // Эпистемология и философия науки. – 2016. – №4. – Т.50. – С. 120–136.
20. *Барышников П.Н.* Феноменальное и вычисляемое в структурах сознания // Вестник РУДН. Серия Философия. – 2017. – № 2 (21). – С. 229–239.
21. *Baryshnikov P.N.* Language, brain and computation: from semiotic asymmetry to recursive rules // RUDN Journal of Philosophy. – 2018. – №22 (2). – P.168–182.
22. *Барышников П.Н.* Метафорические основания компьютеризации в когнитивных науках и философии сознания // Философия науки и техники. – 2018. – Т.24, №2. – С.46–52

**в изданиях, входящих в международные базы цитирований:**

1. *Baryshnikov P.N.* Linguistic Levels of Social Epistemology // Middle-East Journal of Scientific Research. – №14 (7), 2013. – P.925–930

2. *Baryshnikov P.N.* Information, meaning and sense in the linguistic process of consciousness // *Rifl - Rivista Italiana di Filosofia del Linguaggio*. – Vol.6, №2. – P.1–13

**другие научные работы:**

1. *Барышников П.Н.* Пределы репрезентативных моделей языка в философии сознания // *Язык и сознание в семиотическом поле компьютерной виртуальности*; под ред. П.Н. Барышникова. – Пятигорск, ПГЛУ, 2013. – 145 с.
2. *Барышников П.Н.* Кибернетический иммортализм на примере французского трансгуманизма // *Глобальное будущее 2045. Конвергентные технологии (НБИКС) и трансгуманистическая эволюция*; под ред. Д.И. Дубровского. – М.: ООО «Издательство МБА», 2013. – С. 203–228.
3. *Барышников П.Н.* Философия для лингвистов. Сознание язык реальность: горизонты смысла. М.: ЛЕНАНД, 2013. – 380 с.
4. *Барышников П.Н.* Проблема концептомоделирования в языковом сознании (на примере когнитивных метафор) // *Альманах современной науки и образования*. – Тамбов: Грамота, 2010. – №1 (32): в 2-х ч. Ч.2. – С. 26–34.
5. *Baryshnikov P.N.* Metaphorical cognition and actualization of archaic concepts in the daily discourse // *Journal of international scientific publications: language, individual and society*. – 2010. – vol. 4. – part I. – P. 152–160.
6. *Барышников П.Н.* Компьютерные метафоры в теориях сознания // *Философские проблемы информационных технологий и киберпространства: сб. ст. по рез. II Международной междисциплинарной научно-практической конференции*. – Пятигорск, ПГЛУ, 2011. – С. 25–37.
7. *Барышников П.Н.* Dual Role of Metaphor in Scientific Discourse // *Journal of international scientific publications: language, individual and society*. – 2011. – vol. 5. – part I. – P. 182–189.
8. *Барышников П.Н.* Феноменалистский эмпиризм Д. Юма и «кризис объекта» в современной лингвофилософии сознания // *Д. Юм и современная философия: сборник материалов Международной конференции. 15-17 ноября 2011 г.* – М.: Изд-во ИФ РАН, 2011. – С. 60.
9. *Baryshnikov P.N.* Computability of selfness // *Аналитическая философия: проблемы и перспективы развития в России. Тезисы пленарных докладов м секционных выступлений (29-31 мая 2012 г., С.-Петербург)*. – СПб.: Изд-во филос. ф-та СПбГУ, 2012. – С. 81–83.
10. *Барышников, П.Н.* Становление языковых процедур познания // *Сборник научн. статей «Language. Discourse. Text»*. – Ростов-на Дону: Изд-во АкадемЛит, 2012. – С. 51–54.
11. *Барышников П.Н.* Теория кибернетического иммортализма: pro et contra // *Искусственный интеллект: философия, методология, инновации. Сборник трудов VI Всероссийской междисциплинарной конференции студентов, аспирантов и молодых ученых*. г. Москва, МГТУ

- МИРЭА, 29-30 ноября 2012 г. Часть II (секции 4-6). – М: МГТУ МИРЭА. – С. 113–117.
12. *Барышников П.Н.* «Срединный путь» лингвистической теории знака // *Философия. Язык. Культура.* Вып. 4; отв. ред. Горбатова Ю.В. – СПб.: Алетейя, 2013. – С. 119–133.
  13. *Барышников П.Н.* Семантические пределы вычислительной теории сознания // *Перспективные информационные технологии (ПИТ 2013): труды Международной научно-технической конференции; под ред. С.А. Прохорова.* – Самара: Издательство Самарского научного центра РАН, 2013. – С. 455–458.
  14. *Барышников П.Н.* Интенциональность сознания и пределы семантической вычислимости // *Искусственный интеллект: философия, методология, инновации: сборник трудов VII Всероссийской междисциплинарной конференции студентов, аспирантов и молодых ученых. Часть 1. Секции I-III.* г. Москва, МГТУ МИРЭА, 13-15 ноября 2013 г.; под ред. Е.А. Никитиной. – М.: Радио и Связь, 2013. – С. 53–60.
  15. *Baryshnikov P.N.* Informational models of consciousness and systematic descriptions of linguistic processes // *Следование правилу: рассуждение, разум, рациональность; отв. ред. Е.Г. Драгалина-Черная, В.В. Долгоруков.* – СПб.: Алетейя, 2014. – С. 301–317.
  16. *Барышников, П.Н.* Типология бессмертия в теоретическом поле французского трансгуманизма // *Философские проблемы информационных технологий и киберпространства.* – 2014. – № 1. – С. 98–127.
  17. *Барышников П.Н.* Кибернетические проекции внебиологического интеллекта // *Вторые Лемовские чтения: сб. материалов Всероссийской научной конференции с международным участием памяти Станислава Лема; отв. ред. А.Ю. Нестеров.* – Самара: изд-во Самар. гос. аэрокосм. ун-та, 2014. – С. 198–206.
  18. *Барышников П.Н.* Философские проблемы семантики и методология искусственного интеллекта // *Человек в технической среде. Вып. 2. Сб. науч. статей; ред. Н.А. Ястреб.* – Вологда: Изд-во ВолГУ, 2015. – С. 18–21.
  19. *Барышников П.Н., Заварзина С.А.* Вычислительные формы фигуративной семантики: от аналогии к инференции // *Вестник Вологодского государственного университета.* – 2016. – № 1. – С. 25-37
  20. *Барышников П.Н.* Культурные коды в вычислительных моделях метафор и аналогий // *Komunikacja międzykulturowa w świetle współczesnej translatologii: przeKład jaKo dialog Kultur – Konteksty, bariery, wyzwania. Tom IV.* Olsztyn 2016. – P. 61–71.
  21. *Барышников П.Н.* «Язык мысли» в поисках новых философских форм // *Философские проблемы информационных технологий и киберпространства.* – 2016. – № 1. – vol. 11. – С. 121–126.
  22. *Барышников П.Н.* Критика компьютеризации и семантические аспекты феноменального сознания // *Философские проблемы информа-*

- ционных технологий и киберпространства. – 2016. – № 2. – vol. 12.– С. 14–31
23. *Барышников П.Н.* Машинный разум в философском и инженерном измерениях // *Философия науки и техники в России: вызовы информационных технологий: сборник научных статей / М-во образ. и науки РФ, Вологод. гос. ун-т ; под общ. ред. Н.А. Ястреб. – Вологда: ВоГУ, 2017. – С. 16–19.*
24. *Барышников П.Н.* Онтология информации и методологические границы прикладной лингвистики // *Лингвистика XXI века: традиции и новации: сб. трудов по материалам междунар. науч. конф.: памяти проф. В.В. Лазарева / редкол.: Д.А. Аксельруд и др.; Пятиг. гос. ун-т. – Пятигорск: ПГУ, 2017. – С. 36–40.*
25. *Барышников П.Н.* Проблема ментального содержания в аргументах антикомпьюционалистов // *Материалы международной конференции «Перспективные информационные технологии». – Самара: СНИУ, 2017. – С. 1135–1137.*
26. *Барышников П.Н.* Метафорические основания компьютеризации в когнитивных науках и философии сознания // *Субъект, сознание и познание в контексте современной философии когнитивных наук; под ред. Е.О. Труфановой, А.Ф. Яковлевой. – М.: Аквилон, 2017. – С.9–12*

## **Барышников Павел Николаевич**

### **Методологические возможности и границы вычислительных моделей сознания**

Диссертация посвящена исследованию методологических возможностей и границ вычислительных моделей сознания как в философском, так и инженерном ракурсах. В Вычислительной теории сознания особое место занимают исследования семантической концептуализации. Особую роль в вычислительных моделях и методологии искусственного интеллекта занимает проблема содержания функциональных состояний системы, которая не до конца изучена: имеются значительные противоречия между теориями, гипотезами и выводами. Данное исследование основывается на положении о том, что граница между информацией и значением, телесным и ментальным, вычислимым и невычислимым проходит на уровне особых свойств семантической информации. Результаты философского решения проблемы вычислимости ментальных процессов могут применяться как в фундаментальных теоретических областях, так и в прикладных системах по компьютерному моделированию интеллектуальных и когнитивных процессов.

**Pavel N. Baryshnikov**

**Methodological possibilities and limitations of computational models of mind**

The thesis is devoted to the research of methodological possibilities and limitations of computational models of mind both in philosophical and engineering foreshortenings. In the Computational Theory of Mind, a special place is occupied by the research of the semantic conceptualization. A special role in the computational models and methodology of artificial intelligence is occupied by the problem of content of the system's functional states, which has not been fully understood: there are significant contradictions between theories, hypotheses and conclusions. This research is based on the proposition that the boundary between information and meaning, physical and mental, computable and non-computable, passes at the level of specific properties of semantic information. The results of the philosophical solution of the problem of the computability of mental processes can be applied both in fundamental theoretical fields and in applied systems for computer modeling of intellectual and cognitive processes.