

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА
ФИЛОСОФСКИЙ ФАКУЛЬТЕТ

На правах рукописи

Танюшина Александра Александровна

Информационные теории сознания в аналитической философии

09.00.03 – История философии

ДИССЕРТАЦИЯ
на соискание учёной степени
кандидата философских наук

Научный руководитель:
член-корреспондент РАН
доктор философских наук, профессор
Вадим Валерьевич Васильев

Москва – 2021

Оглавление

Введение	4
Глава 1. Первые вычислительные теории сознания	20
1.1. Основные предпосылки появления вычислительных подходов к решению проблемы сознания	20
1.1.1. Математическая теория информации и развитие кибернетики	22
1.1.2. Машины Тьюринга и компьютерное моделирование интеллектуальных функций	30
1.1.3. Когнитивная революция и появление функционализма	34
1.2. Классический компьютеризм: основные идеи и представители	38
1.2.1. Теория машинных состояний Хилари Патнэма	41
1.2.2. Компьютеризм и репрезентативизм в философии Джерри Фодора	47
1.2.3. Интерпретация вычислительных процессов в работах Дэвида Марра и Зенона Пилишина	52
1.3. Нейросетевая парадигма в когнитивных науках и философии сознания ..	57
Заключение главы	62
Глава 2. Наиболее известные информационные теории сознания в аналитической философии	64
2.1. Информационный репрезентативизм Фреда Дретске	65
2.2. Информационно-вычислительная метафора в философии сознания Дэниела Деннета	74
2.3. Кибернетическая теория сознания Кеннета Сейра	86
2.4. Информационный подход к объяснению сознания Д.И. Дубровского	94
Заключение главы	104
Глава 3. Основные аргументы против информационных теорий сознания	106
3.1. Некоторые общие возражения против информационных и вычислительных моделей сознания	107
3.2. Аргумент тривиальности и вычислительной реализации	112
3.3. Вычисление и обработка информации как субъективные процессы	120
3.4. «Китайская комната» и проблема семантики	124
3.5. Критика функционализма и «провал в объяснении»	133

Заключение главы	137
Глава 4. Информационные теории сознания после «трудной проблемы»: «минимальный компьютеризм» и двуаспектный принцип информации	140
4.1. «Минимальный компьютеризм» Дэвида Чалмерса.....	141
4.2. Двуаспектный принцип информации и «трудная проблема» сознания...	147
4.3. Неограниченный двуаспектный принцип информации и панпсихизм	152
4.4. Ограниченный двуаспектный принцип информации.....	156
4.4.1. Феноменальный опыт как результат преобразования информационных структур.....	159
4.4.2. Феноменальный опыт, ассоциированный с особой каузальной деятельностью	163
4.5. «Аспектизм» и дефляционизм по отношению к осведомленности	169
4.6. Информационно-вычислительный подход Д. Чалмерса и концептуальный структурный реализм.....	178
4.7. Онтологический статус информационных пространств и информационный реализм	186
Заключение главы	192
Заключение	195
Библиография	204

Введение

Актуальность исследования

Философия сознания – одна из наиболее активно развивающихся областей современной аналитической философии. В силу своего тесного сотрудничества не только с естественными науками, но также и со многими гуманитарными дисциплинами, философия сознания активно обсуждается научным сообществом, а рассматриваемые ею проблемы получили поистине мировую огласку. Некоторые исследователи полагают, что именно благодаря развитию этой области философии современная наука сняла негласный запрет на использование термина «сознание» и принялась активно изучать природу ментальных состояний.

Пытаясь найти последовательные и непротиворечивые решения проблем, встающих на пути к решению «загадки сознания», многие аналитические философы разрабатывают теории, основанные на различных достижениях современных естественных и математических наук. Одним из наиболее популярных разделов научного знания, снабжающих философские построения своим концептуальным фундаментом, является теория информации.

Философские теории, в рамках которых ментальные состояния трактуются как обладающие информационной природой, стали особенно популярны в середине прошлого столетия, когда начали активно развиваться математическая теория коммуникации, кибернетика и науки об искусственном интеллекте. К информационному описанию природы субъективного опыта обращались многие известные философы, сыгравшие заметную роль в истории аналитической философии. Концептуальный анализ выдвигаемых ими теорий представляется весьма важным для отечественной историко-философской науки, ведь сегодня мы наблюдаем новый всплеск популярности информационных подходов к объяснению сознания не только в философии, но и в рамках современного естествознания. Как отмечает когнитивист Стивен Пинкер, *«ни один уголок науки не остался незатронутым идеей о том, что обработка информации – это*

основная деятельность мозга. Язык нейробиологии насыщен такими терминами теории информации, как “сигналы”, “коды”, “репрезентации”, “трансформации”, “обработка”»¹. Действительно, чтобы убедиться в этом, достаточно взглянуть на такие популярные сегодня теории сознания, как теория интегрированной информации итальянского нейробиолога Джулио Тонони, вычислительные концепции американского биолога Кристофа Коха и психолога Шимона Эдельмана или теория схемы внимания американского нейрочеловека Майкла Грациано. Нельзя не упомянуть популярность информационных теорий сознания и в современных компьютерных науках, уповающих на возможность создания с их помощью модели «сильного» искусственного интеллекта и различных вариантов нейрокомпьютерных интерфейсов. Цифровое моделирование когнитивных функций человека также является одной из основных задач многих прикладных дисциплин, среди которых программирование виртуальных систем, социальное прогнозирование, экономическая информатика и т.д.

Междисциплинарным изучением информационных и вычислительных аспектов работы мозга и сознания занимается большое количество исследовательских центров, среди которых Центр мозга и когнитивных наук при Массачусетском технологическом университете, Институт философии разума и познания при Национальном университете Ян Мин, Ассоциация научного изучения сознания, Центр изучения сознания в университете Аризоны, Центр изучения языка и информации при Стэнфордском университете и др. В России изучением подобных вопросов занимаются Межрегиональная ассоциация когнитивных исследований, Московский центр исследования сознания при философском факультете МГУ им. М.В. Ломоносова, отдел нейронаук НИЦ «Курчатовский институт», Совет РАН по методологии искусственного интеллекта и др. Активная деятельность этих организаций свидетельствует о важности и актуальности тем, так или иначе связанных с вопросом о роли

¹ Пинкер С. Как работает мозг. Пер. О.Ю. Семиной. М.: Кучково поле. 2017. С. 98-99.

информационных теорий в современных междисциплинарных исследованиях мозга и сознания.

Степень разработанности темы исследования

Аналитическая философия сознания является достаточно популярным направлением исследования в российской философской науке. Такие отечественные философы, как Г.У. Брутян², А.Ф. Грязнов³, Д.И. Дубровский⁴, В.А. Лекторский⁵, Н.С. Юлина⁶ и др. были одними из первых, кто начал изучение аналитической философии еще в советское время. В 2000-е годы аналитическая философия стала одной из наиболее актуальных областей исследования, о чем свидетельствует весьма обширная библиография современных отечественных философов и историков философии, работающих в этой области. За последние годы были опубликованы монографии по аналитической философии сознания В.В. Васильева⁷, Д.Б. Волкова⁸, Д.В.Иванова⁹, С.Ф. Нагумановой¹⁰, Н.С. Юлиной¹¹, а также переводы работ зарубежных философов. Анализу этих и других текстов, принадлежащих представителям аналитической философии, посвящено большое количество статей российских исследователей А.П.

² Брутян Г.У. У английских аналитиков // Вопросы философии. №2. 1968. С. 136-143.

³ Грязнов А.Ф. Эволюция философских взглядов Витгенштейна. Критический анализ. М.: Изд. Моск. Ун. 1985.

⁴ Дубровский Д.И. Психические явления и мозг. М. Наука. 1971.

⁵ Лекторский В.А. Аналитическая философия сегодня // Вопросы философии. №2. 1971.

⁶ Юлина Н.С. Аналитическая философия в XX веке // Новейшие тенденции в аналитической философии. М.: ИФ РАН.1985. С. 5-40.

⁷ Васильев В.В. Трудная проблема сознания. М.: Прогресс-традиция. 2009; Васильев В.В. Сознание и вещи: Очерк феноменалистической онтологии. М.: Книжный дом «ЛИБРОКОМ». 2014.

⁸ Волков Д.Б. Бостонский зомби: Д. Деннет и его теория сознания. М.: Книжный дом «ЛИБРОКОМ». 2012.

⁹ Иванов Д.В. Природа феноменального сознания. М.: Книжный дом «ЛИБРОКОМ». 2013.

¹⁰ Нагуманова С.Ф. Материализм и сознание: анализ дискуссии о природе сознания в современной аналитической философии. Казань. 2011.

¹¹ Юлина Н.С. Головоломки проблемы сознания. Концепция Дэниела Деннета. М. 2004.

Беседина¹², А.В. Кузнецова¹³, С.М.Левина¹⁴ Е.В. Логинова¹⁵, М.А. Секацкой¹⁶ и др.

Однако исследований, ставящих своей целью проследить историю развития какой-либо одной группы теорий сознания (в нашем случае – информационных) крайне мало. Информационные теории сознания являются, по преимуществу, объектом различных междисциплинарных исследований, основанных на идеях нейробиологии, информатики и кибернетики. Сюда можно отнести работы таких исследователей, как А.Ю. Алексеев¹⁷, П.Н.Барышников^{18,19}, Д.Э. Гаспарян²⁰, К.К. Колин²¹, В.А. Ладов²², Д.Ф.Люгер²³,

¹² Беседин А.П. Насколько силен эволюционный аргумент против эпифеноменализма? // Вестник Томского государственного университета. Философия. Социология. Политология. 2016. № 2. С. 32-40.

¹³ Кузнецов А.В. Аргумент против конститутивного панпсихизма // Вестник Томского государственного университета. Философия. Социология. Политология, № 4. 2016. С. 60-68.

¹⁴ Левин С. М. Нарративный подход: фикция и реальность // Философский журнал. 2018. Т. 11. № 3. С. 184-187.

¹⁵ Логинов Е.В. «Бонхефферовский физикализм» Дмитрия Волкова // Философия. Журнал Высшей Школы Экономики. 2019. Т.3. №3. С. 329-340.

¹⁶ Секацкая М.А. Пересадка мозга и тождество личности: альтернативная интерпретация одного мысленного эксперимента // Эпистемология и философия науки. 2014. №4 (42). С. 67-76.

¹⁷ Алексеев А.Ю., Ефимов А.Р., Финн В.К. Будущее искусственного интеллекта: тьюринговая или посттьюринговая методология? // Искусственные общества. М.: Издательство Федеральное государственное бюджетное учреждение науки Центральный экономико-математический институт РАН. 2019. Т. 14. № 4; Алексеев А.Ю. Нейрокомпьютерный тест Черчлендов // Нейрокомпьютеры: разработка, применение. М.: Издательство "Радиотехника". 2015. № 5. С. 12-13.

¹⁸ Барышников П.Н. Компьютерные метафоры в теориях сознания // Философские проблемы информационных технологий и киберпространства. 2011. № 2. С. 24-36; Барышников П.Н. Критика компьютеризации и семантические аспекты феноменального сознания // Философские проблемы информационных технологий и киберпространства. 2016. № 2 (12). С. 14-30; Барышников П.Н. Феноменальное и вычислимое в структурах сознания // Вестник Российского университета дружбы народов. Серия: Философия. 2017. Т. 21. № 2. С. 229-239.

¹⁹ В 2018 году П.Н. Барышниковым была защищена докторская диссертация на тему «Методологические возможности и границы вычислительных моделей сознания», в рамках которой автор подробно рассматривает существующие сегодня вычислительные теории сознания. В рамках данного исследования основной упор делается на анализ основных положений позиции, которую автор именует «радикальным компьютеризмом», а также так называемой «семантической проблемы», встающей перед большинством подобных теорий.

²⁰ Гаспарян Д. Э. Методология кибернетики второго порядка в применении к эпистемологическим проблемам философии сознания // Вестник Томского государственного университета. Философия. Социология. Политология. 2020. № 54. С. 5-18.

²¹ Колин К.К. Философия информации и фундаментальные проблемы современной информатики // *Alma mater* (Вестник высшей школы). 2010. №1. С. 29-35.

²² Ладов В.А. Плюрализм философских интерпретаций принципов разумной деятельности в контексте исследований в области искусственного интеллекта // Вестник Томского государственного университета. Философия. Социология. Политология. 2007. № 395. С. 29-34.

²³ Люгер Д.Ф. Искусственный интеллект: стратегии и методы решения сложных проблем. М.: Издательский дом «Вильямс». 2003.

И.Ф. Михайлов²⁴, А.И. Ракитов²⁵, В.Г. Редько²⁶, Г.Н. Раппопорт и А.Г. Герц²⁷, М.А.Сущин²⁸, А.Д. Урсул²⁹ и др. (Стоит отметить, что рамках данных работ по преимуществу анализируются классические вычислительные когнитивные теории, в то время как информационные концепции, развиваемые современными аналитическими философами остаются практически незатронутыми). В 2010 году был проведен «круглый стол» по теме «Информационный подход в междисциплинарной перспективе»³⁰, в котором приняли участие некоторые из упомянутых выше отечественных философов (материалы данного «круглого стола» будут упоминаться в настоящем исследовании). Наиболее основательно информационные теории сознания разобраны в трудах А.А.Братко, А.Н.Кочергина³¹ и Д.И.Дубровского³², в которых, однако, авторы уделяют больше внимания развитию собственных информационных концепций.

В англоязычной литературе исследования информационных подходов к решению философской проблемы сознания являются более распространенными. Большинство из них представлено в виде статей, посвященных либо обзору существующих информационных теорий, либо их критическому анализу (в рамках настоящего исследования будут упомянуты статьи Фрэнка Адамса³³,

²⁴ Михайлов И.Ф. Концепции вычислений в современных науках о человеческом познании // Философские проблемы информационных технологий и киберпространства. 2018. № 1(14). С. 4–22.

²⁵ Ракитов А.И. Философия компьютерной революции. М.: Политиздат. 1991.

²⁶ Редько В.Г. Эволюция, нейронные сети, интеллект: модели и концепции эволюционной кибернетики. М.: Книжный Дом "ЛИБРОКОМ". 2013.

²⁷ Раппопорт Г.Н., Герц А.Г. Искусственный и биологический интеллект. Общность, общность структуры, эволюция и моделирование процессов познания. М.: КомКнига. 2005.

²⁸ Сущин М.А. Байесовский разум: новая перспектива в когнитивной науке // Вопросы философии. 2017. №3. С. 77-88

²⁹ Урсул А.Д. Природа информации. М.: Политиздат, 1968; Урсул А.Д. Проблема информации в современной науке. М.: Наука. 1975.

³⁰ Бодякин В.И., Дубровский Д.И., Колин К.К., Лекторский В.А., Мелик-Гайказян И.В., Пружинин Б.И., Урсул А.Д. Информационный подход в междисциплинарной перспективе. (материалы «круглого стола») // Вопросы философии. 2010. №2. С. 84-113.

³¹ Братко А. А, Кочергин А.Н. Информация и психика. М.: Наука. 1977.

³² Дубровский Д. И. Психические явления и мозг: философский анализ проблемы в связи с актуальными задачами нейрофизиологии, психологии и кибернетики. М., 1971; Дубровский Д. И. Информация, сознание, мозг. М., 1980; Дубровский Д. И. Проблема идеального. М.,1983; Дубровский Д. И. Обман. Философско-психологический анализ. М., 1994; Дубровский Д. И. Сознание, мозг, искусственный интеллект. М.: 2007.

³³ Adams F. The Informational Turn in Philosophy // Minds and Machines. №13. 2003. P. 471-501.

Тима Шаллис³⁴, Джека Коупланда³⁵, Марка Спревака³⁶, Михаэля Рескорла³⁷ и др.³⁸). В трудах данных исследователей разбираются работы аналитических философов, которые можно условно поделить на несколько групп.

Во-первых, это работы представителей так называемого «классического компьютеризма», которые одними из первых начали использовать вычислительный подход в качестве инструмента для описания работы мозга и сознания. Среди этих исследователей такие выдающиеся философы как Хилари Патнэм³⁹, Джерри Фодор⁴⁰, Зенон Пилишин⁴¹ и др. Именно первые вычислительные концепции оказали существенное влияние на становление и развитие всех последующих информационных теорий сознания. Несмотря на то, что идеи классического компьютеризма сегодня признаются большинством исследователей неактуальными, вычислительные подходы продолжают развиваться в рамках коннекционистских и постконнекционистских моделей, которые представлены в работах таких знаменитых философов и когнитивистов, как Патриция и Пол Черчленды⁴², Стивен Пинкер⁴³, Джеффри Хинтон⁴⁴ и многих других.

³⁴ Shallice T. Information-processing models of consciousness: Possibilities and problems // *Consciousness in Contemporary Science* (eds. J. Anthony, M. Bisiach). Oxford University Press, 1988.

³⁵ Copeland J. What is Computation? // *Synthese*, 1996. №108. P. 335–359.

³⁶ Sprevak M. Triviality Arguments About Computational Implementation // *The Routledge Handbook of The Computational Mind* (ed. by M. Sprevak, M. Colombo). NY: Taylor&francis Group, 2019. P. 175-192.

³⁷ Rescorla M. A Theory of Computational Implementation // *Synthese*, 2014. №191. P. 1277–1307; Rescorla M. The Computational Theory of Mind // *The Stanford Encyclopedia of Philosophy* (N. Zalta ed.). 2019. [Электронный ресурс]. URL: <https://plato.stanford.edu/entries/computational-mind/#AltConCom> (дата обращения: 04.12.2020).

³⁸ Daiseanu F. Informational Model of Consciousness: From Philosophic Concepts to an Information Science of Consciousness // *Philosophy Study*, №9(4), 2019. P. 181-196; Gamez D. Are Information or Data Patterns Correlated with Consciousness? // *Topoi* №35(1), 2016. P. 225-239; Hardcastle V.G. A Critique of Information Processing Theories of Consciousness // *Minds and Machines*, №5(1), 1995. P. 89-107; Piccinini G., Scarantino A. Computation vs. Information processing: Why their Difference Matters to Cognitive Science // *Studies in History and Philosophy of Science*. 2010. №41. P. 237–246; Pockett S. Problems with theories that equate consciousness with information or information processing // *Frontiers in Systems Neuroscience*, №8(255), 2014.

³⁹ Putnam. H. *Mind, Language, and Reality* // *Philosophical Papers*, №2. Cambridge: Cambridge University Press, 1975; Putnam. H. *Representation and Reality*, Cambridge, MA: MIT Press, 1988.

⁴⁰ Fodor J. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT/Bradford, 1987; Fodor J. *Information and Representation* // *Information, Language, and Cognition* (ed. P. Hanson). Vancouver: University of British Columbia Press, 1990. P. 175–190.

⁴¹ Pylyshyn Z. *Computation and Cognition. Toward a Foundation for Cognitive Science*. Cambridge, Massachusetts: Bradford Books/MIT Press, 1984.

⁴² Churchland P.S. *Neurophilosophy*, Cambridge: MIT Press, 1986; Churchland P.M., Sejnowski T.J. *The Computational Brain*. Cambridge, MA: MIT Press, 1992.

⁴³ Pinker S. *How the Mind Works*. New York: Norton. 1997.

⁴⁴ Hinton G.E., McClelland J.L., Rumelhart D.E. Distributed representations // *Parallel Distributed Processing* (Rumelhart D.E., McClelland J.L. & PDP Research Group eds.). Cambridge, MA: MIT Press, 1986. P. 77-109.

В следующую группу входят работы философов, начавших создавать свои теории практически параллельно с развитием классического компьютеризма: данные философы либо полностью заменили распространенное в компьютеризме понятие «вычисление» (англ. *computation*) на выражение «обработка информации» (англ. *information-processing*), либо стали использовать данные понятия как взаимозаменяемые (в силу того, что информационное объяснение природы сознания играет в их концепциях центральную роль, в современной философской литературе именно теории этих философов чаще всего называют «информационными»). Сюда можно отнести работы Фреда Дретске, создавшего теорию информационного репрезентативизма⁴⁵, Дэниела Деннета, активно использующего информационно-вычислительную метафору в своих теоретических построениях⁴⁶, Кеннета Сейра, который один из первых выдвинул теорию информационного нейтрального монизма⁴⁷ и др.

Еще один класс текстов принадлежит аналитическим философам, чье обращение к информационному концептуальному аппарату во многом обусловлено выходом в свет работ Дэвида Чалмерса, в рамках которых австралийский философ представил структурно-каузальную трактовку вычислительных процессов, а также двуаспектный принцип информации, чье применение, с его точки зрения, может способствовать решению т.н. «трудной проблемы» сознания⁴⁸. Достаточно большое количество современных исследователей проявило особый интерес к идеям Чалмерса, что послужило росту числа статей и монографий, посвященных их осмыслению и развитию.

⁴⁵ Dretske F. Knowledge and the Flow of Information. MA: Massachusetts Institute of Technology, 1981; Dretske F. Naturalizing the Mind. MA: Massachusetts Institute of Technology, 1995.

⁴⁶ Dennett D.C. Content and consciousness. N.Y., 1969; Dennett D.C. Consciousness explained. Boston, MA, 1991; Dennett D.C. From Bacteria to Bach and Back: The Evolution of Minds. NY: W.W. Norton Company, 2017.

⁴⁷ Sayre K.M. Cybernetics and the Philosophy of Mind. Routledge and Kegan Paul, 1976.

⁴⁸ Chalmers D.J. Computing the Thinkable // Behavioral and Brain Sciences, 1990. №13. P. 658-659; Chalmers D.J. On Implementing a Computation // Minds and Machines. 1995. №4. P. 391-402; Chalmers D. The Consciousness Mind: In Search of a Fundamental Theory. NY: Oxford University Press, 1996; Chalmers D.J. Computational Foundation for the Study of Cognition // The Journal of Cognitive Science, №12, 2011. P. 323-357; Chalmers D.J. The Varieties of Computation: A Reply // The Journal of Cognitive Science. 2012. №13. P. 213-248.

Среди наиболее примечательных отметим работы Вильяма Сигера⁴⁹, Макса Вэлманса⁵⁰, Сэма Колмана⁵¹ и Грегга Розенберга⁵², выдвигающих теории, в рамках которых информационный концептуальный аппарат служит основанием различных версий панпсихистских и двуаспектных теорий.

Анализ работ вышеупомянутых философов составляет основу настоящего диссертационного исследования.

Также стоит выделить еще две достаточно большие группы зарубежных исследований, в рамках которых так или иначе рассматриваются информационные теории сознания, авторство которых, однако, не принадлежит представителям философии сознания. В первую группу входят работы исследователей из таких областей философского знания, как философия информации, философия искусственного интеллекта, философия науки и т.д. В настоящей диссертации будут упоминаться труды таких философов, как Лили Кей⁵³, Лучано Флориди⁵⁴, Уильям Хармс⁵⁵, Маргарет Боден⁵⁶ и Филлип Брей⁵⁷: их работы оказали заметное влияние на ход дискуссий, затрагивающих вопросы о правомерности применения информационного концептуального аппарата в целях описания природы сознания в рамках современных философских и когнитивных наук.

Вторая, еще более обширная группа текстов принадлежит исследователям, работающим в сферах, отличных от философии. Спектр этих дисциплин весьма

⁴⁹ Seager W.E. Consciousness, Information and Panpsychism // *Journal of Consciousness Studies*, №2(3), 1996. P. 272-88.

⁵⁰ Velmans M. Is human information processing conscious? // *Behavioral and Brain Sciences*. 1991. №14(4). P. 651-669; Velmans M. Reflexive monism // *Journal of Consciousness Studies*, 2008. №15(2). P. 5-50.

⁵¹ Coleman S. The Real Combination Problem: Panpsychism, Micro-Subjects, and Emergence // *Erkenntnis*, 2013. №(1). P. 1-26.

⁵² Rosenberg G. A Place for Consciousness. NY: Oxford University Press, 2004; Rosenberg G.H. Causality and the combination problem // *Consciousness in the Physical World: Perspectives on Russellian Monism* (T. Alter, Y. Nagasawa, eds.). Oxford University Press. 2014. P. 224-245.

⁵³ Kay L. *Who Wrote the Book of Life? A History of the Genetic Code*. Stanford University Press. 2000.

⁵⁴ Floridi L. Open Problems in the Philosophy of Information. Metaphilosophy LLC & Blackwell Publishing Ltd. 2004; Floridi I. Informational Realism // *Conferences in Research and Practice in Information Technology* (Eds. J. Weckert, Y. Al-Saggaf). № 37. 2005; Floridi I. *The Philosophy of Information*. Oxford University Press, 2011.

⁵⁵ Harms W.F. The use of information Theory in Epistemology. *Philos. Sci.*, №65, 1998. P. 472-501.

⁵⁶ Boden M.A. *Mind as Machine: A History of Cognitive Science*. Oxford University Press. 2006; Boden M.A. Information, Computation and Cognitive Science // *Handbook of the Philosophy of Science. Vol.8: Philosophy of Information* (ed. by P. Adrians, J. van Benthem). Amsterdam: Elsevier. 2008. P. 741-763.

⁵⁷ Brey P., Soreker J. Philosophy of Computing and Information Technology // *Handbook for Philosophy of Science*. (Ed. A. Meijers, gen. ed. D. Gabbay, P. Thagard, J. Woods). №14. 2009.

широк, хотя большинство из них относится к нейрокогнитивным наукам и биологии. Данные работы нередко оказывают значительное влияние на аналитических философов сознания, что подтверждается многочисленными ссылками последних на современные естественнонаучные исследования. По этой причине в рамках настоящей диссертации подобные работы также будут иногда упоминаться в качестве дополнительного источника теоретического и эмпирического материала⁵⁸.

Цель исследования

Целью исследования является концептуальный анализ и сопоставление наиболее известных информационных теорий сознания, существующих в аналитической философии, а также разбор ключевых проблем, встающих у них на пути к решению загадки сознания.

Задачи исследования

Достижение цели предполагает выполнение следующих задач:

1. Выявить историко-философские и научные предпосылки возникновения первых вычислительных теорий сознания.
2. Разобрать основные идеи классического компьютеризма, представленные в работах Хилари Патнэма, Джерри Фодора, Дэвида Марра и Зенона Пилишина; указать на их ключевые концептуальные проблемы.
3. Рассмотреть стратегии решения вопроса о природе сознания с помощью информационного подхода на примере концепций Фреда Дретске, Дэниела Деннета, Д.И. Дубровского и Кеннета Сейра.

⁵⁸ Помимо уже упомянутых работ Тонони, Эдельмана, Коха и Грациано отметим: Biehl M., Chang A.Y.C., Kanai R., Yu Y. Information Closure Theory of Consciousness // *Front Psychol.* 2020. №11. P. 1504-1522; Davidson R.J. Consciousness and Information Processing: A Biocognitive Perspective // *The Psychology of Consciousness*. Ed. R. Davidson et al. NY. 1980; Deacon T.W. *Incomplete Nature: How Mind Emerged from Matter*. New York: W.W. Norton & Company. 2011; Navon D. The Function of Consciousness or of Information? // *Behavioral and Brain Sciences*, №14(4), 1991. P. 579-594; P. 207-210; Jansen F.K. The Hard Problem of Consciousness from Bio-Psychological Perspective // *Philosophy Study*, №7(11), 2017. P. 579-594; Gabora L. Amplifying Phenomenal Information: Toward a Fundamental Theory of Consciousness // *Journal of Consciousness Studies*, 2002. №9(8). P. 3-29.

4. Разобрать основные группы возражений против информационных теорий сознания.

5. Обозначить ключевые положения информационно-вычислительного подхода к решению «трудной проблемы» сознания Дэвида Чалмерса; выявить и оценить основные преимущества предлагаемого философом подхода перед другими популярными информационными теориями сознания.

6. Указать возможные пути развития данного подхода, используя теоретические наработки современных философов и ученых.

Представленные выше задачи соотносятся со структурой диссертации, в первой главе которой представлены предпосылки возникновения и история развития первых вычислительных концепций сознания, во второй анализируются ключевые идеи наиболее известных информационных теорий сознания, развиваемых аналитическими философами с 1960-х г.г., в третьей разобраны основные группы аргументов против данных теорий, а в четвертой описаны предлагаемые современными философами стратегии решения проблем, характерных для информационных теорий сознания.

Объект исследования

Объектом исследования являются монографии, статьи, сборники и другие материалы, посвященные информационным теориям сознания, таких аналитических философов, как Х.Патнэм, Дж. Фодор, З. Пилишин, Ф. Дретске, Д. Деннет, К. Сейр, Д.И. Дубровский, Д.Чалмерс, Г. Розенберг, Р. Мали и др.

Предмет исследования

Предметом исследования являются информационные теории сознания в аналитической философии.

Научная новизна исследования состоит в следующих пунктах:

1. Впервые в отечественной философской литературе подробно разобрана кибернетическая концепция сознания американского мыслителя Кеннета Сейра,

чьи труды до сих пор оставались практически неизвестными российским исследователям.

2. Обоснованы различия между информационными концепциями Фреда Дретске, Дэниела Деннета, Кеннета Сейра, Д.И. Дубровского и вычислительными подходами к объяснению сознания, разрабатываемыми в когнитивных науках.

3. Разобраны основные положения модели «минимального компьютеризма» Дэвида Чалмерса, а также обоснован анализ информационно-вычислительного подхода философа к решению «трудной проблемы» сознания в контексте его структуралистской философии и эпистемологии.

4. Выявлена специфика двуаспектной онтологии, рассмотрен так называемый «ограниченный» двуаспектный принцип информации и обозначены возможные пути его развития.

Теоретическое и практическое значение диссертации

Настоящее исследование позволит более прицельно взглянуть на особенности современных информационных теорий сознания и сформулировать основные требования, необходимые как для их дальнейшего концептуального развития, так и для плодотворного взаимодействия аналитической философии сознания с другими дисциплинами, изучающими смежные вопросы. Теоретическая значимость настоящей диссертационной работы состоит не только в исследовании достаточно малоизученной в отечественной литературе проблематики, связанной с философской интерпретацией информационных теорий сознания, но также и в потенциальной применимости ее результатов в таких областях теоретического знания, как онтология, эпистемология, когнитивная лингвистика, теория искусственного интеллекта и др.

Последовательный концептуальный анализ современных информационных теорий сознания способен привнести вклад в развитие таких прикладных направлений научной деятельности, как нейросетевое и

нейродинамическое моделирование интеллектуальных функций, роботехническая и компьютерная инженерия, программирование, создание систем виртуальной реальности и т.д.

Материалы исследования также могут быть использованы в практике преподавания учебных курсов по истории философии, философии сознания, философии информации, философии науки, теории искусственного интеллекта, истории когнитивных наук и др.

Методологическая основа исследования

Рассматриваемая в рамках данной работы тема является весьма непростой с точки зрения своей проблематики: она необходимым образом включает в себя изучение культурно-исторического контекста, в котором развивались первые информационные теории сознания, разбор основных положений данных теорий, а также их сопоставление со смежными теоретическими построениями, существующими в нефилософских дисциплинах.

В связи с этим настоящая диссертационная работа строится на методологии историко-философского анализа и предполагает изучение информационных теорий сознания в контексте аналитической философии второй половины XX – начала XXI в.в. Важная роль в исследовании также отводится анализу основных источников, среди которых статьи и монографии философов, разрабатывающих информационные теории сознания; по этой причине базовым методом настоящей диссертационной работы также является источниковедческий анализ. Кроме того, автором разбирается достаточно широкий круг комментаторской литературы, в рамках которой отражены различные точки зрения исследователей по вопросам, касающимся темы диссертации.

Наконец, особое место в исследовании занимают концептуальный анализ ключевых идей информационных теорий сознания и компаративистский анализ, подразумевающий сравнение данных теорий между собой, а также с другими

концепциями сознания, разрабатываемыми в рамках аналитической философии и когнитивной науки.

Положения, выносимые на защиту

1. Первые информационные теории сознания появились в ходе так называемой «когнитивной революции» 1950-х–1960-х г.г. в результате кризиса бихевиоризма и типовой теории тождества. Тем не менее, сторонники информационных теорий не отказались от ключевой для упомянутых выше философских программ задачи, выраженной в необходимости анализа ментальных состояний в тематически нейтральных категориях. Роль подобных категорий стали выполнять такие понятия, как «вычисление», «обработка информации», «код», «символ», «сигнал» и т.д., активно используемые в когнитивных науках до сих пор.
2. Такие философы, как Фред Дретске, Дэниел Деннет, Кеннет Сейр и Д.И. Дубровский создали оригинальные информационные теории сознания, которые объединяет не только единая ориентация на интеграцию философских построений с достижениями современных эмпирических и математических наук, но и желание сократить концептуальный разрыв между естественнонаучным и философским объяснением сознания посредством использования категории информации. Несмотря на это, ни одна из этих теорий не содержит удовлетворительного решения проблемы, связанной с объяснением природы сознания.
3. Перед информационными теориями сознания встает ряд сложных проблем, которые можно условно разделить на две группы: проблемы, связанные с концептуальным разьяснением и интерпретацией основополагающих для подобных теорий понятий и категорий (проблема соотношения понятий «информация», «информационная обработка» и «вычисление», проблема

тривиальности и вычислительной реализации, проблема синтаксиса и семантики и др.); проблемы, касающиеся описания связи информационных процессов с феноменальными переживаниями.

4. Дэвид Чалмерс представил основу потенциальной теории сознания, опирающуюся на идеи «минимального компьютеризма», двухаспектного принципа информации и концептуального структурного реализма. Использование подобных подходов и принципов способно пролить свет на некоторые проблемы, связанные с трактовкой информационных и вычислительных состояний, среди которых проблемы тривиальности, вычислительной реализации и др.

5. Информационная интерпретация природы сознания является онтологически нейтральной и может быть потенциально совмещена с различными популярными сегодня метафизическими концепциями. Однако, в силу того, что большинство данных концепций сталкивается с рядом характерных трудностей, говорить о создании удовлетворительной информационной теории сознания пока рано.

Степень достоверности результатов исследования и апробация диссертации

Настоящая диссертация опирается на достаточно широкую базу первоисточников и комментаторской литературы, при анализе которой соблюдаются принципы объективности и непредвзятости: автором выявляются особенности концептуального языка рассматриваемого им круга философов и ученых, сопоставляются разные точки зрения на исследуемую проблему и обосновывается собственная позиция, подкрепляемая ссылками на первоисточники или цитатами из них.

Положения настоящего диссертационного исследования отражены в 5 публикациях автора, отвечающих требованиям п. 2.3 Положения о присуждении

ученых степеней в Московском государственном университете имени М.В. Ломоносова:

- 1) Танюшина А.А. Философские проблемы теории интегрированной информации Д. Тонони // Вестник Московского государственного университета. Серия 7. Философия. 2020. №1. С. 13-31;
- 2) Танюшина А.А. Нейтральный монизм и информационный реализм в философии сознания К. Сейра // Вестник Воронежского государственного университета. 2020. № 1. С. 84-91;
- 3) Танюшина А.А. Метафизика сознания и информационный реализм: проблема дальнейшего развития двухаспектного принципа информации // Философия науки. 2020. № 4. С. 71-82;
- 4) Танюшина А.А. Компутационализм, информация и сознание: на пути к решению «трудной проблемы» // Вестник Томского государственного университета. Философия. Социология. Политология. 2021. № 59. С. 42-50.
- 5) Танюшина А.А. Двухаспектизм в современной аналитической философии сознания: основные трудности и перспективы развития теории // Вестник Московского государственного университета. Серия 7. Философия. 2021. №2. С. 23-33.

Основные положения и выводы диссертационного исследования были представлены автором в выступлениях на следующих научных конференциях:

- 1) *Нерешенные проблемы в философии информации*. Ломоносов-2019: XXVI Международная конференция студентов, аспирантов и молодых ученых. Москва, 11 – 13 апреля 2019, устный доклад 11 апреля 2019 года.
- 2) *«Игра знаний» Лучано Флориди*. Международная молодежная междисциплинарная конференция «Философия искусственного интеллекта». Москва, 30 мая - 20 июня 2019, устный доклад 31 мая 2019 года.
- 3) *Что такое «ограниченный двухаспектный принцип информации» Д. Чалмерса?* Ломоносов-2020: XXVII Международная конференция студентов,

аспирантов и молодых учёных «Ломоносов». Москва, 10-27 ноября 2020, устный доклад 12 ноября 2020 года.

4) *Перспективы развития информационных подходов к решению проблемы сознания*. Второй международный конгресс Русского общества истории и философии науки «Наука как общественное благо». Санкт-Петербург, 27-29 ноября 2020, устный доклад 28 ноября 2020 года.

5) *«Трудная проблема» информационно-вычислительных теорий сознания*. Философская мастерская. Седьмая встреча в рамках цикла "Лаборатория" (ИФ РАН). Москва, устный доклад 3 марта 2021 года.

Диссертационная работа была обсуждена на заседании кафедры истории зарубежной философии философского факультета Московского государственного университета имени М.В. Ломоносова 31 мая 2021 года и рекомендована к защите.

Структура исследования.

Диссертационная работа состоит из введения, четырех глав, разделенных на параграфы, заключения и библиографии.

Глава 1. Первые вычислительные теории сознания

1.1. Основные предпосылки появления вычислительных подходов к решению проблемы сознания

Возникновение вычислительных теорий сознания⁵⁹ в аналитической философии является закономерным следствием целого ряда факторов, связанных как с общим развитием философской мысли в XX веке, так и со становлением новых областей научного знания, оказавших непосредственное влияние на подобные теории.

Говоря о культурном и философском контекстах, в которых сформировались интересующие нас теории сознания, достаточно сложно строго очертить круг идей, повлиявших на их становление: многие положения данных теорий являются своеобразным развитием как идей классической философии, выраженных в трудах Лейбница, Юма, Канта и других великих мыслителей, так и подходов, появившихся лишь в начале XX века. Так, идеи ранней аналитической философии, представленные в работах таких философов и логиков, как Готлоб Фреге, Людвиг Витгенштейн, Джордж Эдвард Мур, Бертран

⁵⁹ Несмотря на то, что тема настоящего диссертационного исследования звучит как «Информационные теории сознания в аналитической философии», в данной работе будут встречаться такие выражения, как «вычислительные теории», «информационно-вычислительные теории», а также такое понятие как «компьютериализм» (от англ. «computation», вычисление). В силу того, что во многих современных философских трудах вышеперечисленные понятия используются как взаимозаменяемые, достаточно сложно говорить о наличии четкой границы между «компьютериалистскими», «вычислительными» и «информационными» теориями в аналитической философии сознания: если критерии приверженности того или иного мыслителя к положениям «классического компьютеризма» описаны более-менее четко, то очертить круг философов, являющихся сторонниками некоторых иных версий вычислительного подхода, представляется крайне непростой задачей во многом в силу абстрактности и метафоричности используемых ими объяснительных стратегий и аналогий.

Чтобы избежать терминологической путаницы, в настоящей диссертации понятие «компьютериализм» будет применяться исключительно по отношению к теориям философов, являющихся представителями «классического» (символического) вычислительного подхода (таких как Хилари Патнэм, Джерри Фодор, Зенон Пилишин и др.), а также по отношению к концепциям, авторы которых предпочитают именовать их таким образом (как в случае с «минимальным компьютеризмом» Дэвида Чалмерса).

Далее в настоящем исследовании под «информационными теориями сознания» будет подразумеваться совокупность философских теорий, в рамках которых ментальные состояния трактуются как обладающие информационной природой. Вопрос о соотношении «вычислительных» и «информационных» состояний является одним из центральных для настоящего исследования. В большинстве случаев, когда концептуальное разграничение понятий «вычисление» и «информация»/«обработка информации» не является необходимым, выражения «информационный» и «вычислительный» будут использоваться как синонимичные; в ряде случаев, когда потребуется концептуальное разграничение понятий «вычисление» и «обработка информации», об этом будет сказано отдельно.

Рассел, Рудольф Карнап и др., обеспечили интеграцию методов математической логики в философский инструментарий и заложили основы лингвистического анализа, что, как известно, оказало значительное влияние на развитие аналитической философской методологии. Взаимодействие положений позитивистски-ориентированного эмпиризма с различными версиями феноменологического подхода, активно развивавшегося в то время на территории континентальной Европы, привело к осознанию необходимости поиска тематически нейтрального терминологического аппарата, способного пролить свет на решение ряда философских проблем, связанных с построением научной психологии и трактовкой понятия «сознание». Уже начиная с середины XX века ученые и философы пришли к выводу, что многие предложенные ими методы являются непригодными для создания удовлетворительных теорий сознания: кризис лингвистического подхода, выраженный в его неспособности очистить философию от метафизических проблем, способствовал своеобразному обновлению философской и научной исследовательских программ, что, в свою очередь, сделало вновь актуальным поиск релевантной объяснительной стратегии и подходящего ей терминологического аппарата.

Параллельно с постепенным упадком прежних философских подходов и методологий происходило формирование ряда естественнонаучных и математических дисциплин: развитие математической логики и инженерии сказалось на становлении таких новых областей научного знания, как теория информации, теория вычислительных систем и кибернетика. Активное использование представителями данных наук нового информационного концептуального аппарата, обеспечившего в середине прошлого века существенный прирост теоретического знания, привлек к нему внимание аналитических философов, что и послужило основной причиной формирования информационных подходов к интерпретации природы сознания.

Далее будут более подробно рассмотрены такие три ключевые предпосылки развития подобных подходов, как становление теории информации, появление теории вычислительных систем, а также так называемая

«когнитивная революция», подразумевавшая, в частности, отказ от бихевиористской парадигмы и поиск новых научных и философских концепций сознания.

1.1.1. Математическая теория информации и развитие кибернетики

История понятия «информация» является достаточно непростой: после появления в качестве своеобразного технического термина в античных и средневековых текстах оно на какое-то время практически исчезло из философского дискурса, но стало достаточно активно использоваться в разговорной речи, вследствие чего приобрело статус абстрактного понятия, значение которого трактуется неоднозначно и напрямую зависит от контекста. Некоторые историки философии полагают, что сама идея информации, вероятно, играет важную роль во многих древних и классических философских учениях⁶⁰. В Стэнфордской философской энциклопедии можно найти следующее рассуждение голландского ученого Питера Адрианса:

«В ходе исторического анализа можно было бы спросить, имеет ли понятие “informatio”, обсуждаемое Августином, какую-либо связь с информацией Шеннона, кроме непосредственного сходства терминов. В то же время можно задаться вопросом, является ли “исторический, ясный метод” Локка важным вкладом в возникновение современной концепции информации, хотя в своих работах Локк вряд ли использует термин “информация” в техническом смысле. Существует целый конгломерат идей, связанных с понятием информации, который развивался с древности до недавнего времени, но нам все еще

⁶⁰ Так, в 2021 году вышел сборник статей под редакцией Криса Мейнса, представляющий собой первое обширное исследование истории философских идей, так или иначе связанных с концепцией информации (Information and the History of Philosophy (ed. by C. Meyns). Routledge. 2021).

необходимо дальнейшее изучение истории понятия информации»⁶¹.

Адрианс считает, что можно выделить по меньшей мере три различных исторических значения понятия «информация»:

1. «Информация» как процесс получения знания.

Это самый древний смысл, который можно найти еще в трудах перипатетиков, утверждающих, что при познании некоторого объекта, его «форма» закладывается в уме познающего; в том же смысле можно было бы сказать, что скульптор создает статую, «сообщая» куску мрамора «информацию» о статуе⁶². Позднее христианский богослов Аврелий Августин в своем трактате «О Троице» будет анализировать зрение как имеющее три аспекта: физическое воплощение предмета во внешнем мире, информация о нем, получаемая зрением, и результирующая форма в уме⁶³. Как отмечает Адрианс, этот смысл довольно долго сохранялся в западноевропейском дискурсе⁶⁴.

2. «Информация» как состояние агента.

Переход к этому современному значению, по-видимому, происходил постепенно, начиная с середины XV века: так, в эпоху Возрождения ученого можно было назвать «человеком информации».

⁶¹ Adriaans P. Information // The Stanford Encyclopedia of Philosophy (N. Zalta ed.), 2019. [Электронный ресурс]. URL: <https://plato.stanford.edu/archives/spr2019/entries/information/> (дата обращения: 09.05.2020).

⁶² Более подробное описание генезиса слова «информация», а также исследование применения данного понятия в античных философских текстах представлено в работах историка философии Рафаэля Капурро (Capurro R., Birger H. The Concept of Information // Annual Review of Information Science and Technology (ed. by V. Cronin). 2003. №37. P. 343–411).

⁶³ Августин пишет: «[...] само ощущение образуется (formatur) так, что оно уже не есть просто ощущение, которое может быть невидимым даже и во тьме, пока неведимы глаза, но ощущение воображенное (informatus), каковое называется видением. Следовательно, видение порождается видимым предметом, но не им одним, а также и видящим. Поэтому видение порождается видимым и видящим так, что от видящего (в нем) – ощущение зрения и внимания взора и созерцания, воображение же ощущения (informatio sensus), которое называется видением, отпечатывается лишь телом, которое видится, т. е. каким-либо видимым предметом» (Августин. О Троице. Пер. А.А. Тащиана. Краснодар: Издательство ГЛАГОЛ. 2004. С. 241).

⁶⁴ Так, еще Джордж Беркли в диалоге «Алкифрон, или Мелкий философ» пишет: «I love information upon all subjects that come in my way, and especially upon those that are most important» (Berkeley G. Alciphron, or the Minute Philosopher in Focus. (Berman D., ed.). London: Routledge, 1997. P. 30). В русскоязычном переводе, однако, понятие «информация» не фигурирует: «Я люблю узнавать обо всем, что встречаю на своем пути, в особенности о тех предметах, которые принадлежат к числу самых важных» (Беркли Дж. Алкифрон, или Мелкий философ. Работы разных лет. Перевод с английского, составление и комментарии А. А. Васильева. СПб.: Алетейя, 2000. С. 19.).

3. «Информация» как предрасположенность к информированию (способность объекта информировать агента).

Подобное представление основано на следующем рассуждении: так как акт обучения, например, теореме Гёделя, дает человеку информацию об этой теореме, то вполне естественно предположить, что текст, в котором изложена данная теорема, действительно «содержит» некую информацию, то есть обладает способностью информировать; таким образом, под «информацией» подразумевается нечто, что можно хранить и измерять.

Последняя интерпретация понятия «информация» получила наиболее широкое распространение, в то время как связь с такими абстрактными философскими понятиями, как «форма», перестала играть существенную роль, и предыдущие значения исчезли из массового употребления. В этом новом виде понятие было воспринято исследователями в XX веке, когда были впервые описаны формальные методы измерения информации.

Принято считать, что последовательная теория информации появилась в середине прошлого века. Так, в 1948 году было опубликовано знаменитое сочинение американского математика и инженера Клода Шеннона «Математическая теория связи»⁶⁵, описывающее совокупность математических формул, с помощью которых возможно устанавливать значение количества информации, транслируемой по каналам связи (отметим, что еще до 1948 года информационная теория получила некоторое развитие в рамках различных областей научного знания: ее пионерами также можно назвать венгеро-американского математика Джона фон Неймана, активно разрабатывавшего инструменты теории вероятности, американского ученого Ральфа Хартли, определившего «меру информации» с помощью логарифмической формулы, шведского ученого Гарри Найквиста и некоторых других известных исследователей). Согласно теории Шеннона, источником информации может быть любой физический объект, порождающий сигналы-сообщения, которые

⁶⁵ Shannon C. E. A Mathematical Theory of Communication // Bell System Technical Journal, № 27, 1948. P. 379—423; 623—656.

после сжатия специальным кодером передаются по каналу связи; окружающая среда может создавать шумовые помехи и исказить передаваемое сообщение, которое впоследствии подвергается обратному декодированию и доходит до получателя в удобном для него виде.

В 1949 году выходит работа Шеннона «Математическая теория коммуникации»⁶⁶, написанная им в соавторстве с американским математиком Уорреном Уивером; в ней ученые более подробно описывают основные принципы передачи информации, а также вводят понятие информационной энтропии и единицу измерения количества информации – бит. На основании построенной модели также формулируются три ключевые проблемы, встающие перед информационной теорией: это техническая проблема, связанная с точностью передачи сообщения по каналу связи и уменьшением шумовых помех, семантическая проблема, затрагивающая вопрос о передаче содержания сообщения, и проблема эффективности, касающаяся влияния полученного сообщения на поведение реципиента (также известная как «прагматическая проблема»). Сама информация определялась авторами как противоположность энтропии, то есть как уменьшение неопределенности: чем большим объемом информации обладает некоторая физическая система, тем больше степень ее организационной упорядоченности⁶⁷.

Несмотря на то, что сами авторы математической теории коммуникации предупреждали, что предложенная им трактовка информации является сугубо технической и имеет мало общего с такими понятиями, как «значение» или «содержание»⁶⁸, их теория все равно получила двусмысленные толкования.

⁶⁶ Shannon C. E., Weaver W. The Mathematical Theory of Communication. University of Illinois Press, 1963.

⁶⁷ Разумеется, коммуникационная теория Шеннона-Уивера не является единственной известной сегодня теорией информации. Наравне с ней существует целый ряд других концепций, в рамках которых информация трактуется иначе (среди них алгоритмический подход А.Н. Колмогорова, модальный подход Д. Барвайза и Д. Перри, семантическая информационная теория Л. Флориди и многие другие). Тем не менее, именно математическая теория Шеннона оказала наибольшее влияние на когнитивные науки того времени.

⁶⁸ Так, сам Шеннон писал: «Часто сообщения имеют значение, то есть они относятся или соотносятся в рамках какой-либо системы с определенными физическими или концептуальными объектами. Эти семантические аспекты коммуникации не имеют отношения к инженерной проблеме. Важным аспектом является то, что фактическое сообщение выбирается из набора возможных сообщений» (Shannon C. The mathematical theory of communication // Bell System Technical Journal. 1948. №27. P. 379). Уоррен Уивер также отмечал: «Слово “информация” в рамках данной [математической] теории используется в особом смысле, который не следует путать с обыденным словоупотреблением данного понятия. В частности, информацию

Историк науки Лили Э. Кей полагает, что на протяжении всего периода оформления информационной теории к используемому в ней термину «информация» так или иначе примешивались семантические и прагматические коннотации, обыкновенно встречающиеся в разговорной речи:

«Разумеется, теория информации, как и большинство форм знания (в том числе научного), неизбежно дискурсивна и основана на метафорах. Теория информации метафоризирует традиционное понятие информации, заимствуя семиотику человеческого языка для описания высокотехнических, строгих и нечеловеческих процессов. И все же богатство, обеспечиваемое многозначностью и двусмысленностью информационной катахрезы, оказалось непреодолимым. Понятия информации, ее хранения и передачи вызвали убедительные и обманчиво доступные образы коммуникации, быстро изменившие научные и популярные представления о природе и обществе»⁶⁹.

Действительно, формальные построения Шеннона-Уивера тут же были восприняты не только в математических и инженерных науках, но также и зарождавшейся кибернетике. Кибернетика, как наука о закономерностях получения, хранения, кодирования и передачи информации в сложных самоуправляющихся системах, представляла собой совокупность исследований в рамках разнообразных дисциплин, включавших в себя теорию систем, математическую логику, информатику, эволюционную биологию, антропологию и др. Большинство данных исследований появилось в результате работы конференций Мэйси, проходивших в Нью-Йорке между 1941 и 1960 годами (конференции, посвященные непосредственно вопросам кибернетики, проходили в период с 1946 по 1953 г.г.). На данных конференциях обсуждались вопросы, касающиеся работы саморегулирующихся и целенаправленных систем, компьютерных и органических механизмов обратной связи, искусственных

не следует путать со значением» (Weaver W. Recent Contributions to the Mathematical Theory of Communication. University of Illinois Press. 1962. P. 99).

⁶⁹ Kay L.E. Who Wrote the Book of Life? A History of the Genetic Code. Stanford University Press. 2000. P. 21.

языков, аналоговых и цифровых интерпретаций работы мозга, теории принятия решений, а также затрагивались проблемы рефлексии, обучения и восприятия. Активное участие в этих дискуссиях принимали такие известные ученые как Грегори Бейтсон, Маргарет Мид, Джон фон Нейман⁷⁰ и др. Именно благодаря конференциям Мейси состоялось научное сотрудничество мексиканского физиолога Артуро Розенблюта, американского математика Норберта Винера и компьютерного инженера Джулиана Бигелоу, что привело их к совместному написанию в 1943 году статьи, представляющей большой интерес для изучения генезиса кибернетики, – «Поведение, целенаправленность и телеология»⁷¹. Данная статья, в свою очередь, оказала существенное влияние на содержание опубликованной спустя пять лет знаменитой монографии Винера «Кибернетика, или управление и связь в животном и машине»⁷², в которой уже можно встретить следующие два характерных определения информации: *«Информация – это обозначение содержания, полученное нами из внешнего мира в процессе приспособления к нему нас и наших чувств»* и *«Информация – это не материя и не энергия, информация – это информация»*⁷³.

Как уже было отмечено, информационная теория Шеннона была объектом достаточно большого количества дискуссий, проводившихся в рамках конференций: именно эту теорию многие исследователи расценивали в качестве необходимого концептуального фундамента, призванного объединить математические науки с теорией систем, биологией, психологией и даже социологией. Тем не менее, данная теория нередко критиковалась различными участниками конференций, особенно когда стал активно подниматься вопрос об описании возможных семантических свойств информации и связанных с ними процессов рефлексии, характерных для высокоразвитых живых организмов. Так

⁷⁰ Отметим, что именно Джон фон Нейман в 1958 году издает книгу «Компьютер и мозг», оказавшую существенное влияние на становление и развитие компьютерной метафоры в науках (Von Neumann J. The Computer And The Brain. New Haven: Yale University Press. 1958).

⁷¹ Rosenbluelh A., Wiener N., Bigelow J. Behavior, Purpose and Teleology. // Philosophy of Science. Baltimore. 1943. № 10. P. 18–24.

⁷² Wiener N. Cybernetics or Control and Communication in the Animal and the Machine. John Wiley & Sons Inc., New York. 1948.

⁷³ Винер Н. Кибернетика, или управление и связь в животном и машине (2-е изд.). М.: Наука. 1983. С. 344.

называемые рефлексивные петли обратной связи стали весьма популярной темой более поздних обсуждений, происходящих в рамках конференций Мейси, особенно когда дело касалось интерпретации поведения человека. Так, нейроченые Уоррен Мак-Каллок и Уолтер Питтс активно работали над созданием первой математической модели нейрона, основанной на предположении о том, что каждый нейрон имеет некий пороговый уровень активации, который должен быть достигнут посредством приема входящих сигналов возбуждения: лишь после преодоления данного порога нейрон сможет передать свой собственный сигнал другим нейронам⁷⁴. Британский исследователь Уоррен Мак-Каллок предположил, что разряжение нейрона может быть связано с событием или взаимодействием, происходящим во внешнем мире, обеспечивающим сенсорный стимул, который затем улавливается и обрабатывается нервной системой. Ученый также показал, как сигнальный путь нейронной сети может быть настроен рефлекторно, заставляя нейроны непрерывно возбуждать друг в друга в круговом контуре обратной связи без каких-либо новых дополнительных входящих сигналов. Нейроченый утверждал, что это объясняло ментальные явления, в частности те, в которых представление воспринимаемого внешнего стимула было искажено, как у людей, страдающих от галлюцинаций или синдрома фантомной конечности. В связи с данными обсуждениями был поднят вопрос о природе субъективных переживаний, на который в рамках конференций Мейси ученые так и не смогли найти последовательный ответ. Тем не менее, исследователи преуспели в решении смежных проблем, касающихся работы механизмов рефлексии, обучения, восприятия и т.д.; именно благодаря данным успехам, во многом обусловленным междисциплинарным взаимодействием, наука того времени заметно продвинулась в изучении процессов мышления.

Таким образом, путь, который прошло понятие информации, прежде чем попасть в биологию и когнитивные науки, оказался тернистым. Уже упомянутая

⁷⁴ McCulloch W.S., Pitts W. A Logical Calculus of Ideas Immanent in Nervous Activity // Bull/ Mathematical Biophysics, №5, 1943.

выше исследовательница Лили Кей отмечает, что состоялся своеобразный двойной перенос данного понятия: сначала из обыденного языка оно попало в математическую теорию коммуникации, а затем – в смежные научные области, включая когнитивистику:

«При метафорическом применении к биологическим явлениям понятие “информация” становится еще более проблематичным: оно восстанавливает свой первоначальный смысл как “знание” и “значение”, но при этом нарушает принципы теории информации, которая предположительно и узаконила биологические приложения. Таким образом, это понятие становится метафорой метафоры, катахрезой и обозначающим без обозначаемого»^{75, 76}.

То, что Кей называет лингвистическим «зеркальным коридором»⁷⁷, сказалось на процессе формирования нового языка биологии в середине прошлого века. Ссылаясь на американского философа Ричарда Бойда, исследовавшего феномен научной метафоры, Кей отмечает, что биология отныне признавала следующие положения:

- мышление является своеобразным процессом «обработки информации», а мозг – своего рода компьютером;
- определенные когнитивные процессы являются «запрограммированными»;
- существует «внутренний язык мышления», с помощью которого выполняются когнитивные «вычисления»;
- информация «закодирована» в «хранилище» мозга (памяти);
- сознание является феноменом «обратной связи»

⁷⁵ Kay L.E. Who Wrote the Book of Life? A History of the Genetic Code. Stanford University Press. 2000. P. 24.

⁷⁶ Стоит отметить, что подобного мнения придерживаются и многие другие современные ученые и историки науки. См. например: Bergstrom C., Rosvall M. The Transmission sense of Information // Biology and Philosophy. 2011. №26. P. 159–176; Griffiths P.E. Genetic Information: A Metaphor in Search of a Theory // Philosophy of Science. 2001. №68. P. 394–412; Jablonka E. Information: its Interpretation, its Inheritance, and its Sharing // Philosophy of Science. 2002. №69(4). P. 578–605; Lean O. Getting the Most out of Shannon Information // Biology and Philosophy. 2014. №29. P. 395–413; Levy A. Information in Biology: A Fictionalist Account // Noûs. 2011. №45(4). P. 640–657; Maynard Smith J. The Concept of Information in Biology // Philosophy of Science. 2000. №67(2). P. 177–194.

⁷⁷ Там же. P. 28.

- и др.⁷⁸

Безусловно, некоторые из подобных метафор оказались достаточно продуктивными, однако вопрос о том, каким образом следует толковать используемую в них информационную терминологию остается весьма насущным, ведь данные метафоры предполагают совмещение синтаксического, семантического и прагматического аспектов толкования информации и, как следствие, оставляют нерешенными проблемы, связанные с определением сущности информационных структур, их онтологического статуса, степени зависимости от условного получателя информации и т.д. Именно этот факт в дальнейшем окажет заметное влияние и на ход дискуссий вокруг философских теорий сознания, активно использующих информационный концептуальный аппарат, о чем более подробно будет сказано в последующих главах.

1.1.2. Машины Тьюринга и компьютерное моделирование интеллектуальных функций

Понятие «вычисление» (англ. «computation»), обозначающее некий процесс, позволяющий эффективно преобразовывать входные данные с помощью определенной алгоритмической процедуры, появилось в науке и философии несколько столетий назад, несмотря на то, что множество привычных нам вычислительных устройств стало частью нашей повседневной жизни лишь относительно недавно. Некоторые историки науки относят начало активного употребления этого понятия к рубежу XVI-XVII веков, когда были сконструированы первые механические вычислительные устройства, среди которых знаменитые «вычислительные часы» Шикарда и счетная машина Паскаля⁷⁹. Эти механические калькуляторы представляли особый интерес для ученых и философов, так как они использовались не для управления другими

⁷⁸ Там же. Р. 24.

⁷⁹ Williams M.R. A History of Computing Technology. CA: Computer Society Press. 1997; Augarten S. Bit by Bit. An Illustrated History of Computers. London: George Allen and Unwin. 1985.

машинами или механическими устройствами, а для выполнения операций с числами: так как вычисления не могли быть реализованы непосредственно на числах, представляющих собой абстрактные сущности, они должны были быть опосредованы через нечто физическое; так, в результате соотнесения величин физического измерения (например, длин физических объектов) с числами, вычисления могли быть реализованы путем физического манипулирования этими объектами. Однако вскоре эти достаточно сложные манипуляции уступили место операциям с использованием знаковых представлений, очевидное преимущество которых стало причиной быстрого развития данного метода, впоследствии ставшего парадигмой для мышления в целом⁸⁰. Данный поворот в математике и механике отчасти послужил развитию механистического мировоззрения, в рамках которого вскоре появилась идея, что не только вычисление, но и человеческое мышление предполагает манипулирование некоторыми знаками. В своей радикальной форме эта идея была обобщена и представлена Томасом Гоббсом в его работе 1655 года «О теле», в которой английский философ пишет: *«Под рассуждением я подразумеваю [...] исчисление. Вычислять - значит находить сумму складываемых вещей или определить остаток при вычитании чего-либо из другого. Следовательно, рассуждать значит то же самое, что складывать и вычитать»*⁸¹. Некоторые историки философии отмечают, что толкуемое в логико-математическом ключе понятие «вычисление» применительно к процессам рассуждения активно использовали также Рене Декарт, Готфрид Вильгельм Лейбниц, Жюльен Офре де Ламетри и многие другие философы Нового времени, активно обсуждавшие связь между операциями, выполняемыми механическими калькуляторами, и работой мышления, оперирующего ментальными образами⁸².

⁸⁰ Согласно историку Вернону Пратту, таким образом в науке был сделан своеобразный переход от «аналоговых» к «цифровым» представлениям (Pratt V. *Thinking Machines. The Evolution of Artificial Intelligence*. Oxford: Basil Blackwell, 1987. P. 32).

⁸¹ Гоббс Т. Сочинения в 2 т. Т.1. М. Мысль, 1989. С. 74.

⁸² См. например: Alistair M.C. *Computational Thought from Descartes to Lovelace* // *The Routledge Handbook of the Computational Mind* (ed. by M. Spevak, M. Colombo). New York. 2019. P. 9-12; Arthur R.T.W. *Leibniz as a precursor to Chaitin's Algorithmic Information Theory* // *Information and the History of Philosophy* (ed. by C. Meyns). Routledge. 2021. P. 153-177/

Более основательно вычислительный подход к описанию ментальных состояний стал разрабатываться лишь спустя два столетия, параллельно с ростом эмпирических исследований в области нейрофизиологии. В середине XIX века вместе с постепенным развитием наук о мозге появились и математические теории, призванные объяснить вычислительную природу мышления. Значительный вклад в подобные исследования внес английский логик Джордж Буль, опубликовавший в 1854 году книгу «Исследование законов мышления»⁸³, в которой он поставил задачу изучения фундаментальных свойств мышления и рассуждения с последующим представлением данных свойств в форме формальных исчислений. Формальная алгебраическая система Буля, определяемая бинарными операциями и двумя элементами («истина»/«ложь») послужила, в свою очередь, основой для последующих научных теорий, оперирующих понятием «вычисление».

Данное понятие, долгое время используемое многими исследователями на интуитивном уровне, получило свою формальную интерпретацию лишь в 1930х годах, став предметом развитого логического анализа. Сложно найти ученого, оказавшего большее влияние на зарождающиеся в это время вычислительные теории, чем английский логик и математик Алан Тьюринг, который в 1936 году заложил основы так называемой «теории вычислимости», находившейся на стыке математической логики, теории алгоритмов и информатики. Как нетрудно догадаться из названия, данная теория занималась сравнением различных моделей вычислений, а также вопросами о возможности доказательства (выводимости) утверждений в рамках тех или иных теорий. В том же году совместно с американским математиком Алонзо Черчем Тьюринг доказал, что многие известные в математике формализмы применяют вычисления одного и того же класса функций, что означало, что различные механические процессы могут иметь одинаковые вычислительные возможности (сходство этих формализмов описывается в определении, известном как «тезис Черча-Тьюринга»); в связи с этим Тьюрингом была предложена абстрактная модель

⁸³ Boole G. An Investigation of the Laws of Thought. London: Walton and Maberly. 1854. P. 3.

вычислительной машины, способной решать любые алгоритмические задачи, – так называемая «машина Тьюринга», – которая сразу вдохновила многих исследователей на создание похожих теоретических моделей, объясняющих устройство и функционирование человеческого когнитивного аппарата.

Согласно схеме, предложенной Тьюрингом, подобная машина обладает бесконечным количеством линейно расположенных ячеек памяти и центральным процессором, сканирующим данные ячейки и выполняющим четыре элементарные операции: запись символа в ячейку, стирание символа из ячейки, перемещение к следующей ячейке, возврат в предыдущую ячейку. Машинная таблица определяет, какую элементарную операцию выполняет центральный процессор, учитывая его текущее состояние и символ, к которому он обращается в данный момент. Таким образом, таблица содержит конечный набор рутинных четких инструкций, управляющих вычислениями. Помимо этого, математик также доказал возможность существования «универсальной машины Тьюринга», способной имитировать любую другую машину Тьюринга.

Быстрый прогресс в области компьютерных наук побудил многих исследователей, в том числе и самого математика, задуматься о том, возможно ли построить компьютер, способный к выполнению некоторых интеллектуальных операций характерных для человеческого мышления. Именно этот вопрос ученый поднял в своей работе «Вычислительные машины и разум»⁸⁴, написанной и опубликованной им в 1950 году: в ней Тьюринг сформулировал правила знаменитой «игры-имитации» и оценил вероятность создания «думающих» машин. Ссылаясь на ограниченные возможности нашего перцептивного и когнитивного аппарата, ученый утверждал, что любой символический алгоритм, выполняемый человеком, может быть воспроизведен подходящей машиной Тьюринга, где мыслительная деятельность может быть попросту отождествлена с состоянием компьютерной системы, которое определено лишь своей ролью в продуцировании последующих состояний или выводом соответствующих данных, представленных, в случае человека,

⁸⁴ Turing A. Computing Machinery and Intelligence // Mind LIX, №236, 1950. P. 433—460.

вербально-поведенческой реакцией. Математик приходит к выводу, что формализм машины Тьюринга, несмотря на свою крайнюю простоту, достаточно силен, чтобы охватить все выполняемые мозгом механические процедуры над символическими конфигурациями.

Уже в середине 1950-х годов ученые стали задумываться над созданием машин, способных выполнять основные умственные задачи, такие как решение проблем, рассуждение, принятие решений и т.д. С целью изучения возможности проектирования подобных машин в 1956 году состоялся семинар Дартмутского колледжа, в котором приняли участие такие ученые, как Джон Маккарти, Марвин Ли Минский, Аллен Ньюэлл, Герберт Александер Саймон и др. На данном семинаре ученые разбирали достижения современного им компьютерного программирования и оценивали перспективы создания искусственного интеллекта (само понятие «artificial intelligence» вошло в оборот именно благодаря Дартмутскому семинару). Ранние исследования в области искусственного интеллекта основывались на компьютерном моделировании логических и дедуктивных способностей, и первые успехи в этом направлении вдохновили огромное количество исследователей, в том числе и философов. Вскоре термин «вычисление» и непосредственно связанное с ним выражение «обработка информации» стали центральными для многих когнитивных теорий, и мозг стал все чаще уподобляться своеобразному компьютеру, чья функция заключалась в получении информации из окружающей среды, ее обработке и использовании в целях управления.

1.1.3. Когнитивная революция и появление функционализма

Как уже было отмечено выше, параллельно с развитием теории информации, кибернетики и первых моделей искусственного интеллекта шел еще один немаловажный процесс: теория бихевиоризма стала постепенно терять свою популярность среди представителей научного и философского сообщества.

Как известно, бихевиоризм был чрезвычайно популярной исследовательской программой среди психологов и философов примерно с 1930х годов⁸⁵. Однако в 1950-1960-х годах классический вариант бихевиоризма стал терять свое влияние: сомнению подвергся его основополагающий тезис о том, что поведение может быть объяснено без учета внутренней ментальной деятельности (стратегия, которую философ Гилберт Райл охарактеризовал как «изгнание духа из машины»⁸⁶). Для многих критиков бихевиоризма стало очевидным, что характер человеческого поведения зависит прежде всего не от истории подкрепления индивида, а от того факта, что окружающая среда или история обучения особым образом представляются самим индивидом. Бихевиоризм, игнорировавший роль перцептивного восприятия и субъективных ощущений, стал, в конце концов, противоречить сам себе: один из ключевых представителей данного направления, гарвардский психолог Беррес Фредерик Скиннер в своей книге 1957 года «Вербальное поведение» допускал использование различных ментальных терминов, что сразу же было замечено многими критически-настроенными учеными, в частности американским лингвистом Ноамом Хомским. В своей рецензии на работу Скиннера Хомский отметил, что бихевиористские модели описания языка не могут объяснить различные аспекты его изучения: кажется совершенно неверным, что изучение языка зависит от применения детального подкрепления, ведь язык как таковой, по-видимому, изучается без явного и подробного обучения, и бихевиоризм не предлагает объяснения того, как это может быть. Хомский отмечал, что человеческое поведение и способности часто превосходят ограничения

⁸⁵ Идеи бихевиоризма поддерживали такие философы, как Гилберт Райл, Рудольф Карнап, Уиллард Куайн, Карл Густав Гемпель и другие. Куайн, например, полагал, что понятие психологической или умственной деятельности не имеет места в научном объяснении ни происхождения, ни смысла речи, а говорить в научной манере о значении высказывания – значит говорить о стимулах высказывания, то есть о его так называемом «стимульном значении» (Quine W. *Word and Object*, Cambridge, MA.: MIT Press, 1960. P. 43-44). Гемпель, в свою очередь утверждал, что все психологические утверждения, которые имеют смысл можно перевести в утверждения, которые не включают психологические понятия, но только понятия, характеризующие физическое поведение (Hempel C. *The Logical Analysis of Psychology // Readings in Philosophical Analysis* (eds. N. Feigl and W. Sellars), New York: Appleton-Century-Crofts, 1949. P. 18). Однако уже в работе 1966 года философ заметил: «Чтобы описать поведенческие паттерны, склонности или способности [...] нам нужен не только подходящий бихевиористский словарь, но и психологические термины» (Hempel C. *Philosophy of Natural Science*. N. J.: Prentice-Hall, 1966. P. 110).

⁸⁶ Ryle G. *The Concept of Mind*. Routledge. 2009. P. 1-12.

индивидуальной истории подкрепления, которая часто слишком бедна, чтобы однозначно определить, что и как мы делаем; большая часть обучения, по видимому, требует уже существующих или врожденных репрезентативных структур или принципиальных ограничений, в рамках которых происходит обучение⁸⁷. Последовательная критика концепции Скиннера Ноамом Хомским способствовала оформлению идей так называемой «когнитивной революции», легших в основу альтернативных научных и философских теорий, призванных пролить свет на проблему мышления и сознания.

В 1956 году в Массачусетском технологическом институте прошел симпозиум, посвященный проблемам обработки информации, на котором были сделаны три доклада, заложившие, по мнению большинства исследователей, основы изучения человеческого познания с позиций когнитивистики и информационного подхода: это был доклад Хомского «Три модели описания языка», в котором обсуждались математико-ориентированные лингвистические модели, доклад американского психолога Джорджа Миллера «Семь плюс-минус два, или о некоторых ограничениях наших способностей перерабатывать информацию», в рамках которого была описана модель рабочей памяти, основанная на компьютерной метафоре, и доклад Герберта Саймона и Аллана Ньюэлла, в котором ученые представили свою первую рабочую модель искусственного интеллекта «Логик-Теоретик». В этом же году вышел ряд основополагающих работ по психологии, нейробиологии, лингвистике и психологии, а уже в 1960 году был открыт первый центр когнитивных исследований в Гарварде. Интенсивное развитие когнитивной науки стало толчком к развитию новых философских теорий, призванных объяснить природу сознания в отличных от бихевиористских терминах.

Одним из наиболее популярных подходов, послуживших своеобразной альтернативой бихевиоризму, идентифицирующему ментальные состояния с поведенческими диспозициями, стал функционализм, утверждающий, что со всяким ментальным состоянием соотносится некоторое функциональное

⁸⁷ Chomsky N. Review of Verbal Behavior //Language. 1959. №35. P. 26–58.

состояние системы или живого организма, определяемое его каузальным влиянием на другие состояния организма⁸⁸. Таким образом, можно сказать, что необходимые условия возникновения ментальных переживаний в рамках функционализма намного строже, чем в бихевиоризме: если в последнем для появления некоторого желания X системе или организму необходимо и достаточно обладать лишь определенным набором правил обработки входной и выходной информации, то в функционализме допустимо, что система может обладать этим набором правил, но при этом не желать X (как отметил в свое время философ Нэд Блок, функционалисты, обвиняли бихевиористов в излишнем «либерализме», заключавшемся в приписывании ментальности объектам, которые ей в действительности не обладают⁸⁹). В то же время, теория тождества расценивалась функционалистами как «шовинистская», так как отвергала возможность наличия ментальных свойств у систем с физической организацией, отличной от человеческой: функционалисты, как правило, не исключают возможности наличия психики у организмов, не имеющих, к примеру, мозга, ведь обладание соответствующими ментальными состояниями может быть обеспечено различными видами функциональных состояний. Это базовое положение функционализма, также известное как идея множественной реализуемости, делает данную философскую позицию практически метафизически-нейтральной и совместимой не только с редуктивным физикализмом.

⁸⁸ Функционализм был не единственной популярной концепцией, пришедшей на смену бихевиористским идеям. Практически параллельно с развитием первых теорий функционализма австралийские философы Уллин Плейс и Джек Сمارт, а также австриец Герберт Фейгл стали создавать теории тождества, трактующие ментальные процессы как тождественные физическим процессам, происходящим в мозге. Хотя представители теории тождества не отказались от задачи анализа ментальных состояний в тематически нейтральном ключе, эти философы поставили перед собой цель анализа ощущений в отличных от диспозициональных терминах. Однако теории тождества сразу стали предметом жесткой критики: исследователи отмечали, что подобные теории не способны решить проблему каузальной релевантности ментальных качеств, а также неверно трактуют природу самого тождества. Особенно известен модальный аргумент Сола Крипке против т.н. «типовой» теории тождества, в рамках которой любой тип ментального состояния признается тождественным определенному типу физических состояний (Kripke S. Naming and Necessity. Cambridge, MA: Harvard University Press. 1980. P. 152), перед которым, однако, по утверждению некоторых исследователей, способна устоять «конкретная» теория тождества, в рамках которой постулируется, что каждому конкретному ментальному переживанию тождественно конкретное физическое состояние, что, в некотором смысле, сближает данную теорию с функционализмом (см. например: McGinn C. Anomalous Monism and Kripke's Cartesian Intuitions // Analysis. 1977. №37(2). P. 78-80).

⁸⁹ Block N. Troubles with Functionalism // Minnesota Studies in the Philosophy of Science, №9, 1978. P. 263.

Тезис множественной реализуемости сознания вкупе с идеями Тьюринга стал основополагающим для философской позиции, представляющей мозг в виде своеобразной вычислительной системы, также известной как классический компьютеризм⁹⁰.

1.2. Классический компьютеризм: основные идеи и представители

Итак, в рамках классического компьютеризма механизмы мышления уподобляются некоторым логико-алгоритмическим операциям, а разного рода когнитивные способности, такие как решение проблем или обучение, – вычислениям наподобие абстрактной машины Тьюринга. В популярных философских работах нередко встречается утверждение о том, что сторонники классического компьютеризма обычно прибегают к так называемой «компьютерной метафоре», в рамках которой работа мышления уподобляется работе своеобразного компьютера. Однако многие современные ученые и философы полагают, что подобное сопоставление неверно, так как, во-первых, когнитивный аппарат здесь сравнивается не с компьютером как таковым, а с вычислительной системой⁹¹, а во-вторых, в рассматриваемой философской позиции не идет речь о метафорическом уподоблении мозга вычислительной системе: согласно классическому компьютеризму мозг действительно является вычислительной системой с определенным набором вычислительных функций, которые потенциально могут быть реализованы различными способами.

⁹⁰ Отметим, что само определение «классический» стало активно использоваться в когнитивных науках и философии лишь начиная с 1980-х годов, когда появилась необходимость подчеркнуть некоторую теоретическую обособленность символического вычислительного подхода от зарождавшихся в это время нейросетевых концепций. Сами представители данного направления поначалу именовали себя просто «компьютеризмистами».

⁹¹ К примеру, Дэвид Чалмерс полагает, что уподобление некоторой системы компьютеру подразумевает, что данная система является программируемой. Однако сторонники компьютеризма не обязательно должны соглашаться с утверждением о программируемости, ведь большинство машин Тьюринга, с работой которых сопоставляется работа мышления, не являются программируемыми (Chalmers D. A Computational Foundation for the Study of Cognition // The Journal of Cognitive Science, №12, 2011. P. 323–357).

Кроме того, некоторые историки философии отмечают, что утверждение о соответствии функциональной организации мозга функциональному устройству машины Тьюринга присуще далеко не всем теориям классического компьютеризма. Так, современный американский философ Михаэль Рескорла приводит ряд аргументов, демонстрирующих своеобразную ограниченность формализма машины Тьюринга. Во-первых, отмечает философ, машины Тьюринга выполняют исключительно символические вычисления, так как входные и выходные данные представляют собой символы, вписанные в ячейки памяти, в то время как мозг использует информацию и посылает сигналы, стимулирующие определенные моторные действия; полная теория должна описывать, как ментальные вычисления взаимодействуют с входными сенсорными данными и двигательными реакциями. Во-вторых, подобная машина обладает неограниченным объемом памяти, в то время как живые организмы могут похвастаться лишь весьма ограниченным объемом. В-третьих, она имеет центральный процессор, который, последовательно выполняя операции по переходу, получает доступ только к одной ячейке на ленте. Человеческий мозг же способен на параллельное выполнение разнообразных когнитивных операций. Наконец, вычисления, свойственные машине Тьюринга, всегда детерминированы в силу того, что каждое прошлое вычислительное состояние машины определяет всякое последующее ее состояние, хотя сторонники классического компьютеризма не отрицают, что для описания мышления скорее характерны стохастические определения вычислительных процессов, реализующихся в системе лишь с определенной долей вероятности. Рескорла заключает, что в рамках классического компьютеризма когнитивная деятельность уподобляется вычислительной системе, лишь отдаленно напоминающей стандартную машину Тьюринга, что позволяет данной позиции отклоняться от некоторых формализмов, разработанных знаменитым английским математиком⁹².

⁹² Rescorla M. The Computational Theory of Mind // The Stanford Encyclopedia of Philosophy (Spring 2017 Edition), Edward N. Zalta (ed.), 2015. URL: <https://plato.stanford.edu/archives/spr2017/entries/computational-mind/> (дата обращения: 04.12.2020).

Ряд исследователей утверждает, что вычислительный процесс, представленный в рамках классической компьютерно-ориентированной модели, можно в целом охарактеризовать как символический, алгоритмический, формальный и интерпретируемый⁹³: символическость означает, что вычисления можно описать в терминах символических манипуляций, где сами символы представляют собой набор репрезентаций, входящих в состав некоторой системы, организованной в соответствии с определенными синтаксическими и семантическими законами; алгоритмичность предполагает наличие особых конечных и четко заданных правил или инструкций, в соответствии с которыми вычислительная система манипулирует данными символами; формальный характер вычислительного процесса обеспечивается главенствующей ролью синтаксических свойств, а интерпретируемость, в свою очередь, подразумевает, что, несмотря на свой формальный характер, вычислительный процесс обладает семантическими качествами, вследствие чего он может быть описан в терминах перехода от одного значимого состояния в другое.

Однако большинство современных историков философии и науки сходятся во мнении, что компьютерно-ориентированные теории сознания представляют собой группу достаточно разрозненных концепций, которые весьма сложно объединить рядом общих положений⁹⁴. По этой причине наиболее верной стратегией изучения классического компьютерно-ориентированного подхода представляется разбор отдельных философских концепций. Далее будут рассмотрены основные идеи ключевых представителей классического компьютерно-ориентированного подхода, среди которых Хилари Патнэм, Джерри Фодор, Зенон Пилишин и Дэвид Марр.

⁹³ Samuels R. Classical Computational Models // The Routledge Handbook of the Computational Mind (ed. M. Spivak, M. Colombo). New York, 2019. P. 104.

⁹⁴ Boden M.A. Mind as Machine: A History of Cognitive Science. Oxford University Press. 2006; Clark A. Mindware: An Introduction to the Philosophy of Cognitive Science, Oxford: Oxford University Press. 2014; Marcus G. The Algebraic Mind, Cambridge. MA: MIT Press. 2001; Piccinini G. Computationalism in the Philosophy of Mind // Philosophy Compass. 2009. №4(3). P. 515-532.

1.2.1. Теория машинных состояний Хилари Патнэма

Одним из первых философов, разработавших последовательный вычислительный подход к решению проблемы сознания, был американский философ Хилари Патнэм. Именно Патнэма принято считать автором теории вычислительного функционализма, основанной, как уже было сказано ранее, на тезисе множественной реализуемости. Свой подход к описанию работы мышления философ начал разрабатывать с 1960-х г.г. Впоследствии Патнэм вспоминал об этом периоде своей философской работы следующее:

«В те дни я часто объяснял идею “машины Тьюринга” в своих курсах математической логики. Меня поразило, что в работах Тьюринга, как и в современной теории вычислений, “состояния” воображаемого компьютера (машины Тьюринга) описывались совершенно иначе, чем это принято в физической науке. Состоянием машины Тьюринга можно назвать такое вычислительное состояние, которое определяется его ролью в определенных вычислительных процессах, независимо от того, как оно физически реализуется. Построенный в девятнадцатом веке вычислительный механизм, работающий с бумагой и карандашом, и современный электронный компьютер могут находиться в одном и том же вычислительном состоянии, не находясь в одном и том же физическом состоянии. Я начал применять образы, предложенные теорией вычислений, к философии разума, и в лекции, прочитанной в 1960 году, я предложил многообещающую гипотезу под названием “функционализм”, согласно которому психические состояния человека являются вычислительными состояниями мозга»⁹⁵.

Предложенная Патнэмом теория была призвана продемонстрировать истинную сущность проблемы сознание-тело, которую, с точки зрения

⁹⁵ Putnam H. A Half Century of Philosophy, Viewed From Within // Daedalus, 1997. №1. P. 180-181.

философа, возможно свести к достаточно простым лингвистическим и логическим составляющим.

В своей статье 1960 года «Сознание и машины» Патнэм предпринял попытку ответить на вопросы о соотношении ментальных и физических свойств с помощью упомянутой аналогии с машиной Тьюринга. «Машина, – пишет Патнэм, – способная распознавать, по крайней мере, часть собственных структурных состояний, находится приблизительно в таком же положении, что и человек, который может определять с различной степенью достоверности только некоторые дисфункции своего организма [...] Ответ [машины] может содержать ссылку как на физическую структуру машины (“органы чувств” и т.п.), так и на ее “логическую структуру” (программу “чтения” и “интерпретации” входной ленты)»⁹⁶. Подобно тому, как работу абстрактной машины Тьюринга можно представить двумя разными способами – с помощью структурного чертежа инженера или посредством «машинной таблицы» – так и психика человека может быть описана либо с помощью бихевиористского подхода, подразумевающего полный разбор поведения в физических терминах, либо в виде абстрактного описания в терминах ментальных состояний и восприятий. Сопоставление машинных логических состояний с ментальными переживаниями предполагает наличие у подобных состояний особых характеристик, отличающих их от физических состояний: к ним философ отнес функциональную организацию, вербализацию и открытость «программы» для рациональной критики. Философ полагает, что всякое утверждение о соотношении физических и ментальных состояний в рамках проблемы «сознание-тело» в некоторой степени схоже с соответствующими утверждениями о структурно-организационных и логических состояниях машины Тьюринга. Следовательно, заключает Патнэм, подобные утверждения представляют собой не более чем концептуальное и вербальное открытие (verbal discovery).

⁹⁶ Патнэм Х. Сознание и машины. Пер. Л.Б. Макеевой // Философия сознания. М.: Дом интеллектуальной книги. 1999. С. 36.

В своей статье «Природа ментальных состояний» философ продолжил свое рассуждение о проблеме тождества физических и ментальных состояний. Пытаясь ответить на вопрос «Является ли боль состоянием мозга?», Патнэм отбрасывает основные аргументы в пользу теории тождества, среди которых аргумент нарушения нормы естественного языка, аргумент эмпирической редукции и аргумент инвариантной корреляции, и строит свое доказательство ложности подобной теории. Патнэм полагает, что боль не является состоянием мозга (или какой-либо другой ограниченной физической системы тела), но является функциональным состоянием всего организма. Философ ссылается на следующие положения:

«1) Все организмы, способные ощущать боль, есть вероятностные автоматы.

2) Любой организм, способный ощущать боль, имеет по крайней мере одно описание определенного вида (т.е. способность ощущать боль есть обладание функциональной организацией соответствующего вида).

3) Ни один организм, способный испытывать боль, не предполагает разложения на части, которые по отдельности имеют описание указанного в (2) вида.

4) Для каждого указанного в пункте (2) описания существует подмножество сенсорных входных данных таких, что организм, имеющий такое описание, ощущает боль тогда и только тогда, когда некоторые из его сенсорных входных данных принадлежат к этому подмножеству»⁹⁷.

Функциональное объяснение, по утверждению Патнэма, обладает рядом преимуществ по отношению к простому отождествлению субъективных переживаний и ощущений с физическими состояниями: так, оно совместимо с некоторыми нефизикалистскими воззрениями, и даже если допустить, что

⁹⁷ Патнэм Х. Психологические предикаты (Природа ментальных состояний). Пер. Л.Б. Макеевой // Философия сознания. М.: Дом интеллектуальной книги. 1999. С. 60.

некоторая система имеет дуалистическую природу, то она тоже может трактоваться как вероятностный автомат. К тому же, функциональное состояние, по мнению философа, не просто коррелирует с поведением организма, но и объясняет это поведение.

В статье 1967 года «Ментальная жизнь некоторых машин» Патнэм изучает применимость некоторых ментальных понятий к деятельности расширенных конечных автоматов⁹⁸ и снова приходит к выводу о неудовлетворительности редукционистских трактовок сознательного опыта:

«В случае машин Тьюринга функциональная организация задается машинной таблицей. Описание функциональной организации человека может быть совершенно иным и куда более сложным. Но важно то, что описания функциональной организации системы по своему типу логически отличаются от описаний ее физико-химического строения, так и от описаний ее реального и потенциального поведения. Думаю, что исследования в философии сознания потому так часто удивляют своей неудовлетворительностью, что в них не учитывается понятие функциональной организации или же оно смешивается с понятиями совершенно иного рода»⁹⁹.

Машинный функционализм Патнэма вскоре подвергся жесткой критике. Так в статье 1972 года «Чем не являются психологические состояния»¹⁰⁰ философы Нэд Блок и Джерри Фодор утверждают, что предложенный Патнэмом подход значительно умаляет возможности человеческого мышления: уподобляя мозг вероятностному автомату, пусть даже обладающему неограниченным временем и объемом памяти, мы признаем конечность набора ментальных состояний, в которых может пребывать человек, а такое рассуждение, считают

⁹⁸ Говоря о расширенном автомате, Патнэм имеет в виду такую абстрактную машину Тьюринга, которую возможно дополнить органами чувств и органами движения, обеспечивающими ее возможность эффективно взаимодействовать с окружающим миром.

⁹⁹ Патнэм Х. Ментальная жизнь некоторых машин. Пер. Л.Б. Макеевой // Философия сознания. М.: Дом интеллектуальной книги. 1999. С. 87.

¹⁰⁰ Block N., Fodor J. What Psychological States Are Not // The Philosophical Review. 1972. №81. P. 159–181.

философы, никак не соответствует представлениям о продуктивности человеческой мысли. Блок и Фодор также отмечают, что мышление является систематичным, и каждое ментальное состояние должно соотноситься с другими ментальными состояниями: например, мысль «Джон любит Мэри» должна соотноситься с мыслью «Мэри любит Джона». Однако машинный функционализм соотносит ментальные состояния с машинными состояниями, в которых отсутствуют необходимые систематические отношения друг с другом. Таким образом, возражение философов предполагает, что концепция Патнэма пренебрегает существенными свойствами нашей психики, а потому не может являться релевантной теорией ментального.

Патнэм, однако, вскоре и сам пришел к подобному заключению. В 1973 году философ выступил на Ферстеровском симпозиуме, посвященном теме «Компьютеры и мышление», с докладом «Философия и наша ментальная жизнь», в котором отметил, что его утверждения о возможности сопоставления психических состояний человека и состояний условной машины Тьюринга не оправданы. Патнэм признал проблемы, встающие перед предложенным им вычислительным подходом (а именно упомянутые Блоком и Фодором проблемы недостаточной продуктивности и систематичности), а также упомянул несколько концептуальных трудностей, связанных с вопросом о правомерности использования машинных метафор в науке: несмотря на то, что подобные метафоры и соответствующие им аналогии способствуют лучшему пониманию особенностей функционального устройства различных сложных систем, их чрезмерное употребление недопустимо для философии и когнитивистики.

Поначалу философ пытался разработать другие функциональные подходы к описанию свойств ментальных состояний путем развития психологической теории, основанной на формальных построениях теории вычислений, но и данная версия «психологического компьютеризма» потерпела неудачу и была вскоре отброшена философом, так как, с его точки зрения, ученым вряд ли когда-нибудь удастся обнаружить весь спектр психофизических законов, позволяющих отличать функциональные корреляты одних психических

состояний от других. В период после 1980-х годов он подвергнул серьезному критическому разбору как собственные теоретические наработки, так и многие другие философские концепции сознания, выдвигаемые его коллегами. В своем небольшом тексте под названием «Химера когнитивной науки», подготовленном специально для доклада на IV Российском философском конгрессе в 2005 году, философ пишет следующее:

«Размышления о научных возможностях, которые мы не способны реализовать в настоящее время, сами по себе безвредны; но я понял, что возможность “идеальной психологической теории” такого рода есть “то, не знаю что”. Ни у кого нет ни малейшей идеи, как можно было бы конструировать подобную теорию. В наши дни мы постоянно слышим разговоры о “когнитивной науке”, но надо отличать научную теорию от плохой научной фантастики. Если я прав, то идея теоретической редукции в данном случае – редукции всей психологии, имплицитно присутствующей в нашей повседневной практике атрибутирования ментальных состояний, к вычислительной психологии – бессодержательна»¹⁰¹.

Однако, несмотря на последующий отказ Патнэма от своих ранних идей, его теория вычислительного функционализма оказала огромное влияние на когнитивную науку и философию 1960-1980-х годов. Это влияние прослеживается не только в работах сторонников классического компьютеризма, но также и многих других исследователей, развивавших альтернативные вычислительные подходы к описанию работы мышления и сознания.

¹⁰¹ Патнэм Х. Химера когнитивной науки // Тезисы докладов и выступлений IV Российского Философского конгресса (Москва, 24-28 мая 2005). Пер. В.В. Васильева. [Электронный ресурс]. URL: https://www.academia.edu/4091972/Споры_о_сознании_в_аналитической_философии_предисловие_к_публикации (дата обращения: 06.04.2021).

1.2.2. Компьютационализм и репрезентативизм в философии Джерри Фодора

Будучи учеником и одним из главных критиков Хилари Патнэма, американский философ и психолингвист Джерри Алан Фодор, развивал некоторые идеи своего учителя, показавшиеся ему наиболее пригодными для описания работы мышления. Выделив и проанализировав наиболее проблемные места теории машинного функционализма, Фодор пришел к выводу, что любая непротиворечивая вычислительная теория сознания должна быть также дополнена теорией репрезентативизма, определяющей содержание ментальных переживаний посредством указания на их интенциональные и семантические качества. Для этого вычислительные процессы должны признаваться символическими (то есть определенными на множестве репрезентаций) и формальными (то есть применимыми к репрезентациям в силу своих синтаксических свойств): согласно философу, два ментальных состояния могут быть различны по своему содержанию, только если они соотносятся с различными (с формальной точки зрения) репрезентациями¹⁰².

Чтобы дополнить идеи Патнэма Фодор разработал репрезентативистскую теорию сознания, согласно которой мышление представляет собой вычисления в стиле машины Тьюринга, производимые над языком мышления. Данный язык мышления (Language of Thought; или Ментализ) Фодор определил как систему, состоящую из простых и сложных ментальных репрезентаций¹⁰³. Подобная

¹⁰² Fodor J. Representations. Philosophical Essays on the Foundations of Cognitive Science. Harvester Press, 1981. P. 226.

¹⁰³ Предложенное Фодором описание Ментализа как непосредственно «языка», предполагающее, что все ментальные символы похожи на выражения естественного языка, может вводить в заблуждение. Сегодня многие исследователи полагают, что существуют возможные «неязыковые» (то есть непропозициональные) форматы ментальных символов. Так, некоторые современные сторонники теории вычислительного репрезентативизма полагают, что когнитивные вычисления также оперируют изображениями, картами, диаграммами и другими представлениями. Подобная плюралистическая линия представляется особенно пригодной для описания процессов восприятия, а также для объяснения когнитивных процессов, характерных для животных. Так, уже упомянутый выше философ Михаэль Рескорла исследует феномен когнитивных карт (Rescorla M. Cognitive Maps and the Language of Thought // The British Journal for the Philosophy of Science, 2009. №60. P. 377–407), а философ Эдуард Кэмп, ссылаясь на биологические исследования социальных взаимодействий некоторых животных, утверждает, что те могут кодировать социальные отношения с помощью особых древовидных визуальных представлений (Camp E. A Language of Baboon Thought? // The Philosophy of Animal Minds (ed. R. Lurz). Cambridge: Cambridge University Press, 2009).

теория постулировала существование не только конечного набора примитивных ментальных выражений, объединенных в потенциально бесконечную систему сложных выражений, но и особые систематические связи, по отношению к которым субъект может занимать различные установки (таким образом Фодор объяснял продуктивность и систематичность мышления, чего, по его мнению, не хватало в теории машинного функционализма Патнэма).

Фодор, безусловно, понимал, что подобная версия компьютеризации должна быть снабжена релевантной натуралистической теорией, призванной объяснить, каким образом у описанных им ментальных репрезентаций появляются семантические и интенциональные свойства. В работе «Психосемантика» философ отстаивает позицию интенционального реализма и отмечает, что интенциональные свойства присущи репрезентациям, каузально связанным с их объектами из внешнего мира¹⁰⁴. Так как подобными интенциональными свойствами теоретически могут обладать не только высокоразвитые живые существа, но и, возможно, так называемые «расширенные компьютеры», дополненные сенсорными системами или аналогами органов чувств, философ допускает, что искусственные системы также могут каузально взаимодействовать с окружающей средой и являться полноправными интенциональными агентами.

Экстерналистская трактовка ментальных содержаний, предложенная Фодором, получила название теории асимметричной зависимости (*asymmetric dependency theory*). Согласно данной теории, причина, определяющая содержание некой ментальной репрезентации, является фундаментальной, в то время как причины, не определяющие ее содержание, полностью зависят от фундаментальной причины; иными словами, причины, не определяющие содержание, не существовали бы, если бы не фундаментальная причина. Таким образом, существует асимметричная зависимость между законами репрезентации: к примеру, собака является фундаментальной причиной

¹⁰⁴ Fodor J. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT/Bradford, 1987. P. 141.

ментального представления «собака», в то время как лиса может быть нефундаментальной причиной схожей репрезентации лишь потому, что лис нередко путают с собаками¹⁰⁵.

Итак, асимметричные каузальные зависимости признаются философом своеобразной основой, на которой базируется семантическая составляющая любого ментального содержания. Однако, ряд исследователей отмечает, что при этом сами зависимости также могут являться продуктом (или побочным продуктом) подобных содержаний: предполагается, что Y вызывают «X» только потому, что X вызывают «X», а не в силу каких-либо семантических фактов об «X»; в этом случае, однако, возможна ситуация, когда не только Y, но и многие другие причины способны вызывать «X» совершенно независимо от того, факта, что X вызывают «X». Чтобы избежать подобных следствий, мы вынуждены утверждать, что только Y может вызывать «X», так как, скажем, только лисы у нас ассоциируются с собаками. Этот ответ, однако, предполагает вывод асимметричных причинно-следственных зависимостей из значения репрезентации «собака», что нарушает исходное предположение проекта семантической натурализации¹⁰⁶.

Другая смежная проблема, встающая перед каузальной теорией Фодора, связана с природой номологической связи, якобы существующей между ментальным представлением и его непосредственным референтом: если допустить, что ментальный символ «X» номически связан с объектом X, то получается, что каждый раз, когда в окружающей среде появляется X, у нас возникает соответствующее представление. Однако кажется очевидным, что это далеко не так. По этой причине теории необходимо определить условия,

¹⁰⁵ Теорию асимметричной зависимости можно выразить следующим образом:

- 1) Объект X является фундаментальной причиной репрезентации «X»;
- 2) Для любых объектов Y, не являющихся X, если Y является причиной «X», то «Y, вызывающие “X”», асимметрично зависимы от «X, вызывающих “X”».

¹⁰⁶ Подобные замечания к теории Фодора встречаются у многих авторов. См. например: Adams F., Aizawa K. Fodorian Semantics // *Mental Representations* (S. Stich, T. Warfield (eds.). Oxford: Basil Blackwell, 1994. P. 223–242; Wallis C. Asymmetrical Dependence, Representation, and Cognitive Science // *The Southern Journal of Philosophy*, 1995. №33. P. 373–401; Gibson M. Asymmetric Dependencies, Ideal Conditions, and Meaning // *Philosophical Psychology*, 1996. №9. P. 235–259.

необходимые для появления того или иного представления (помимо условия наличия во внешнем мире самого референта).

Так как подобные условия должны быть описаны в терминах, которые не будут апеллировать к семантическому значению самой репрезентации, Фодор предпринял попытку адаптировать информационный подход для их характеристики. В «Психосемантике» философ описывает механизмы получения информации из внешнего мира следующим образом: после того, как информация об окружающей среде была перекодирована в электрические сигналы, которыми оперирует наш мозг, эта информация проходит определенные процедуры обработки и интеграции с уже имеющейся у субъекта информацией, в результате чего генерируется ментальный образ¹⁰⁷. Вычислительные механизмы, обеспечивающие фиксированную корреляцию ментальных содержаний с соответствующими объектами внешнего мира, могут быть описаны каузально-синтаксически, то есть без апелляции к семантическим структурам данных содержаний: *«Картина такова, – пишет философ, – что между органами восприятия и имеющимися убеждениями находится компьютер, и обозначение определенных психофизических понятий происходит в ходе выполнения компьютером определенных вычислений, которые, в свою очередь, способствуют появлению значения “протон” (или “лошадь”, или любого другого значения)»*¹⁰⁸.

Несмотря на то, что в некоторых своих работах Фодор различает вычислительные механизмы, отвечающие за фиксацию репрезентаций, и механизмы, детерминирующие непосредственно содержания этих репрезентаций, философ описывает данные механизмы весьма поверхностно, так как считает эти и смежные вопросы сугубо «инженерными», а не

¹⁰⁷ Fodor J. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT/Bradford, 1987. P. 115-118.

Отметим, что философ практически не опирается на положения классической информационной теории, хотя в некоторых своих статьях он активно использует термин «информация», который трактуется им в духе Шеннона (например: Fodor J. *Information and association* // *Notre Dame Journal of Formal Logic*, 1986. №27. P. 307-323; Fodor J. *Information and representation* // *Information, Language and Cognition* (ed. P. P. Hanson). University of British Columbia Press, 1990).

¹⁰⁸ Там же. P. 123.

философскими. По этой причине предложенная философом вычислительная теория представляется неполной и весьма абстрактной.

Слабость теоретических построений Фодора отразилась и в его полемике с еще одним известным сторонником вычислительного подхода к описанию работы мышления и сознания – американским когнитивистом Стивеном Пинкером. В качестве своеобразного ответа на вышедшую в 1997 году книгу Пинкера «Как работает сознание»¹⁰⁹, основанную на идеях так называемого «Нового синтеза»¹¹⁰, Фодор пишет работу «Сознание работает не так»¹¹¹, в которой критикует основные выводы, сделанные Пинкером. Согласно Фодору, несмотря на значительные успехи вычислительной нейронауки, далеко не вся работа мозга может быть описана в вычислительных терминах, особенно если мы пользуемся объяснительными моделями, основанными на формализме машины Тьюринга. Помимо этого, философ приводит ряд аргументов, призванных показать несостоятельность вычислительных моделей, которые использует Пинкер, а также доказывает ограниченность объяснительного потенциала вычислительной психологии, которая, по его мнению, заключается в неспособности показать, как происходит сложное комплексное и абдуктивное мышление. В ответ на критику Фодора Пинкер публикует статью под названием «Так как же работает сознание?»¹¹², в которой обвиняет Фодора в неверной трактовке обсуждаемых им вычислительных моделей и, в частности, формализма машины Тьюринга, отмечая, что *«никто никогда не рассматривал всерьез идею о том, что когнитивная архитектура - это классическая архитектура Тьюринга»*¹¹³. Являясь сторонником идеи обобщенного вычисления (англ. *generic computation*), в рамках которого символы (состояния физических элементов, подобных микропроцессорам в компьютере или

¹⁰⁹ Pinker S. How the Mind Works. New York: Norton. 1997. (В русском переводе «Как работает мозг»).

¹¹⁰ Под «Новым синтезом» в данном случае имеется в виду синтез вычислительных концепций, появившихся во времена когнитивной революции в 1950-1960-х г.г., положений эволюционной биологии, а также теории генного отбора, впервые представленной в 1970-х г.г.

¹¹¹ Fodor J. The Mind Doesn't Work This Way. The Scope and Limits of Computational Psychology. MIT Press, 2000.

¹¹² Pinker S. So How Does the Mind Work? // Mind & Language. 2005. №1(20). P. 1–24.

¹¹³ Там же. P. 6.

нейронам в мозге) обладают как семантическими, так и каузальными свойствами, Пинкер показывает, что выдвигаемые Фодором аргументы несостоятельны, и многие положения предлагаемой философом теории сознания неудовлетворительны (в частности, Пинкер отмечает, что выдвигаемая Фодором модульная теория, в рамках которой когнитивный аппарат описывается как совокупность специфичных и независимых друг от друга модулей, не может быть верной, пока сохраняются утверждения о функциональной замкнутости и информационной инкапсулированности подобных модулей, ведь данные положения опровергаются эмпирическими науками).

Тем не менее, несмотря на эти и многие другие замечания, направленные в адрес теории Фодора, некоторые его идеи все же остаются весьма популярными не только среди философов, занимающихся проблемами сознания, но и среди нейробиологов и когнитивистов, развивающих предложенную им теорию репрезентативизма и по сей день.

1.2.3. Интерпретация вычислительных процессов в работах Дэвида Марра и Зенона Пилишина

К началу 1980-х годов происходит переосмысление некоторых положений и методологических оснований классического компьютеризма, связанное с необходимостью уточнения эпистемологических возможностей полюбившейся многими учеными и философами аналогии мышления и работы абстрактной машины Тьюринга.

Наибольшее влияние на многих исследователей оказал подход английского ученого Дэвида Марра, работавшего на стыке нейробиологии, психологии восприятия и исследований искусственного интеллекта и являвшегося одним из основателей современной вычислительной биологии. В своей работе, посвященной возможностям компьютерного моделирования зрения, Марр уделил особое внимание разработке дескриптивных установок, необходимых для верной интерпретации вычислительных процессов,

реализуемых во время работы восприятия и мышления. Ученый отмечает, что для изучения функционирования любого устройства, решающего некоторую задачу посредством обработки соответствующей информации, требуется его трехуровневое описание, в которое входят следующие составляющие:

«1) Информационная теория. На данном уровне работа устройства описывается как некоторое преобразование информации одного вида в информацию другого вида, формальные свойства которого являются точно определенными. Работая с этим уровнем, необходимо ответить на следующие вопросы: что является целью вычислительного процесса, почему именно этот процесс должен использоваться и в чем заключается логика стратегии, обеспечивающей его реализацию?»

2) Представление и алгоритм. Данный уровень связан с выбором представления для входной и выходной информации и выбором алгоритма, который должен быть использован для преобразования одного типа информации в другой. Требуется ответить на вопрос о том, каким образом можно реализовать существующую информационную теорию? В частности, как следует представлять входную и выходную информацию и что представляет собой алгоритм преобразования?»

3) Техническая реализация. Третий уровень описывает подробности физической реализации выбранных алгоритмов и представлений – детальную архитектуру вычислительной машины»¹¹⁴.

Данные уровни, по утверждению Марра, являются логически и каузально взаимосвязанными, но связи между ними свободны, что, в частности, означает

¹¹⁴ Марр Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов. М.: «Радио и связь». 1987. С. 272.

возможность корректной интерпретации тех или иных эмпирических данных на каком-либо одном необходимом уровне. Ученый пишет:

«Хотя эмпирически алгоритмы и механизмы (аппаратура) более доступны, именно высший уровень, т.е. уровень информационной теории, имеет решающее значение с точки зрения обработки информации. Причина заключается в том, что характер вычислений (процедур обработки информации, лежащих в основе восприятия) в большей степени зависит от задач обработки информации, подлежащих решению, а не от той конкретной аппаратуры, с помощью которой соответствующие решения находятся. Другими словами, алгоритм, вероятно, легче понять, исследуя характер решаемой задачи, чем изучая устройство (и его аппаратную часть), в котором он реализуется. Аналогичным образом попытка понять восприятие исключительно на основе изучения нейронов подобна попытке понять природу полета птиц, изучая лишь их оперение. Это просто невозможно. Для того чтобы осознать природу полета птиц, необходимо владеть аэродинамикой»¹¹⁵.

Идея трехуровневого описания работы вычислительных систем, предложенная Марром, была воспринята многими исследователями. Так, канадский когнитивист и философ Зенон Пилишин, являющийся одним из наиболее известных представителей классического компьютеризма, в книге 1984 года «Вычисление и познание», предлагает свою трактовку некоторых выдвигаемых Марром положений и разбирает основные трудности, присущие современным ему версиям вычислительных теорий мышления. Одна из подобных трудностей связана с вопросом о физических коррелятах вычислительных процессов: Пилишин отмечает, что формально-универсальный характер рассматриваемого подхода, несомненно, привлекает к нему большое количество сторонников из разных научных областей, но, в то же время, делает

¹¹⁵ Там же. С. 274.

само вычислительное описание весьма абстрактным и метафоричным. Безусловно, любые действия компьютера или любой другой вычислительной системы можно описать как каузально-взаимосвязанную последовательность физических состояний, переходы между которыми обусловлены физическими законами, однако, как отмечает философ, сама последовательность физических состояний системы ничего не говорит о происходящем в ней вычислительном процессе. Это объясняется, во-первых, тем, что далеко не все физические свойства системы связаны с ее вычислительной функцией, и, во-вторых, тем, что невозможно полное описание всех лежащих в основе вычислительных процессов физических состояний; более того, в силу тезиса множественной функциональной реализуемости, возможно говорить о потенциально бесконечном наборе подобных физических состояний¹¹⁶. По этим причинам Пилишин считает необходимым ввести особую функцию (философ именует ее «*instatiation function*»), зависящую от условного наблюдателя, позволяющую интерпретировать последовательность номологически обусловленных изменений в физических состояниях системы как вычисление.

Компьютеры, по утверждению Пилишина, помогают нам не только представлять когнитивные состояния в терминах символьных выражений, систематически зависящих от их интерпретации, но и соотносить систему данных выражений с каузальной структурой физических систем, лежащих в их основании. Кроме того, так как одной из наиболее важных черт вычислительных и когнитивных процессов является, с точки зрения философа, их зависимость от репрезентации своих состояний, следует говорить также и о семантической интерпретационной функции. Таким образом, любое описание когнитивного процесса, характеризующее его как последовательность семантически-интерпретируемых состояний, может быть непосредственно соотнесено с вычислительным описанием данного процесса. Для любых функционально-релевантных семантических состояний существуют соответствующие им

¹¹⁶ Pylyshyn Z. *Computation and Cognition. Toward a Foundation for Cognitive Science*. Cambridge, Massachusetts: Bradford Books/MIT Press, 1984. P. 55-56.

синтаксические состояния, «закодированные» в физически-реализованных символических структурах.

Пилишин, выделяет три автономных уровня описания когнитивных процессов:

- 1) биологический (или физический);
- 2) символический (синтаксический или функциональный);
- 3) семантический (интенциональный).

Несмотря на то, что каждый из приведенных уровней описания обладает собственным теоретическим и терминологическим аппаратом, их взаимодействие четко определено: так, физико-биологические факторы влияют на функциональную архитектуру когнитивной системы, и, как следствие, на ее вычислительные ресурсы. Функциональная архитектура, в свою очередь, ограничивает спектр возможных интенциональных состояний системы (при этом далеко не каждое функциональное состояние обеспечивает появление соответствующих семантико-репрезентативных состояний).

Свой вычислительный подход исследователь, подобно Марру, применил в разрабатываемой им теории зрения (известной также как «теория визуальной индексации»), призванной объяснить механизмы распределения зрительного внимания и познания объектов в пространстве. В рамках данной теории Пилишин выдвигает гипотезу о предконцептуальном механизме, с помощью которого объекты могут быть индивидуализированы, отслежены и закодированы в результате некоторых когнитивных процессов. Философ полагает, что данная гипотеза способна пролить свет на некоторые вопросы, касающиеся особенностей визуального восприятия, а также возможности проектирования нейрокомпьютерных интерфейсов¹¹⁷.

Как нетрудно заметить, дескриптивные стратегии, предложенные Марром и Пилишиным, были призваны продемонстрировать неоднозначность вычислительной интерпретации процессов, осуществляемых когнитивными системами: будучи осведомленными об основных критических программах, направленных против классической вычислительной установки, в своих работах

¹¹⁷ Pylyshyn Z. Seeing and Visualizing: It's Not What You Think. Bradford Press, 2003.

исследователи попытались обозначить условия описания мыслительных процессов, при которых подобная установка могла избежать проблем, которые позже в философской литературе получают название проблемы тривиальности и вычислительной реализации (более подробно данные проблемы будут рассмотрены в Главе 3 настоящего исследования).

1.3. Нейросетевая парадигма в когнитивных науках и философии сознания

Главной альтернативой символическому компьютеризму, долгое время господствовавшему в когнитивных науках и философии, в 1980-х годах стал подход, апеллирующий к объяснительной парадигме нейросетевой теории, получивший название коннекционизм.

Самые ранние нейросетевые модели были основаны на исследованиях нейробиологов Мак-Каллока и Питтса, доказавших способность нейросетей к решению простейших логических задач¹¹⁸. Значительный вклад в развитие коннекционистской концепции внесла Исследовательская группа по изучению параллельной распределенной обработки информации» (PDP Research Group), члены которой активно изучали свойства искусственных сетей нейронов в 1980-х годах. Ее участники, среди которых были такие ученые, как Дэвид МакКлеланд, Дэвид Румелхарт, Джеффри Хинтон и др., предложили новый подход к описанию работы мозга, который, с их точки зрения, был намного убедительней классического компьютеризма, в рамках которого познавательные процессы уподоблялись символическому вычислению цифрового компьютера. Коннекционистский подход к толкованию вычислительных процессов казался многим ученым-нейробиологам более приемлемым, так как не нуждался в строгом делении вычислительной системы на функциональные модули и не уподоблял когнитивные процессы

¹¹⁸ McCulloch W., Pitts W. A Logical Calculus of Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics, 1943. №5. P. 115-133.

неконтинуальным операциям, осуществляемым над набором предзаданных символов. Исследователи также апеллировали к достижениям искусственных нейронных сетей, среди которых моделирование процессов обучения, которые, по их мнению, не способна повторить ни одна машина Тьюринга. По этим причинам некоторые сторонники коннекционизма предложили полностью отказаться от формализма машины Тьюринга и в дальнейшем опираться лишь на концептуальный инструментарий сетевого подхода (к сторонникам подобного «элиминативного» коннекционизма относятся Дэвид Румелхарт, Пол и Патриция Черчленды, Теренс Хорган и др.). Ряд исследователей пошел еще дальше и предложил отталкиваться исключительно от работы мозга и данных нейробиологии, вовсе не изучая работу искусственных вычислительных систем, – этот подход получил название вычислительной нейронауки (computational neuroscience).

Однако некоторые ученые и философы все же предпочли оставить аналогию с цифровым компьютером, чья правомерность зависит, как продемонстрировали Марр и Пилишин, от уровня описания той или иной вычислительной системы: так, нейронное описание может являться низкоуровневым (физическим), а символическое описание – высокоуровневым. Некоторые исследователи предостерегали своих коллег от чрезмерной заинтересованности в изучении работы отдельных нейронов, так как находили подобный уровень объяснения нерелевантным для исследования высших когнитивных функций¹¹⁹. По данной причине были предложены коннекционистские модели, основанные на понятии субсимвольного вычисления, где субсимволы трактуются как более низкоуровневые сущности по отношению к символам, которыми оперирует вычислительная система наподобие машины Тьюринга, но более высокоуровневыми, чем единичные нейроны. Вычисление, с точки зрения ученых, происходит на уровне сетевой структуры, свойства отдельных узлов которой не влияют существенным образом

¹¹⁹ См. например: Smolensky P. On the Proper Treatment of Connectionism // Behavioral and Brain Sciences, 1988. №11. P. 1–74; Marcus G. The Algebraic Mind. Cambridge: MIT Press, 2003.

на общий характер поддерживаемой ими связи. Семантические качества подобной структуры обеспечиваются системой распределенных репрезентаций, состоящих из паттернов активности определенных групп сетевых узлов, в то время как сами эти узлы лишены каких-либо семантических качеств.

Стоит отметить, что основное отличие нейросетевой модели от классического символического компьютеризма заключается отнюдь не в синтаксической архитектуре данных систем (и нейросетевая, и символическая вычислительная система наподобие машины Тьюринга являются универсальными и могут выполнять одинаковые вычисления); скорее, данные модели отличаются тем, как их синтаксические структуры соотносятся с потенциальными семантическими структурами. В классической символической модели предполагается, что обе данные структуры совпадают: каждая базовая репрезентация соответствует определенной синтаксической структуре, а более сложные репрезентации представляют собой совокупность более простых. В коннекционизме, однако, синтаксический и потенциальный семантический уровни не совпадают, и репрезентации появляются на следующем уровне как результат определенной структурной организации базовых элементов вычислительной системы. Обладая сложной структурой, подобные организационные паттерны представляются сторонникам коннекционизма лучшими кандидатами на роль обладателей внутренних семантических свойств по сравнению с отдельными символами или комбинациями символов.

Данный довод, разумеется, оспаривался многими представителями классической вычислительной установки: так, Фодор и Пилишин отмечали, что сетевые репрезентации не обладают должной внутренней структурой, необходимой для появления внутреннего семантического содержания; по их мнению, сетевые паттерны не могут обеспечивать репрезентации необходимым уровнем систематичности, в то время как символические модели могут обладать стабильными репрезентациями в силу своих особых композиционных свойств. Фодор и Пилишин говорят о систематической работе мышления на примере языка: наша способность обрабатывать одно языковое выражение связана со

способностью обрабатывать другие выражения с аналогичной структурой. Например, любой, кто способен понять фразу «Джон любит Мэри», также понимает и фразу «Мэри любит Джона». Представители классического компьютеризма объясняют систематичность тем, что носители языка «вычисляют» значение «Джон любит Мэри», так как знают значения слов «Джон», «любит» и «Мэри», а также правило, позволяющее определить значение фразы по значениям ее составляющих. Если это верно, то способность носителя языка понять фразу «Мэри любит Джона» появляется автоматически. Философы допускают, что коннекционисты могут обучить сети распознаванию как фразы «Джон любит Мэри», так и «Мэри любит Джона», однако, ничто в нейросетевой архитектурной модели не гарантирует, что успех в одной из этих задач влечет за собой успех в другой. А поскольку сама по себе коннекционистская модель не объясняет систематичность, она не является адекватной для объяснения особенностей функционирования нашего языка и мышления¹²⁰.

Однако довод Фодора и Пилишина был впоследствии оспорен многими философами, доказавшими, что нейросетевые модели могут обладать должным уровнем семантической систематичности, которая обеспечивает систему набором релевантных репрезентативных форм¹²¹. Исследователями, в частности, отмечалось, что в коннекционистских моделях информация фиксируется в паттернах единовременной активации всех нейронов. Хотя эти паттерны могут иметь компоненты, такие как активность отдельных нейронов, эти компоненты являются «субсимволическими» в том смысле, что нет очевидного способа присвоить им какое-либо репрезентативное содержание. Распределенная активация, которая происходит для такого сложного выражения, как «Джон любит Мэри», не обязательно должна содержать какое-либо явное представление его частей. Хотя информацию о составляющих «Джон», «любит»

¹²⁰ Fodor J.A., Pylyshyn Z. Connectionism and cognitive architecture: A critical analysis // *Cognition*, 1988. №28. P. 3-71.

¹²¹ Pollack J.B. Recursive distributed representation // *Artificial Intelligence*, 1990. №46. P. 77-105; Chalmers D. Connectionism and Compositionality: Why Fodor and Pylyshyn Were Wrong // *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, 1993. P. 340-347; Aizawa, K. Explaining Systematicity // *Mind and Language*, 1997. №12. P. 115– 136.

и «Мэри» можно обнаружить с помощью изучения паттернов активации сети, нет никаких признаков того, что сеть должна явно извлекать эту информацию сама для выполнения задач обработки языка¹²².

Начиная с середины 1990-х годов рядом исследователей были разработаны новые гибридные сетевые модели, а также альтернативные подходы к описанию когнитивных процессов, основанные на теории динамических систем, в рамках которых вычислительные состояния системы определяются с помощью набора дифференциальных уравнений, описывающих, как изменяется состояние системы во времени. Появление энактивистских концепций, ключевые положения которых сформулировали исследователи Франсиско Варела, Элеонор Рош и Эван Томпсон в своей работе 1991 года «Воплощенный разум»¹²³, способствовало пересмотру ряда положений как классического компьютеризма, так и коннекционизма: настаивая на существовании тесной связи между сознанием агента и опытом его взаимодействия с окружающей средой, сторонники энактивизма отказывалась признавать, что ментальные явления обладают вычислительными и репрезентативными свойствами. Поддерживая идеи так называемого «воплощенного познания», в рамках которого постулируется неразрывное единство разума, сознания и тела, данные исследователи выдвинули ряд характерных возражений против вычислительного подхода к объяснению природы ментальных состояний (суть данных возражений будет подробнее описана в Главе 3 настоящего исследования).

Несмотря на развитие альтернативных когнитивных теорий, коннекционизм не сдал своих позиций: сетевые подходы остаются весьма популярными, и дальнейшие исследования работы естественных и искусственных нейронных сетей представляются необходимыми большинству современных ученых и философов.

¹²² Chalmers D. Syntactic transformations on distributed representations // Connection Science, 1990. №2. P. 53-62.

¹²³ Varela F. J., Thompson E., Rosch E. The Embodied Mind: Cognitive Science and Human Experience. Cambridge, MA: The MIT Press. 1991.

Заключение главы

Появление первых вычислительных теорий сознания стало закономерным следствием становления новых областей математического и естественнонаучного знания, оформившихся в 1940-1950-х годах прошлого столетия. Математическая концепция информации Шеннона и теория вычислимости Тьюринга оказали значительное влияние практически на все научные дисциплины того времени, в результате чего неизбежно привлекли к себе внимание философов, искавших новые концепции, призванные объяснить природу человеческого сознания в терминах, отличных от используемых в рамках бихевиоризма или теории тождества. Информационно-вычислительный поворот послужил основанием для сближения когнитивных наук и философии, быстро включивших в свой терминологический аппарат такие понятия, как «вычисление», «обработка информации», «код», «символ», «сигнал» и т.д.

Классический компьютеризм, уподоблявший мышление вычислительной системе наподобие машины Тьюринга, стал весьма популярным подходом к объяснению мышления, который применяли в своих теориях многие известные философы, оказавшие значительное влияние на развитие всей аналитической философии сознания. Так, Хилари Патнэм, автор теории машинного функционализма, предложил идею множественной реализуемости, ставшую впоследствии ключевой для большинства вычислительных теорий, а также предположил, что определенные ментальные состояния человека сопоставимы с логическими состояниями вычислительной машины. Ученик Патнэма Джерри Фодор в своих работах уделил особое внимание обоснованию возникновения семантических свойств символов, которыми манипулирует вычислительная система, и создал теорию ментальных содержаний, постулирующую существование особого «языка мышления»; сочетая данную теорию с вычислительной установкой, Фодор приходит к выводу, что умственная деятельность включает в себя манипуляции над «языком мышления» в соответствии с определенными механическими правилами.

Философ и когнитивист Зенон Пилишин, тесно сотрудничавший с Фодором, сделал значительный вклад в развитие дескриптивных и эпистемологических установок классического компьютеризма, вслед за нейрочеловеком Дэвидом Марром выделив и представив уровни возможного описания вычислительных процессов, происходящих в результате работы когнитивной системы человека.

Первые вычислительные концепции были далеки от полноценных теорий сознания: большинство из них объясняло лишь некоторые аспекты когнитивной деятельности, такие как память, логический вывод, принятие решений и т.д. Более того, многие исследователи отмечали, что модель, предполагающая сравнение когнитивных процессов с вычислительной машиной наподобие машины Тьюринга, является крайне неправдоподобной с точки зрения биологии. В силу большого количества теоретических проблем, связанных с интерпретацией используемой в рамках классического компьютеризма аналогии мозга и машины Тьюринга, к концу 1980-х годов этот подход стал стремительно терять своих сторонников как среди ученых, так и среди философов. Исследователи начали разрабатывать новые когнитивные теории: так появились первые последовательные нейросетевые вычислительные модели, а также альтернативные концепции, основанные на идеях энантизма и «воплощенного познания».

Глава 2. Наиболее известные информационные теории сознания в аналитической философии

Как уже было отмечено, обращение аналитических философов к информационному словарю в 1960-х годах было во многом обусловлено своеобразной «модой» на информационно-вычислительные подходы к объяснению тех или иных явлений, диктуемой другими научными дисциплинами. Однако далеко не все философы интегрировали конкретные вычислительные формализмы в свои рассуждения о природе мышления и сознания, отдавая предпочтение более абстрактным описаниям, опирающимся на концептуальный аппарат информационной теории.

Далее будут рассмотрены наиболее известные и показательные концепции философов, активно использующих данный концептуальный аппарат и характерное для него выражение «обработка информации» для объяснения природы сознания: это теории Фреда Дретске, Дэниела Деннета, Кеннета Сейра и Д.И. Дубровского. Идеи именно этих мыслителей чаще всего упоминаются в контексте рассуждений об «информационных» концепциях сознания. Однако, так как сам Деннет предпочитает не называть свою теорию «информационной» (о чем философ говорит в некоторых своих интервью¹²⁴), а Кеннет Сейр по отношению к предложенной им концепции упоминает определение «кибернетический», то непосредственно «информационными» чаще именуют лишь теории Дретске и Дубровского. Однако задача настоящего исследования заключается не столько в развешивании терминологических ярлыков, сколько в попытке понять, с какой целью рассматриваемые философы обращались к

¹²⁴ См. например: The Computational View. A Talk with Daniel C. Dennett. 2001. [Электронный ресурс] // Edge.com. URL: https://www.edge.org/conversation/daniel_c_dennett-the-computational-perspective; Uzgalis B. Interview with Daniel Dennett // Minds and Machines. 2006. №16. P. 7–19.

информационной терминологии при описании природы ментальных состояний, – по этой причине в данной главе мы сконцентрируемся именно на этой задаче.

2.1. Информационный репрезентативизм Фреда Дретске

Как уже было продемонстрировано на примере теоретических построений Джерри Фодора, к 1980-м годам одной из наиболее актуальных задач аналитической философии стал поиск релевантной натуралистической теории, призванной объяснить появление сознательного опыта и его содержания в терминах естественных отношений, возникающих между подобным содержанием и окружающей средой. Подобно Фодору попытку адаптировать информационную теорию к описанию механизмов возникновения интенциональных и репрезентативных качеств сознания предпринял ряд других сторонников так называемых каузальных и телеосемантических концепций ментального контента, описывающих природу связи между ментальными содержаниями и некоторой частью внешнего мира¹²⁵. Одним из самых известных представителей данных концепций является американский философ Фред Дретске, построивший репрезентативистскую теорию сознания, основанную на информационном подходе. Желая объединить известные науке данные о свойствах мышления с математической теорией коммуникации, философ поставил перед собой задачу доказать, что данный подход способствует решению многих проблем, встающих на пути у современной когнитивистики.

В своей наиболее известной книге «Знание и поток информации»¹²⁶, увидевшей свет в 1981 году, философ отмечает, что информация нередко путается с такими понятиями, как «значение» или «смысл» (meaning), хотя, с его точки зрения, ее не следует трактовать таким образом, ведь «смысл» принято

¹²⁵ См. например: Tye M. Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind. MIT Press. 1995; Loewer B. From Information to Intentionality // Synthese, №70, 1987. P. 287-317; Millikan R. What Has Natural Information to Do with Intentional Representation? // Naturalism, Evolution and Mind (ed. D.M. Walsh). Cambridge: Cambridge University Press, 2001. P. 105-125; Pineda D. Information and Content // Philosophical Issues, №9, 1998. P. 381-387.

¹²⁶ Dretske F. Knowledge and the Flow of Information. MA: Massachusetts Institute of Technology. 1981.

расценивать как вторичный продукт, непосредственно связанный с деятельностью высокоразвитых когнитивных агентов, в то время как синтаксические информационные структуры присутствуют на любом уровне физической организации. В подобной путанице, с его точки зрения, и заключается проблема, связанная с невозможностью совмещения концептуальных словарей когнитивистики и математических наук. Ее решение Дретске видит в построении своеобразного концептуального моста между данными дисциплинами, основанного на релевантном толковании информационной теории. Для выполнения данной задачи философ предпринял попытку дополнить теорию, предложенную Шенноном, концепцией семантической информации, в рамках которой содержание информационного сообщения трактуется вне зависимости от свойств реципиента данного сообщения. Описывая положения данной концепции, Дретске отмечает, что содержание любого информационного сигнала обладает семантическими свойствами по умолчанию; ответ на вопрос о том, как «количество» информации может соответствовать некоторому «содержанию», заключается в том, что даже синтаксическая информация передается посредством номических зависимостей между конкретными физическими свойствами окружающего мира, и само получение подобной информации уже обязано нести некоторое содержательное знание об этих свойствах¹²⁷.

В своей работе философ предлагает следующее определение содержания информационного сигнала:

¹²⁷ Трактовка информационных состояний как обладающих некоторым естественным смыслом (англ. «natural meaning») заимствуется Дретске из влиятельных работ философа и лингвиста Пола Грайса, предположившего, что естественный смысл включает в себя надежные контрфактические корреляции (к примеру, ствольные кольца дерева коррелируют с его возрастом, - в разговорной речи мы описываем эти кольца как носители информации о возрасте дерева) (Grice H.P. Meaning // The Philosophical Review, 1957. №66. P. 377–388; Grice H.P. Studies in the Ways of Words. Cambridge: Harvard University Press. 1989). Нередко подход Дретске также относят к группе «вероятностных теорий информации», определяющих семантическую информацию в терминах распределения вероятностей и инвертированного отношения между информацией, содержащейся в некотором сообщении/событии X, и вероятностью X. Подобный подход также представлен в работе «Эскиз теории семантической информации» израильского философа Йегошуа Бар-Хиллела, написанной в соавторстве с Рудольфом Карнапом (Bar-Hillel Y., Carnap R. An Outline of a Theory of Semantic Information // Technical Report, 1952. №247).

«Выражение: “Сигнал R несет информацию о том, что S есть F” означает, что условная вероятность того, что S является F, учитывая особенности сигнала R (а также уже имеющуюся у субъекта информацию), равна 1»¹²⁸.

К примеру, для красного света (R), несущего информацию о том, что цель (S) была достигнута (есть F), вероятность того, что цель действительно была достигнута, учитывая, что свет действительно красный и учитывая имеющиеся у нас знания о мире, должна быть 1 (но меньше 1, если бы субъект, интерпретирующий сигнал, исходил исключительно из своих знаний; подобную зависимость информации от знаний получателя информационного сообщения Дретске называет «релятивизацией»)¹²⁹.

Далее, опираясь на подобное представление информационного контента, философ дает характеристику информационно-обусловленному убеждению:

«Если убеждение является видом внутреннего состояния, содержание которого можно выразить как “S есть F”, то оно обусловлено информацией о том, что “S есть F” тогда и только тогда, когда физические свойства сигнала, несущего эту информацию, являются каузально эффективными для производства подобного убеждения»¹³⁰.

С точки зрения философа, сигналы, осуществляющие передачу некоторой информации, должны в обязательном порядке обладать особыми физическими

¹²⁸ Dretske F. Knowledge and the Flow of Information. MA: Massachusetts Institute of Technology. 1981. P. 65.

¹²⁹ Стоит отметить, что современные исследователи относятся к подобной трактовке достаточно критично: так, философ Олимпия Ломбарди отмечает, что предложенный Дретске подход к толкованию семантической информации базируется на интерпретации природы информации, которая сильно отличается от той, что обычно используется в физических науках и эпистемологии (Lombardi O. Dretske, Shannon's theory and the interpretation of information // Synthese. 2005. №144. P. 23–39). Философ Лучано Флориди полагает, что теория семантической информации скорее должна основываться на значениях истинности некоторого информационного сигнала, а не на распределениях вероятностей его содержания (Floridi L. Outline of a theory of strongly semantic information // Minds and Machines. 2004. №14(2). P. 197-221). Исследователи Дэниэл Хутто и Эрик Мин также отмечают, что отстаиваемое Дретске представление об информации как о вероятностной ковариации ничего не говорит об отношении подобной информации к истинному положению дел, а лишь свидетельствует о некоторых каузальных функциях подобного информационного сигнала, и, следовательно, не может способствовать объяснению природы ментальных содержаний (Hutto D., Myin E. Radicalizing Enactivism: Basic Minds without Content. Cambridge, MA: MIT Press, 2013).

¹³⁰ Dretske F. Knowledge and the Flow of Information. MA: Massachusetts Institute of Technology. 1981. P. 87-88.

характеристиками, определяющими степень «видимости» подобных сигналов для данного когнитивного агента: так, восприятие одной и той же информации, передаваемой посредством разных носителей, зависит от свойств данных носителей.

Как только реципиент принимает некоторый информационный сигнал, содержательная составляющая данного сигнала может зафиксироваться, что ведет к появлению стабильного ментального контента. Для описания последовательности подобной фиксации ментального содержания философ прибегает к таким понятиям, как первичная репрезентативная зависимость, цифровая перекодировка и период обучения.

Понятие первичной репрезентативной зависимости используется им для описания причины, в силу которой, к примеру, когнитивный агент слышит свист вскипевшего чайника, а не вибрацию своих барабанных перепонки, являющуюся, однако, также обусловленной этим свистом. Безусловно, восприятие свиста снабжает реципиента информацией сразу о нескольких факторах, и для описания разных информационных уровней, содержащихся в одном сигнале, Дретске вводит понятие информационной вложенности (nested information):

«Сообщение “Т есть G” является вложенным в сообщение “S есть F”, если и только если “S есть F” несет информацию о том, что “Т есть G”, то есть условная вероятность того, что “Т есть G” равна 1, и условная вероятность того, что “S есть F” тоже равна 1 (при заданном сигнале R)»¹³¹.

Таким образом когнитивный агент получает сразу несколько уровней информации, представленных в рамках всего одного информационного сигнала. Однако почему ментальное состояние агента свидетельствует именно о свисте вскипевшего чайника, ведь сам по себе подобный сигнал несет и другую информацию?

¹³¹ Там же. Р. 71.

После получения реципиентом некоторого сигнала для изъятия из него релевантного информационного содержания происходит его кодовое преобразование, призванное очистить сигнал от лишних содержательных составляющих и шумовых помех. Философ отмечает, что любой сигнал может существовать в аналоговой или цифровой форме: в случае, если сигнал несет в себе дополнительную («вложенную») информацию, он представлен в аналоговой форме; если сигнал не несет подобной дополнительной информации, то он представлен в цифровой форме. Данное различие можно проиллюстрировать с помощью следующего примера. Допустим, что на столе перед неким молодым человеком стоит чашка чая. Желая передать информацию об этом своему знакомому, молодой человек говорит ему: «Чашка чая стоит на столе». Из этой фразы не ясно, как выглядит чашка, какой в ней чай, как много его налито и т.д., поэтому знакомый получает сигнал вида «S есть F» без дополнительной информации. Однако, если бы молодой человек вместо того, чтобы озвучивать подобную фразу, отправил бы своему знакомому фотографию стоящей на столе чашки чая, тот бы получил большое количество «вложенной информации» (например, информацию о форме стола, материале, из которого сделана чашка, крепости чая и т.д.). Подобным образом когнитивный агент, получающий тот или иной информационный сигнал, также получает множество дополнительной и ненужной информации, которая затем подвергается соответствующей перекодировке из аналогового вида в цифровой и предстает в релевантной форме. Данное различие между аналоговой и цифровой информационными формами лежит, по мнению философа, в основе различия между процессами восприятия и мышления. Восприятие является, по Дретске, процессом, в рамках которого информация поступает в более богатой (то есть аналоговой) форме, а затем она подвергается обработке соответствующими когнитивными центрами. Мышление же представляет собой более высокоуровневый процесс, предполагающий, что лежащие в его основе когнитивные механизмы оперируют исключительно информацией, подвергшейся значительной цифровой перекодировке.

Ключевой особенностью высокоразвитых организмов является, согласно Дретске, наличие особой функциональной организации, отвечающей за изощренные способы обработки и использования получаемой информации, что, в свою очередь, влияет на уровень интенциональности; основная задача когнитивных агентов состоит в том, чтобы перейти от большого количества нерелевантной информации, содержащейся в любом информационном сигнале, к его однозначному семантическому толкованию. Именно во время периода обучения эти агенты находят наиболее эффективный способ перекодировки поступающего информационного сигнала, способствующий появлению определенного семантического содержания:

«Предположим, что во время периода обучения система развивает способ цифровой перекодировки информации о том, что что-то является F: развивается определенный тип внутреннего состояния, которое избирательно чувствительно к информации о том, что что-то есть F. Пока эта структура развивается, она приобретает, так сказать, свою собственную жизнь и становится способной наделять себя последующими знаками и смысловым наполнением (тем содержанием, которое она приобрела в процессе обучения) независимо от того, имеет ли оно в действительности свое информационное наполнение»¹³².

Так философ разъясняет, каким образом содержание того или иного ментального состояния возникает вне зависимости от истинного положения дел, хотя само информационное сообщение не может быть ложным по определению (иными словами, каким образом мы можем увидеть оазис в пустыне, когда на самом деле это всего лишь мираж). (Отметим, что далеко не все философы были удовлетворены подобным объяснением: так, Джерри Фодор высказал предположение, что в рамках теории Дретске появление ошибочной репрезентации возможно только во время обучения, однако это противоречит нашему опыту, ведь даже когда человек научился отличать, скажем, собаку от

¹³² Там же. Р. 193.

лисы, он все равно может их перепутать в силу каких-либо обстоятельств. Более того, философ полагает, что концепция Дретске не учитывает намерений условного «учителя», чья роль в процессе обучения является, по его мнению, решающей¹³³).

Отдельный вопрос касается связи информационной теории Дретске с вычислительными теориями, существующими в когнитивных науках: некоторые философы полагают, что предлагаемая философом концепция цифровой перекодировки не может трактоваться как вычисление. Так, Михаэль Раскорла отмечает, что обыкновенный термостат, «обрабатывающий информацию» о температуре окружающей среды, не реализует никакой нетривиальной вычислительной модели, из чего следует, что система может обрабатывать информацию в стиле Дретске без выполнения специфических вычислений¹³⁴. Хотя сам Дретске говорит о «более изощренных» способах кодирования и обработки информации, присущих высокоразвитым когнитивным агентам, он нигде не поясняет, как именно происходят подобные процессы информационной обработки и основываются ли они на классической или нейросетевой вычислительных моделях, разрабатываемых в рамках современных ему наук. По этой причине предложенную философом трактовку следует считать самобытной интерпретацией процессов информационной обработки, не основанной на концептуальных схемах, развиваемых в рамках современных ему когнитивных наук.

Представив основы натуралистического подхода к толкованию семантических свойств информационных сигналов, а также описав связь информационных и интенциональных состояний, Дретске приступает к поиску ответов на другие вопросы, касающиеся работы когнитивной системы человека. В своей следующей книге, «Объясняя поведение»¹³⁵, философ адаптирует информационную теорию к решению некоторых вопросов, связанных с

¹³³ Fodor J. Semantics, *Wisconsin Style* // *Synthese*, 1984. №59. P. 231–250.

¹³⁴ Rescorla M. The Computational Theory of Mind // *The Stanford Encyclopedia of Philosophy* (N. Zalta ed.), 2019. [Электронный ресурс]. URL: <https://plato.stanford.edu/entries/computational-mind/#AltConCom> (дата обращения: 04.12.2020)

¹³⁵ Dretske F. *Explaining Behavior: Reasons in a World of Causes*. Cambridge, Massachusetts The MIT Press, 1988.

проблемой ментальной каузальности. Дретске предполагает, что любое убеждение вида «S есть F» коррелирует с состоянием мозга, которое отвечает за часть процессов, вызывающих, например, движение, в силу того, что данное физическое состояние несет информацию «S есть F». Убеждения, таким образом, являются ментальными репрезентациями, которые способствуют производству движения в силу своего информационного содержания (при этом философ отмечает, что состояния мозга, отвечающие за передачу информации, выполняют свою функцию даже если они больше не несут информацию – так Дретске вновь объясняет возникновение ошибок, связанных с искажениями и ложной информацией)¹³⁶.

Вопросам, касающимся интерпретации прочих аспектов ментальной деятельности, Дретске посвящает книгу «Натурализируя сознание», в которой он формулирует так называемый «репрезентативистский тезис», основанный на следующих двух посылках:

- 1) Все ментальные факты являются фактами репрезентаций.
- 2) Все факты репрезентаций являются фактами об информационных функциях¹³⁷.

Функции, отвечающие за передачу и обработку той или иной информации, Дретске разделяет на систематические и приобретенные: первые развиваются в результате филогенетической эволюции организмов и передаются по наследству, в то время как вторые появляются в ходе индивидуального развития

¹³⁶ Дретске предлагает концептуально разделять результат поведения (например, двигательную активность) и само поведение, понимаемое как процесс, каузально вызывающий какую-либо двигательную активность в силу наличия у рассматриваемого организма некоторого внутреннего состояния. Следом за этим разделением Дретске предлагает также разделять и два типа каузального объяснения: объяснение конкретных событий (например, некоторого движения) и объяснение процесса как череды взаимозависимых событий – то, что философ называет объяснением с помощью структурных причин (explanation by structuring causes). Согласно Дретске, объяснение результата поведения (скажем, подъема руки) будет всегда подразумевать эпифеноменальность ментальных свойств; однако, когда мы говорим о непосредственном поведении, а не его результатах, мы полагаем наличие у него некой структурной каузальной действенности. К примеру, мы можем задать следующий вопрос: почему именно данный вид нейронной активности вызывает подъем руки? Для ответа на подобный вопрос, полагает Дретске, мы и апеллируем к ментальным свойствам, которые являются структурными причинами поведения в целом.

Концепцию структурной каузальности Дретске продолжил развивать и в своих последующих работах (Dretske F. Mental events as structuring causes of behavior // *Mental Causation* (eds. J. Heil, A. R. Mele). Oxford University Press, 1993. P. 120-145; Dretske F. Mental causation // *The Proceedings of the Twentieth World Congress of Philosophy* (ed. K. A. Stoehr). Bowling Green: Philosophy Doc Ctr. 1999. P. 81-88).

¹³⁷ Dretske F. *Naturalizing the Mind*. Cambridge, Massachusetts: The MIT Press, 1995. P. 13.

субъекта. На этом различии основывается деление ментальных репрезентаций на неконцептуальные (ощущения и сенсорные переживания, которые Дретске соотносит с квалиа), являющиеся фактами систематических информационных функций, и концептуальные (убеждения, знания), являющиеся, соответственно, фактами приобретенных информационных функций. С помощью подобной категоризации философ пытается продемонстрировать, что его экстерналистский подход к толкованию интенциональных состояний распространяется и на качественный опыт, который, вопреки интерналистской интуиции, полностью зависит от истории каузального взаимодействия организма с окружающей его средой. Подобный опыт, таким образом, редуцируется к ментальным репрезентациям, которые, в свою очередь, описываются в функциональных терминах, что дает основания называть концепцию Дретске редуктивным репрезентативизмом.

В основе главного возражения против подобной позиции лежит указание на тот факт, что информационно-теоретическое описание естественных каузальных и телеофункциональных отношений, возникающих между свойствами объектов внешнего мира и репрезентирующими их интенциональными состояниями когнитивного агента недостаточно для должного объяснения некоторых качественных аспектов ментального опыта. Даже если допустить, что философу удалось объяснить механизм возникновения сенсорных ощущений в терминах неконцептуальных репрезентаций, то феноменальные переживания, трактуемые как «каково это находиться в некотором ментальном состоянии», все еще остаются без должного объяснения в рамках предложенной им концепции (по этой причине некоторые философы ставят Дретске в один ряд с представителями иллюзионизма по отношению к феноменальным свойствам сознания¹³⁸). Более того, исследователи отмечают, что в ряде случаев наше понимание интенциональности так или иначе

¹³⁸ См. например: Lycan W. Representational Theories of Consciousness // The Stanford Encyclopedia of Philosophy (ed. by E.N. Zalta). 2019. [Электронный ресурс] URL: <https://plato.stanford.edu/archives/fall2019/entries/consciousness-representational/> (дата обращения: 18.10.2021).

вынуждено апеллировать к феноменальным категориям, что означает принципиальную невозможность редукции феноменальных переживаний к ментальным репрезентациям¹³⁹.

Таким образом, репрезентативизм Дретске, основанный на идеях информационной телеосемантики, сталкивается с целым рядом концептуальных трудностей и, следовательно, вряд ли способен стать основанием для последовательной философской теории сознания. Тем не менее, несмотря на описанные выше сложности и возражения, не будет преувеличением сказать, что предложенная Дретске концепция оказала значительное влияние на многих аналитических философов, среди которых Дэниел Деннет, Дэвид Чалмерс, Лучано Флориди, а также на исследователей, продолжающих развивать информационные теории ментальных содержаний с опорой на данные современных когнитивных наук¹⁴⁰.

2.2. Информационно-вычислительная метафора в философии сознания Дэниела Деннета

Американский философ Дэниел Деннет всегда высоко ценил объяснительный потенциал современных когнитивных наук, а также подробно изучал подходы к объяснению природы сознания, основанные на той или иной версии вычислительной установки. Так, в 1986 году Деннет посвятил рассуждениям над основаниями вычислительных теорий отдельную статью, в

¹³⁹ Chalmers D.J. The Representational Character of Experience // The Future for Philosophy (ed. by B. Leiter). Oxford University Press. 2004. P. 165.

¹⁴⁰ Среди них такие выдающиеся ученые и философы, как Крис Элиасмит (Eliasmith C. A New Perspective on the Representational Problems // Journal of Cognitive Science. 2005. №6. P. 97-123.), Карен Неандер (Neander K. A Mark of the Mental: In Defence of Informational Teleosemantics. Cambridge, MIT Press. 2017.), Брайан Скримс (Skyrms B. Signals: Evolution, Learning and Information. Oxford University Press. 2010.), Ульрих Стегман (Stegman U. Prospects for Probabilistic Theories of Natural Information // Erkenntnis. 2015. №80. P. 869-893.) и Мариус Ушер (Usher M. A Statistical Referential Theory of Content: Using Information Theory to Account for Misrepresentation // Mind and Language. 2001. №16(3). P. 331-334).

которой выделил и описал три основные догмы, лежащие, по его мнению, в основе позиции классического компьютеризма:

«1) Мышление является процессом обработки информации. Это означает, что само мышление может быть проанализировано в виде совокупности процессов ввода/вывода, решения задач, поиска и т.д.

2) Обработка информации является вычислительным процессом, подразумевающим работу с символами. Системы, производящие обработку информации, могут быть представлены в терминах “вычисления”. Ментальные репрезентации, в свою очередь, могут быть описаны как состоящие из символов, принадлежащих системе, обладающей синтаксисом и набором формальных правил для манипулирования подобными символами.

3) Семантика данных символов соединяет мышление с окружающим миром. Другими словами, существует определенный язык мышления, большинство понятий которого могут быть соотнесены с объектами из внешнего мира»¹⁴¹.

Деннет отмечает, что идеи классического компьютеризма представляются ему непригодными для описания работы когнитивной системы человека, и по этой причине он отдает предпочтение объяснительным моделям коннекционистского типа, которые, однако, по выражению философа, нацелены на анализ работы мозга, а не мышления, и потому нуждаются в более высоком уровне описания, на котором мы можем приписывать мозгу внешние семантические свойства. Деннет предлагает продолжать исследования в области компьютерной науки и искусственного интеллекта, однако отказаться от строго формализованной вычислительной модели, переложив ее в форму мысленных экспериментов. Данную исследовательскую установку, которую, по

¹⁴¹ Dennet D.C. The logical geography of computational approaches: A view from the east pole // The Representation of Knowledge and Belief (eds. M. Brand, R. M. Harnish). University of Arizona Press, 1986. P. 60.

утверждению философа, он позаимствовал у Дэвида Марра, Деннет реализовал в рамках своего подхода к изучению феномена сознания.

Свою теорию философ начинает строить с концепции ментального содержания (как позже он отметит в «Объясненном сознании»: *«Моя главная стратегия всегда заключалась в следующем: сначала объяснить ментальное содержание как нечто независимое и фундаментальное по отношению к сознанию [...] затем построить на этой базе объяснение сознания»*¹⁴²). В отличие от Дретске, предложившего натуралистическую интерпретацию информационной семантики, Деннет полагает, что информационные сигналы обладают некоторым содержанием только благодаря интерпретационной деятельности внешнего наблюдателя; а так как всякая интенциональная система, с точки зрения философа, необходимым образом должна оперировать некоторой информацией, то ее описание также основывается на интерпретационных установках наблюдателей. Руководствуясь паттернами поведения тех или иных систем, мы можем приписывать им обладание некоторыми ментальными состояниями, что, в свою очередь, является достаточно эффективным инструментом для осуществления прогнозов и предсказаний относительно их дальнейшего поведения.

Строя свою теорию на учении о трех установках, – физической, проективной и интенциональной, – Деннет отмечает, что действия искусственных систем, наподобие компьютеров, представляющих собой особого рода информационные хранилища, могут описываться с помощью любой из этих установок (отсюда следует, как отмечают некоторые исследователи, что Деннет является сторонником идеи сильного ИИ, ведь в рамках его концепции информационные процессы, происходящие как в естественных, так и в искусственных системах, можно представить с позиции интенциональной

¹⁴² Dennett D.C. Consciousness explained. Boston, MA. 1991. P. 457.

установки¹⁴³; в то же время, некоторым ученым подобная установка кажется неправдоподобной¹⁴⁴):

«Безусловно, подобное хранилище может быть названо информационным только в силу того факта, что некоторые пользователи смогут интерпретировать полученную информацию. Мы должны оставить термин “разумное хранилище” для хранилища такой информации, которая будет значима для самой системы, а не только для пользователей и создателей системы. Критерием разумности хранилища будет удовлетворительность итогового поведения системы для ее нужд при заданных стимулирующих условиях, изначальных входных данных и заданной окружающей среде, в которой проявляется поведение»^{145, 146}.

Философ отмечает, что объяснение функционирования информационно-обусловленных когнитивных процессов с опорой на положения семантического реализма равносильно признанию существования своеобразного гомункула, «понимающего» содержание данных функциональных состояний, что, в свою очередь, порождает вопрос о том, как сам гомункул понимает содержание, – подобный довод вел к бесконечному постулированию гомункулов внутри гомункулов. *«Мозг, – пишет Деннет, – как показывают теория интенциональных систем и эволюционная биология, является семантическим инструментом: его задача заключается в определении смысла различных*

¹⁴³ Волков Д.Б. Бостонский зомби: Д. Деннет и его теория сознания. М.: Книжный дом «ЛИБРОКОМ», 2012. С.70.

¹⁴⁴ Nanay B. Symmetry between the intentionality of minds and machines? The biological plausibility of Dennett's account // *Minds and Machines*. 2006. №16. P. 57–71.

¹⁴⁵ Dennett D.C. *Content and consciousness*. NY, 1969. P. 45.

¹⁴⁶ Подобный подход нередко сравнивают с теорией британского физика и нейрочеловека Д.М. МакКея, который рассматривал целенаправленные системы как обладающие входными данными, полученными из окружающей среды, и репрезентирующие состояние этой среды в данный момент, а также данными, совмещающими эту репрезентацию с целями системы. МакКей представлял информацию в терминах ее возможности оказывать влияние на «потенциал готовности» организма, и поэтому его можно считать одним из первых теоретиков «прагматической» информации (MacKay D. M. *Information, Mechanism, and Meaning*. Cambridge: MIT Press, 1970). Сам Деннет нередко ссылается на работы МакКея, отмечая, что многие из сделанных ученым наблюдений, касающихся трактовки информационных состояний, являются весьма мудрыми (Dennett D.C. *From Bacteria to Bach and Back: The Evolution of Minds*. NY: W.W. Norton Company, 2017. P. 217).

событий, в этом его предназначение. Но мозг, как подсказывает нам психология и здравый смысл, является просто синтаксическим инструментом»¹⁴⁷. По мнению Деннета, сложные когнитивные системы решают задачи, используя случайные соответствия между структурными закономерностями и усвоенными ими семантическими установками, чье содержание исходит от внешнего наблюдателя, приписывающего системе те или иные состояния с целью предсказания ее поведения¹⁴⁸.

Объяснению появления и функционирования сознания Деннет посвятил последующие работы: так, философ прибегнул к использованию информационно-вычислительной метафоры в теории, известной как модель множественных набросков (multiple drafts model), впервые представленной им в 1981 году в книге «Мозговые штурмы»¹⁴⁹ и более подробно описанной в «Объясненном сознании». Так как своеобразное «картезианство», сводящееся к необходимости выявить и описать когнитивные центры или модули, обеспечивающие объединение информационных процессов в единый «поток сознания», представлялось философу общей чертой большинства современных ему философских и когнитивных теорий, Деннет предпринял попытку представить концепцию, призванную стать альтернативой представлению, известному как «Картезианский театр», и снять проблему гомункула и бесконечного регресса в объяснении. Согласно модели множественных набросков, когнитивная структура представляется в виде конкурирующей неиерархической вычислительной архитектуры: *«Все разновидности восприятия, – пишет философ, – как и все разновидности мышления или умственной деятельности, осуществляются в мозгу параллельными, многодорожечными процессами интерпретации и обработки сенсорных входов.*

¹⁴⁷ Dennett D.C. The intentional stance. A Bradford Book, The MIT Press Cambridge, Massachusetts London, 1989. P. 61.

¹⁴⁸ Отметим, что подобная трактовка ментальных содержаний оказала существенное влияние на многих сторонников вычислительных подходов к объяснению ментального опыта, среди которых Джон Хаугеланд (Haugeland J. Artificial Intelligence: The Very Idea. Cambridge, MA: MIT Press, 1985.), Роберт Камминс (Cummins R. Meaning and Mental Representation. Cambridge, MA: MIT Press, 1989.) и Патриция Черчленд (Churchland P.M., Sejnowski T.J. The Computational Brain. Cambridge, MA: MIT Press, 1992.).

¹⁴⁹ Dennett D.C. Brainstorms: Philosophical Essays on Mind and Psychology. MIT Press, 1981.

Информация, поступающая в нервную систему, постоянно подвергается редакторской правке»¹⁵⁰. Отмечая, что все информационные процессы происходят параллельно, что мешает четкому определению, какой именно из них будет осознанным в тот или иной момент времени, Деннет отрицает наличие какой-либо четкой границы, отделяющей осознанные переживания от неосознанных, равно как и возможность определения их содержания. Иллюзия единого потока сознания возникает в результате осуществления так называемых «проб» информационных процессов, необходимых для формирования у когнитивного агента суждений о собственных ментальных состояниях.

Роль своеобразного «бутылочного горлышка», фильтрующего сознательные состояния, выполняет в теории философа компьютерная метафора. Определяя человеческое сознание как «виртуальную машину мемов», Деннет сравнивает ее с программой, которую устанавливают в компьютер:

«Человеческое сознание представляет собой огромный комплекс мемов (или, точнее, мем-эффектов в мозге), которые лучше всего можно понять как работу “фоннейманновской” виртуальной машины, включенной в параллельную архитектуру мозга [...] Возможности этой виртуальной машины значительно усиливают основные возможности органического “жесткого диска”, на котором она работает, но в то же время многие из ее наиболее любопытных особенностей, в том числе ее ограничения, могут быть объяснены как побочные продукты»¹⁵¹.

Сознающий человеческий ум представляет собой, по Деннету, последовательную виртуальную машину, или, проще говоря, «программу», являющуюся контрольной управляющей инстанцией, основанной на оборудовании, предоставленном нам эволюцией. Именно наличие подобной машины и обуславливает наличие сознания.

¹⁵⁰ Dennett D.C. *Consciousness Explained*. Back Bay Books, 1992. P. 111.

¹⁵¹ Там же. С. 210.

Безусловно, предложенная Деннетом концепция виртуальной машины сразу стала предметом активной дискуссии. Так, известный сторонник коннекционизма, когнитивист Пол Черчленд критиковал используемую Деннетом аналогию мозга с классической вычислительной машиной, обладающей дискретными состояниями, которая, по мнению ученого, значительно уступает в своей объяснительной силе модели рекурсивных нейронных сетей¹⁵². Деннет, однако, отмечает, что представленная им модель является исключительно абстрактной и ее использование необходимо для более высокоуровневого описания вычислительной системы, обладающей нейросетевой структурой¹⁵³, что, как уже было отмечено, соответствует его методологической установке, кратко описанной в статье 1986 года. Ответ Деннета демонстрирует, что в рамках своей теории он использует информационно-вычислительное описание как описание по аналогии, не уделяя при этом особого внимания трактовке лежащих в его основе вычислительных формализмов, а лишь применяя концептуальный аппарат информационной и вычислительных теорий для наглядности¹⁵⁴.

В своих последующих работах Деннет заменяет модель множественных набросков другой моделью, известной как «слава в мозге» («fame in the brain»), основная идея которой близка к идее, выраженной в теории глобального рабочего пространства Бернарда Баарса¹⁵⁵, в рамках которой сознательный опыт описывается как результат пребывания тех или иных информационно-вычислительных сигналов в пространстве рабочей памяти. Ряд исследователей полагает, что подобная модель весьма напоминает уже упомянутую метафору

¹⁵² Churchland P.M. Catching Consciousness in Recurrent Net // Daniel Dennett: Contemporary Philosophy in Focus (ed. A. Brook, D. Ross). Cambridge University Press, 2002. P.79.

¹⁵³ Dennett D.C. Two steps closer on consciousness // Paul Churchland (ed. by Brian L. Keeley). Cambridge University Press, 2005.

¹⁵⁴ В одной из своих работ философ также отмечает: «Важно понимать, что компьютерные аналогии на самом деле являются философскими мысленными экспериментами, помпами для интуиции, а не эмпирическими экспериментами. Они призваны систематически исследовать последствия наших предположений. Философы привыкли проводить свои мысленные эксперименты вручную, по одному за раз. Теперь они могут выполнять тысячи вариаций за час, что является хорошим способом проверки того, что интуиция, которую они накачивают, не является артефактом какой-то произвольной особенности сценария мысленного эксперимента» (Dennett D. Freedom evolves. Allen Lane Publishers. 2003. P. 218).

¹⁵⁵ Baars B. In the Theater of Consciousness: The Workspace of the Mind. NY: Oxford University Press, 1997.

«Картезианского театра», что, в свою очередь знаменует отход Деннета от принципов, постулируемых в рамках более радикальной модели множественных набросков; сам Деннет, однако, считает, что переход к модели «слава в мозге» знаменует собой лишь смену метафор и аналогий, но не смену парадигмы¹⁵⁶.

Пожалуй, наиболее наглядно информационный подход к объяснению природы сознания представлен в одной из последних работ Деннета «От бактерии к Баху и обратно», в которой он обобщает результаты большинства своих предыдущих исследований. В данной книге философ вводит такие понятия, как «компетентность» («competence») и «понимание» («comprehension»), отмечая при этом, что некоторые системы (будь то бактерии или, например, роботы) могут обладать «компетентностью без понимания», что выражается в их способности выполнять сложные действия без понимания соответствующих причин и целей данных действий; способность к пониманию возникает лишь в процессе эволюции и естественного отбора. Таким образом, Деннет постулирует разницу между условным роботом (машиной Тьюринга или машиной фон Неймана) и живым «сознательным» организмом, и почти полностью отказывается от компьютерной метафоры в пользу более абстрактного информационно-ориентированного объяснения сознания.

Очевидным образом вопрос о переходе от «компетентности» к «пониманию» зависит от семантических информационных структур, которыми оперирует некий агент. Рассуждая над сущностью понятия семантической информации, которым оперирует современная биология, философ пишет следующее:

«Использование термина “информация”, встречающегося повсеместно в когнитивных науках (и не только), не означает информацию Шеннона. Пока еще нет единой схемы кодирования – не обязательно бинарной (0 и 1), – которая оцифровывает совокупность вероятностных распределений, а также нет и

¹⁵⁶ Подробнее об этой дискуссии см.: Волков Д.Б. Бостонский зомби: Д. Деннет и его теория сознания. М.: Книжный дом «ЛИБРОКОМ», 2012. С. 129-154.

способа измерить количество информации. Когда-нибудь в будущем мы можем обнаружить, что существует естественная интерпретация передачи [информации] в нервной системе, которая дает меру пропускной способности или емкости памяти в битах, а также меру кодирования всего, что передается, обрабатывается и хранится, но до тех пор понятие информации, которое мы используем в когнитивной науке, – это семантическая информация, то есть информация, идентифицируемая как нечто конкретное: лица, места, или, например, глюкоза»¹⁵⁷.

Семантическую информацию философ метафорически определяет как «дизайн, который стоит заполучить» («design worth getting»), подразумевая, что манипулирование семантической информацией способствует более эффективному использованию когнитивных структур посредством адаптации и преобразования их частей; Деннет также активно использует выражение «различие, производящее различие», отмечая значимость подобного различия для организма, который в процессе эволюции научился преобразовывать «шум» в «сигнал»¹⁵⁸. Таким образом, семантическая информация является относительной и, следовательно, неизмеримой, и не может быть закодирована, передана или сохранена; такое понятие как «дезинформация» означает лишь целенаправленную эксплуатацию одним организмом системы значимых различий другого организма в своих интересах (как, например, это делает вирус,

¹⁵⁷ Dennett D.C. From Bacteria to Bach and Back: The Evolution of Minds. NY: W.W. Norton Company, 2017. P. 208.

¹⁵⁸ До сих пор точно не известно, кому принадлежит авторство выражения «различие, производящее различие»: обычно его приписывают британско-американскому ученому Грегори Бейтсону, активно использовавшему данное определение в своих работах 1970-80-х годов, однако некоторые исследователи полагают, что впервые данное выражение употребил Д.М. МакКей. Например, Лучано Флориди утверждает, что МакКей определял информацию как «distinction that makes a difference», в то время как Бейтсон несколько видоизменил данную фразу, сформулировав ее как «difference that makes a difference» (Floridi L. Information: A Very Short Introduction. Oxford: Oxford University Press, 2010. P. 24). Впоследствии именно определение Бейтсона стало популярным в науке и философии (отметим, что в своей книге Деннет употребляет определение, данное МакКеем).

маскирующийся под клетки организма, в которые он внедряет свою ДНК), и никак не относится к такой категории, как «истина»¹⁵⁹.

Сознание же является, по Деннету, «иллюзией пользователя». Эволюция культурных «мемов», а также необходимость коммуникации с другими индивидами обеспечивает условия, при которых те или иные информационные процессы становятся «видимыми». По мнению философа, наша точка зрения «от первого лица» никак не отличается от нашей точки зрения «от третьего лица»: нам не доступна сложная работа нейронных механизмов, и мы вынуждены довольствоваться интерпретированной пользовательской иллюзией, которая настолько нам знакома, что мы принимаем ее за реальность; хотя нейробиология пытается объяснить, как релевантная информация обрабатывается нашим мозгом, это еще не значит, что данная информация реально имеет некоторое воплощение в виде феноменальных переживаний.

Объяснив сознание таким образом, Деннет обрек свою теорию на жесткую критику со стороны философов, отрицающих возможность подобной «дисквалификации квалиа». Некоторые исследователи полагают, что информационно-вычислительная метафора – это единственное, что спасает концепцию Деннета от жесткого бихевиористского прочтения¹⁶⁰. Более того, понятие «иллюзия», столь часто используемое Деннетом, является неоднозначным, и требуется дополнительный анализ, чтобы показать, к примеру, как возникают столь убедительные иллюзии феноменальных

¹⁵⁹ Данные положения философ резюмирует в следующих тезисах:

1. Семантическая информация является ценной: недостаток информации и дезинформация представляют собой либо патологии, либо паразитические искажения стандартов;
 2. Ценность семантической информации относительна и не может быть измерена никаким объективным способом, хотя может быть подтверждена эмпирическим путем;
 3. Объем семантической информации, передаваемой или содержащейся в любом ограниченном предмете, также не может быть измерен в каких-либо единицах, но поддается сравнительной оценке в конкретных условиях;
 4. Семантическую информацию не нужно кодировать для передачи или хранения».
- (Dennett D.C. From Bacteria to Bach and Back: The Evolution of Minds. NY: W.W. Norton Company, 2017. P. 126).

Стоит отметить, что некоторые исследователи оспаривают предложенную философом трактовку семантической информации, указывая, в частности, на ложность последних двух тезисов (Miłkowski M. Thinking about Semantic Information. Avant. 2019. №11(2). P. 1-10).

¹⁶⁰ Юлина Н. С. Сознание объясненное // Энциклопедия эпистемологии и философии науки. Сост. и общ. ред. И. Т. Касавина. Москва: «Канон+», 2009. С. 898.

переживаний (некоторые философы отмечают, что Деннет не решает проблему сознания, а заменяет ее на не менее сложную «проблему иллюзии»¹⁶¹). Эти и ряд других концептуальных трудностей, характерных для большинства иллюзионистских теорий сознания, являются причиной, по которой позиция Деннета, несмотря на свое широкое признание в кругу представителей когнитивных наук, не удовлетворяет многих философов, предпочитающих относиться к феноменальному сознанию «всерьез».

Отдельная группа замечаний к концепции Деннета состоит из критики его методологической установки, которая, как было уже не раз отмечено, заключается в метафорическом использовании ряда выражений и терминов, подчерпнутых философом из теории информации, вычислительной нейробиологии и компьютерных наук. Хотя некоторые сторонники развиваемого философом подхода поддерживают стратегию использования так называемых «помп для интуиции», содержащих такие выражения, как «виртуальная машина», «мем-эффект», «пробы», «дизайн, который стоит заполучить», «информационные наброски» и т.д., некоторые из подобных абстрактных аллегорий представляются достаточно сомнительными с точки зрения своей полезности для объяснения сознания, так как они затрудняют сопоставление представлений Деннета с более конкретными естественнонаучными теориями и делают некоторые положения его концепции принципиально неverifiedируемыми. Сам Деннет полагает, что на сегодняшний день подобное использование информационных и вычислительных метафор является наиболее полезным философским инструментом, призванным сгладить столь явные границы между нашими привычными представлениями о сущности сознательного опыта и его научным объяснением. Однако можно задаться вопросом о том, как долго подобные метафоры будут оставаться эффективными. Так, американский историк философии Джон Дорболо отмечает:

¹⁶¹ Frankish K. Illusionism as a Theory of Consciousness // *Journal of Consciousness Studies*. 2016. №23 (11-12). P. 11-39.

«Философия действительно переживает вычислительный поворот. По мере того как завершится этот поворот, будет проверено и предположение Деннета о том, что философия будет неуклонно терять свою загадочность. Если он прав, то по мере того, как компьютерные технологии будут становиться все более необходимыми для философской методологии, мы также обнаружим, что сила философских интуиций уменьшается. Когда эти интуиции ослабнут, исчезнут и загадки, связанные с такими понятиями, как сознание, свобода, намерение, творчество и интеллект. Если это так, то философия претерпевает самые значительные изменения со времен “золотого века” древнегреческой философии. Однако, если Деннет ошибается, то механицистам и редукционистам придется подождать следующей волны технического прогресса, чтобы сформулировать новый метафорический вызов природе разума»¹⁶².

Весьма вероятно, что информационно-вычислительные метафоры, применяемые в целях описания тех или иных явлений, связанных с работой сознания, могут быть вскоре вытеснены другими, более наглядными аналогиями, из чего следует, что теория Деннета, чья оригинальность во многом обусловлена использованием подобных метафор, через некоторое время может быть признана устаревшей и нерелевантной. (Возможно именно поэтому в своих последних работах американский философ предпочел отказаться от ряда применяемых им ранее компьютерных аналогий в пользу более абстрактных информационно-вычислительных описаний¹⁶³). Отсюда следует, что информационная теория

¹⁶² Dorbolo J. Intuition Pumps // Minds and Machines. 2006. №16(1). P. 86.

¹⁶³ Данный поворот словно иллюстрирует идеи исследовательницы Маргарет Боден, отстаивающей правомерность использования вычислительного объяснения в рамках современных когнитивных наук. Боден полагает, что компьютеры представляют собой версию бытовой метафоры, и исследователи, выступающие с критикой подобной метафоры, вполне справедливо предполагают, что она вскоре будет вытеснена другими аналогиями, основанными на терминологии следующего технологического поколения: «Если под “компьютерами” критики имеют в виду сегодняшние ИИ технологии, то они в чем-то правы. В конце концов, современные вычислительные концепции и вычислительные машины не снабдили нас ответами на все вопросы когнитивной науки. (Тридцать лет назад правдоподобных/многообещающих ответов было еще

сознания в идеале должна опираться на такие эпистемологические установки, которые бы делали информационную интерпретацию более устойчивой к изменениям, неизбежно происходящим в результате смены наших научно-технических практик.

2.3. Кибернетическая теория сознания Кеннета Сейра

Особый вклад в развитие информационного подхода к решению проблемы сознания внес американский философ Кеннет Сейр, автор большого количества книг и статей, посвященных философскому осмыслению кибернетики, теории искусственного интеллекта и вычислительной нейронауки. Получив не только философское, но также и математическое образование в Гарвардском университете к 1958 году, Сейр устроился на должность профессора философии в Университете Нотр-Дам (США, Индиана), где стал активно изучать научные работы Оливера Селфриджа, Марвина Минского, Клода Шеннона и Норберта Винера. Уже в 1960-х годах он выступил в качестве соредатора таких знаковых сборников текстов по кибернетике, как «Моделирование сознания»¹⁶⁴ и «Философия и кибернетика»¹⁶⁵, а в 1969 году опубликовал монографию «Сознание: Философское исследование умов и машин»¹⁶⁶, многие идеи которой затем лягут в основу его более известной книги «Кибернетика и философия сознания»¹⁶⁷.

меньше, во многом в силу доступности еще более примитивных технологий.) Но если они имеют в виду вычислительные механизмы вообще, они ошибаются. Обычно физикалисты не утверждают, что в мире все объясняется с помощью современной физики, – скорее они имеют в виду некую идеальную физическую теорию; так же когнитивисты полагают, что сознание будет объяснено с помощью теории, которая окажется идеальной теорией компьютеров. Остается только наблюдать, как сильно эта теория будет отличаться от современного понимания. Но, как и в случае с физикой, многие современные положения могут остаться в силе» (Boden M.A. *Information, Computation and Cognitive Science // Handbook of the Philosophy of Science. Vol.8: Philosophy of Information* (ed. by P. Adrians, J. van Benthem). Amsterdam: Elsevier. 2008. P. 756.)

¹⁶⁴ *The Modeling of Mind*. (ed. by F.J Crosson, K.M. Sayre). University of Notre Dame Press. 1963.

¹⁶⁵ *Philosophy and Cybernetics*. (ed. by F.J Crosson, K.M. Sayre). University of Notre Dame Press. 1967.

¹⁶⁶ Sayre K.M. *Consciousness: A Philosophic Study of Minds and Machines*. Random House. 1969.

¹⁶⁷ Sayre K.M. *Cybernetics and the Philosophy of Mind*. Routledge and Kegan Paul. 1976.

Сейр отлично разбирался как в достижениях современных ему когнитивных наук, так и в инженерном деле, что позволяло ему ориентироваться в результатах новейших исследований и находить параллели между различными дисциплинами. Свою главную цель философ видел в совмещении наработок технических наук, биологии и философии, а также в применении подобного междисциплинарного подхода к решению основополагающих практических и теоретических задач кибернетики. Однако на пути к этой цели стояла, по мнению Сейра, серьезная преграда. Еще во введении к сборнику «Моделирование сознания» он отмечает, что проект объединения философии с различными науками не будет доведен до конца, пока не будут найдены «внутренние каналы коммуникации» (channels of intercommunication), обеспечивающие продуктивный диалог между данными областями знания: именно несовпадение концептуальных словарей эмпирических и философских наук является причиной, по которой многие вопросы, касающиеся, в частности, природы сознания, до сих пор остаются открытыми¹⁶⁸. В своих последующих работах философ предпринял попытку найти решение проблемы сознания, заключающуюся, с его точки зрения, в невозможности сопоставления результатов когнитивных исследований с концепциями теоретической метафизики и представлениями «народной философии» об особом феноменальном качестве нашей ментальной жизни. По этой причине философ поставил своей целью найти иные, тематически нейтральные концептуальные категории, которые могут быть применимы в рамках разных подходов и дисциплин.

В книге «Кибернетика и философия сознания»¹⁶⁹ Сейр отмечает, что классическим примером подобного применения нейтральных категорий для описания феноменальных и физических состояний является концепция «сенсibiliий» Бертрана Рассела: «сенсibiliями» британский философ

¹⁶⁸ The Modeling of Mind. Ed. by F.J Crosson, K.M. Sayre. University of Notre Dame Press. 1963. P. 10.

¹⁶⁹ Далее мы будем ссылаться в основном на данную работу Сейра, так как именно в ней, по признанию самого философа, последовательно представлены основные идеи, кратко описанные им в его более ранних текстах.

называл гипотетические объекты, обладающие таким же онтологическим статусом, что и ментальные переживания, но не обязательно принадлежащие индивидуальному сознанию какого-либо субъекта; любой объект, таким образом, может быть представлен как ментально, так и физически, будучи определенным с помощью единой категории сенсibiliй¹⁷⁰. Однако основной слабостью подобной теории является, как полагает Сейр, объяснительная несостоятельность используемого британским философом понятия, которое никоим образом не способствует более плодотворному исследованию феноменальных или физических явлений с практической точки зрения:

«Выходит, что любое объяснение ментальных или физических явлений, сформулированное в нейтральных терминах, должно быть в итоге выведено из нейтрального контекста, в котором оно было изначально разработано. Несмотря на свой онтологический и методологический приоритет, подобная модель не способна объяснить реальные явления, а потому вредна для уже существующих теоретических построений»¹⁷¹.

Сейр полагает, что во избежание подобной проблемы, следует не только найти общую категорию, пригодную для описания как психических, так и физико-биологических явлений, но также и интерпретировать данную категорию с помощью языка науки. По этой причине философ предлагает использовать такое понятие как «информация», которое, в отличие от «сенсibiliй» Рассела, обладает намного большей объяснительной силой в силу своей формальной разработанности в сфере компьютерных наук, кибернетики и теории коммуникации. Кроме того, он отмечает, что подобная категория становится все более популярной в современных ему философских исследованиях, чему во многом способствовали «Материалистическая теория сознания» Дэвида

¹⁷⁰ Russell B., Denonn L.E. Bertrand Russell's dictionary of mind, matter, and morals. Philosophical Library. 2014. P. 162.

¹⁷¹ Sayre K.M. Cybernetics and the Philosophy of Mind. Routledge and Kegan Paul, 1976. P. 13.

Армстронга¹⁷² и «Контент и сознание» Дэниела Деннета, оказавшие значительное влияние и на самого Сейра.

Философ отмечает, что предлагаемый им кибернетический подход¹⁷³ является междисциплинарным, так как в нем совмещены теоретические достижения математики, физики, эволюционной биологии, психологии, лингвистики и социологии. В целом, информационно-ориентированная объяснительная стратегия представляется Сейру следующим образом:

«Объяснение ментального в данном [кибернетическом] контексте означает представление его в качестве результата эволюции, то есть как следствие развития биологических процессов из органических. На концептуальном уровне необходимо объяснить такие ментальные свойства, как субъективность и разумность с точки зрения установления возможности их биологического развития. Основные шаги в этом концептуальном “выводе” заключаются в следующем: отрицательная обратная связь анализируется как информационный процесс, в терминах которого можно определить саму эволюцию; затем поведенческая обусловленность анализируется как форма эволюционного процесса, а перцептивное сознание, в свою очередь, как форма данной обусловленности; язык рассматривается как проявление перцептивных способностей человека, а разум – как язык, свободный от контроля стимулов; наконец, субъективность

¹⁷² В данной работе Армстронг выдвинул теорию центральных состояний (center-state theory), по которой мозг выступает в качестве посредника между внешними информационными стимулами и внутренними реакциями на них человека. Армстронг активно использовал понятие информации для решения проблем целенаправленного поведения, интенциональности и восприятия, последнее из которых он определял как «поток информации, который мы осознаем, будучи интроспективно о нем осведомленными» (Armstrong D.M. A Materialist Theory of the Mind. Routledge&Kegan Paul, 1968. P. 226). При этом философ не давал четкого определения столь часто используемому им понятию «информация»: очень вероятно, что философ был знаком с теорией Шеннона и Уивера, однако их работы не упоминаются в его книгах.

¹⁷³ Отметим, что несмотря на то, что сам Сейр в своей работе использует выражение «кибернетический подход», в современной зарубежной историко-философской литературе применительно к теории данного философа чаще употребляется выражение «информационный» (см. например: Adams F.R. The Informational Turn in Philosophy //Minds and Machines, 13 (4), 2003. P. 471-501; Stubenberg L. Neutral Monism // The Stanford Encyclopedia of Philosophy, 2016. URL: <https://plato.stanford.edu/entries/neutral-monism/> (дата обращения: 08.07.2019)).

оказывается измерением языка и разума, на фоне которого человек мыслит другие объекты, отличные от него самого»¹⁷⁴.

Сознание является, по Сейру, продуктом естественного отбора, то есть своеобразной формой эволюционной адаптации сенсорных систем, отвечающих за обработку информационных сигналов и стимулов:

«Сознание, я полагаю, является видом информационной обработки, при которой поступающие к органам чувств сигналы трансформируются в паттерны мозговых состояний, которые остаются стабильными, пока не изменятся из-за смены требований со стороны системы восприятия»¹⁷⁵.

Эволюционное развитие способствует как увеличению вариативности поведения того или иного организма, так и увеличению эффективности его сенсорных информационных каналов, отвечающих за переработку информации и уменьшение шумовых помех. К примеру, при преобразовании визуальных сигналов в форму уже знакомых образных паттернов требуется значительно меньше энергии по сравнению с последовательным преобразованием обособленных зрительных стимулов, которое осуществляется менее приспособленными живыми организмами или вычислительными машинами.

Подобно Фодору и Дретске, Сейр является сторонником каузального подхода к толкованию ментальных содержаний: философ считает, что особенности структурной организации внешних объектов определяют структуру содержания ментальных представлений. *«Подобно Аристотелю, – пишет Сейр, – мы можем утверждать, что структуры, являющиеся результатом деятельности нашего ума, идентичны объективным структурам, расположенным на противоположной стороне нашего восприятия. В контексте теории информации, структуры, присутствующие в мозгу организма, идентичны структурам, характеризующим объект восприятия»¹⁷⁶.*

(В отличие от многих других сторонников экстернализма по отношению к

¹⁷⁴ Sayre K.M. Cybernetics and the Philosophy of Mind. Routledge and Kegan Paul, 1976. P. 15.

¹⁷⁵ Там же. С. 156.

¹⁷⁶ Там же. С. 155.

содержанию сознания, Сейр не использует в своих теоретических построениях понятие репрезентации, и нигде не говорит о том, что в процессе мышления происходит репрезентация объекта восприятия или его свойств, – философ лишь предполагает, что ментальные представления изоморфны информационной структуре внешних объектов).

Сейр отмечает, что широко распространенный в современных ему естественных науках вычислительный подход не является пригодным для описания свойств когнитивной деятельности, так как подобный подход оперирует исключительно синтаксическим определением информации, игнорируя другие трактовки этого понятия, более пригодные для описания биологических и психологических процессов. В своих более поздних статьях, таких как «Интенциональность и обработка информации»¹⁷⁷ и «Когнитивная наука и проблема семантического содержания»¹⁷⁸, философ представляет свой взгляд на описанную проблему: с его точки зрения, действительно эффективный подход к объяснению механизмов когнитивной деятельности должен начинать с описания базовых этапов обработки информации, понятой в техническом смысле (которые, по его мнению, в равной степени свойственны как механическим, так и биологическим системам, манипулирующим символами), а затем работать над объяснением того, как биологические системы приобретают дополнительную способность для работы с семантической информацией. Скорее всего, полагает Сейр, ментальные представления могут объясняться как паттерны активации нейронов в коре мозговых полушарий, благодаря которым когнитивный агент способен фокусироваться на объектах в постоянно изменяющейся окружающей среде и контролировать «уровень взаимной информации» между данными паттернами и соответствующими объектами. Отсюда следует, что интенциональные функции отвечают за установление и поддержание таких представлений, и семантические свойства добавляются к информационным состояниям как средство эффективного контроля. Философ

¹⁷⁷ Sayre K.M. Intentionality and information processing: An alternative model for cognitive science // Behavioral and Brain Sciences. 1986. №9(1). P.121-138.

¹⁷⁸ Sayre K.M. Cognitive Science and the Problem of Semantic Content. // Synthese, 1987. №70. P. 265.

заключает, что трактуемая подобным образом семантическая обработка информации вполне может быть свойственна и некоторым искусственным системам вроде роботов или компьютеров.

Сейр также пытается закрыть разрыв между трактовкой информации, используемой в рамках коммуникационной теории Шеннона, и семантическим толкованием информации, объясняя развитие сложных смысловых структур на примере человеческой речи: философ считает, что так как все звуки в природе ассоциируются с определенными объектами посредством каузальных взаимодействий, то «смысл» появляется в результате интенциональной направленности человека на какой-либо объект или достижение цели. Так как интенции, которые философ определяет как осознанные цели, направленные на контроль поведения, представлены нейронными конфигурациями индивидуального организма, необходимо понять, как одна и та же интенция может влиять на поведение разных индивидуумов: по мнению философа, смысл, при котором одно и то же значение способно присутствовать в сознании разных индивидуумов, может быть объяснен присутствием в сознании воспринимающего тех же информационных структур, что и в объекте восприятия. Таким образом, с помощью понятия «информационного изоморфизма» философ объясняет наличие у различных субъектов восприятия схожих паттернов организации нейронных структур, влияющих на формирование одинаковых интенциональных состояний и, следовательно, схожих ментальных содержаний.

Итак, подход Сейра с первого взгляда напоминает простой информационно-ориентированный функционализм: философ не только отмечает наличие эволюционных преимуществ у организмов, обладающих сознанием, но также делает попытку объяснить, как информационно-обусловленные ментальные состояния способны каузально влиять на поведение человека¹⁷⁹.

¹⁷⁹ Для этой цели Сейр предлагает собственную модель статистической каузации, которая во многом основывается на каузальной теории немецкого философа и логика Ганса Рейхенбаха (Sayre K.M. *Cognitive Science and the Problem of Semantic Content*. // *Synthese*, 1987. №70. P. 65-87; Sayre K.M. *Statistical Models of Causal Relations*. // *Philosophy of Science*, 1977. № 44 (2). P. 203-214).

Однако, как ни парадоксально, одна из самых любопытных идей его философии представляет собой небольшое рассуждение во введении к работе «Кибернетика и философия сознания», где философ размышляет над вопросом об онтологическом статусе информационных структур. Сейр считает, что так как информация способствует появлению у организма ментальных переживаний, то информационные пространства должны быть онтологически первичны по отношению к сознанию. Философ пишет:

«Данный проект демонстрирует, что онтология информационных состояний является подходящей для объяснения явления сознания, как отличного от онтологии физических событий. Возможно также, что онтология информации является основой и для физических наук»¹⁸⁰.

Сейра нередко причисляют к сторонникам концепции информационного реализма, согласно которой информационные пространства являются самостоятельными сущностями, чье существование независимо от внешнего наблюдателя¹⁸¹. Подобная интерпретация информации, предполагающая толкование данного понятия не только как тематически нейтрального вспомогательного термина, но и как онтологически фундаментальной категории, к которой потенциально могут быть сведены все известные физические и феноменальные явления, сделала теорию Сейра известной и во многом актуальной: так, некоторые исследователи отмечают, что предложенная философом концепция, является одной из наиболее убедительных теорий нейтрального монизма, существующих на данный момент¹⁸². Выводы Сейра о возможности признания информационного реализма также нередко сопоставляют с построениями современной цифровой физики, а концепцию «информационного изоморфизма» – с положениями теории нередуктивного

¹⁸⁰ Sayre K.M. Cybernetics and the Philosophy of Mind. Routledge and Kegan Paul, 1976. P.16.

¹⁸¹ Предположение Сейра можно сопоставить с так называемыми «атрибутивными» концепциями информации, активно развиваемыми некоторыми отечественными исследователями в 1960-1980-х годах, среди которых А.Д. Урсул, И.Б. Новик, Л.Б. Баженов, Л.А. Петрушенко и др.

¹⁸² См. например: Jaworski W. Philosophy of Mind: A Comprehensive Introduction. Wiley-Blackwell. 2011. P. 268.

функционализма Дэвида Чалмерса (сам Чалмерс отмечает, что находит идеи Сейра весьма плодотворными, хотя и утверждает, что ознакомился с книгой «Кибернетика и философия сознания» уже после создания собственной теории¹⁸³).

Несмотря на свою сегодняшнюю актуальность, информационная теория Сейра некоторое время оставалась без должного внимания со стороны историков философии: и по сей день присущая философу ориентация на междисциплинарный подход является помехой для исследователей, делающих попытки вписать идеи Сейра в общий контекст развития философских теорий сознания. Эта ситуация является наглядной иллюстрацией проблемы, связанной с классификацией многих известных сегодня информационных теорий сознания: концепцию Сейра, сочетающую в себе некоторые устаревшие идеи кибернетического подхода к описанию работы мышления, а также ряд нестандартных решений в области лингвистики, психологии и социологии, иногда бывает сложно соотнести с прочими концепциями аналитической философии, а также с известными сегодня когнитивными теориями, что, несомненно, затрудняет историко-философский анализ генезиса информационных теорий сознания в целом.

2.4. Информационный подход к объяснению сознания Д.И. Дубровского

Российский философ Давид Израилевич Дубровский является одним из самых известных сторонников информационного подхода к объяснению природы сознания, чья теория получила широкое признание в кругу как отечественных, так и зарубежных исследователей. Посвятив большую часть своих работ рассуждениям над знаменитой философской проблемой «сознание-тело», Давид Израилевич разработал последовательную концепцию, в рамках которой внимание уделяется не только информационной трактовке различных

¹⁸³ Чалмерс Д. Сознание и ум: В поисках фундаментальной теории. Пер. В.В. Васильева. М.: УРСС: Книжный дом «ЛИБРИКОМ», 2013. С. 483

аспектов мыслительной деятельности, но также и анализу смежных вопросов, возникающих в сферах онтологии, эпистемологии, кибернетики, теории искусственного интеллекта и др.

Еще в 1960-х годах вышел ряд знаковых статей философа, в которых он настаивает на необходимости активного изучения работы мозга, а также на том, чтобы ученые относились «всерьез» к проблеме сознания и связанным с ней философским вопросам. Так, в статье 1968 года «Мозг и психика»¹⁸⁴, ставшей впоследствии поводом для оживленной дискуссии Д.И. Дубровского с Э.В. Ильенковым¹⁸⁵, для описания особых качественных свойств ментальных состояний философ впервые использует понятие «субъективная реальность», которое соотносится с понятием «объективная реальность», служащим, в свою очередь, для характеристики физических структур, выступающих своеобразными «носителями» субъективных качеств. Философ поставил перед собой задачу дать сугубо научное объяснение природы психических явлений, которые, с его точки зрения, зависят от индивидуальных особенностей структурного устройства головного мозга человека, динамически организующего воспринимаемый им поток разнородной информации. Будучи зафиксированной в нейродинамической системе, эта информация и является тем самым качественным содержанием психики, которую философ именует субъективной реальностью. Полагая, что именно нейродинамическое описание должно лежать в основе адекватной трактовки природы психических переживаний, Дубровский отмечает, что перед современными учеными стоит задача исследования нейродинамического кода, лежащего в основе подобных субъективных явлений. Данная задача, сводящаяся к обнаружению

¹⁸⁴ Дубровский Д.И. Мозг и психика // Вопросы философии. 1968. № 8.

¹⁸⁵ Плеханов И.А. Д.И. Дубровский и Э.В. Ильенков: неоконченный спор о природе психики и идеальном // Вестник Нижегородского университета им. Н.И. Лобачевского, 2007. №1(6). С. 153-162. Иванов Д.В. Психофизическая проблема и проблема идеального в дискуссии Э.В. Ильенкова и Д.И. Дубровского // Психологические исследования. 2019. Т. 12. № 65-66. С.5.

специфических закономерностей информационных процессов, происходящих в мозге, будет еще не раз упоминаться философом в его последующих работах¹⁸⁶.

Полагая, что психофизическая проблема, сводящаяся к вопросу о природе связи между мозгом и субъективными переживаниями, представляет собой одну из самых насущных проблем современных естественных наук, Дубровский, однако, отмечает, что сами феноменальные явления не входят в предмет исследования эмпирических дисциплин и претендуют на отдельный, философский анализ; в силу несовпадения средств описания физических и ментальных явлений, используемых в современной науке, перед исследователями также стоит нелегкая задача, связанная с проблемой толкования понятия «сознание». Для решения этой и ряда других проблем философ предлагает концепцию, во многом основанную на идеях классической информационной теории Шеннона.

Интерпретации информационной теории Дубровский уделяет особое внимание в своей монографии 1971 года «Психические явления и мозг. Философский анализ проблемы в связи с некоторыми актуальными задачами нейрофизиологии, психологии и кибернетики»: ссылаясь на теоретические разработки советских и зарубежных исследователей, философ отмечает, что основанный Шенноном математический подход к толкованию информационных структур требует значительного преобразования, подразумевающего интеграцию синтаксической теории информации с семантическими и прагматическими концепциями. Философ полагает, что осуществить подобное преобразование можно путем ограничения класса физических объектов, с которыми могут быть ассоциированы информационные свойства. С его точки зрения, подходы, описывающие информацию как свойство всякого физического объекта (так называемые «атрибутивные» подходы), трактуют данное понятие слишком широко, в силу чего информационные термины сильно сближаются с такими физическими категориями, как, например, каузальность; Дубровский

¹⁸⁶ Дубровский Д. И. Проблема нейродинамического кода психических явлений // Вопросы философии. 1975. № 6; Дубровский Д. И. Расшифровка кодов (Методологические аспекты проблемы) // Вопросы философии. 1979. № 12.

отмечает, что информационные свойства следует приписывать только сложным самоорганизующимся системам, способным не только к хранению информации, но также к ее передаче, обработке и использованию в своих целях. Даже искусственные вычислительные системы, способные, казалось бы, выполнять многие из перечисленных функций, не являются, по мнению философа, полноправными информационными агентами, так как они не обладают целями, отличными от тех, что были изначально запрограммированы человеком. Дубровский пишет:

«[...] человек способен придавать любому объекту любое информационное значение, носящее конвенциональный характер. Это обстоятельство и та его более широкая основа, которая именуется очеловечиванием природы и опредмечиванием человека, также связанная с наделением естественных объектов информационными значениями, по-видимому, и питают теоретическую иллюзию о вездесущности информационных процессов»¹⁸⁷.

По этим причинам Дубровский склоняется к более узкой интерпретации информационных пространств, в рамках которой информационные процессы свойственны только системам, обладающим достаточно высоким уровнем функциональной организации. Подобные процессы являются, согласно философу, обратной стороной деятельности системы, ассоциированной с управлением, и представляются системе как явления субъективной реальности, то есть феноменальные переживания.

Опираясь на подобную трактовку информации, Д.И. Дубровский создает свою концепцию сознания. Несмотря на то, что данная концепция за более чем полвека претерпела ряд изменений, ее ключевые идеи, остались практически неизменными¹⁸⁸. В основе подхода философа, опирающегося, по его словам, на

¹⁸⁷ Д.И. Дубровский. Психические явления и мозг: философский анализ проблемы в связи с некоторыми актуальными задачами нейрофизиологии, психологии и кибернетики. М.: Наука, 1971. С. 257.

¹⁸⁸ Далее идеи философа будут представлены в тех же формулировках, которые использованы в его статьях из сборника 2007 года «Сознание, мозг, искусственный интеллект», а также более поздних работах, среди которых книга 2015 года «Проблема “Сознание и мозг”: теоретическое решение» и др.

последние достижения эволюционной биологии, лежат две базовые посылки и одно соглашение:

«1) Информация с необходимостью воплощена в физическом носителе, который выступает в качестве ее кода.

2) Информация инвариантна по отношению к физическим свойствам своего носителя и может иметь разнообразные кодовые воплощения.

3) Феномены субъективной реальности могут интерпретироваться как информация»¹⁸⁹.

В случае сознания, информация, трактуемая как «содержание сообщения» или «значение сигнала», имеет конкретный носитель – мозг, являющийся комплексной нейродинамической системой. Связь между ментальными явлениями и их носителем является функциональной, а потому не может быть объяснена исключительно физическими свойствами носителя: так как подобная связь представляет собой «кодую зависимость», она может быть воплощена на других носителях, обладающих иными физическими свойствами. Кодовая зависимость, исторически сформировавшаяся еще в филогенезе и онтогенезе сознательного организма, определяет функциональную организацию мозговых процессов, лежащих в основе некоторого психического явления (именно по этой причине философ считает, что расшифровка данных кодов является первейшей задачей естественных наук). Поскольку информационная структура может быть реализована на носителях, отличающихся по своим физическим характеристикам, то в сложной самоорганизующейся системе каузальными свойствами обладают не конкретные физические свойства данного носителя, а именно содержащаяся в нем информация. Так философ вводит понятие «информационной причинности».

¹⁸⁹ Дубровский Д. И. Зачем субъективная реальность, или «почему информационные процессы не идут в темноте?» (Ответ Д. Чалмерсу) // Сознание, мозг, искусственный интеллект: сб. статей. М.: 2007. С. 146.

Феноменальное сознание, как полагает Дубровский, связано с определенным способом представления информации индивиду, который отличается, во-первых, данностью информации в «чистом» виде, означающем, что индивид не знает физических характеристик ее носителя (то есть мозга), во-вторых, доступностью данной информации для управления. Философ отмечает, что далеко не все информационные процессы предстают в форме качественного опыта, а только те, которые были подвергнуты определенной кодовой переработке:

«Для того, чтобы информация приобрела форму субъективной реальности, необходимо, по крайней мере, двойное, или, лучше сказать, двухступенчатое кодовое преобразование на уровне эго-системы: первое из них представляет информацию как таковую (которая пребывает пока “в темноте”), второе преобразование “открывает” и тем самым актуализует ее для “самости”, делает доступной для оперирования и использования в целях управления»¹⁹⁰.

Философ утверждает, что в результате подобного рода преобразования возникает «информация об информации», представляющая собой такой вид информационной интеграции, который и обеспечивает появление нашей внутренней «виртуальной реальности». В книге «Проблема “Сознание и мозг”» философ выделяет шесть параметров данной реальности: временной, содержательный (связанный со свойством сознания быть интенционально направленным на некоторый объект), формальный (всякое явление субъективной реальности должно быть оформленным и структурированным), истинностный, ценностный, и деятельностный (волевой)¹⁹¹.

Дубровский полагает, что представленная им теория решает многие проблемы, связанные с описанием природы сознания, и обладает значительными преимуществами по отношению к некоторым другим аналитическим

¹⁹⁰ Там же. С. 153.

¹⁹¹ Дубровский Д.И. Д.И. Дубровский. Проблема «Сознание и мозг»: теоретическое решение. М.: «Канон+», 2015. С. 42.

концепциям. Так, философ критикует теорию Деннета, в которой, по его словам, должным образом не исследован вопрос об онтологии сознания: редукционистская установка американского философа, основанная на стремлении элиминировать субъективные переживания, связана, с точки зрения Дубровского, с узким пониманием Деннетом гносеологической составляющей психофизиологической проблемы и признанием главным признаком наличия подобных переживаний функционирование информационно-обусловленных когнитивных процессов. Философ отмечает, что сознание несводимо к простой когнитивной компетентности, и идеи функционализма, лежащие в основе рассуждений Деннета, не препятствуют возможности постулирования реальности субъективных переживаний¹⁹². По этой причине идеи другого американского философа, Джона Серла, сводимые, с точки зрения Дубровского, к тезису о присущей сознанию «неэлеминируемой субъективной онтологии», представляется отечественному философу более приемлемыми, хотя и не признаются им как новые: он отмечает, что подобные предположения были высказаны задолго до публикации работ Серла. Дубровский также полагает, что постулируемая Серлом онтология сознания как «онтология от первого лица» требует отдельной философской рефлексии и концептуального анализа. Особый интерес у отечественного философа также вызвала критика Серлом информационно-вычислительных подходов, в приверженности которым американский философ усматривает грубейшую ошибку современной когнитивной науки. Дубровский отмечает, что, вопреки утверждениям Серла, далеко не все философы являются сторонниками грубой аналогии мозг-компьютер, в рамках которой сознательный опыт полностью нивелируется (такие рассуждения, полагает философ, свойственны разве что Деннету). В целом, Дубровский критикует Серла за общую концептуальную

¹⁹² Дубровский Д. И. В «Театре» Дэниэла Деннета (об одной популярной концепции сознания) // Сознание, мозг, искусственный интеллект: сб. статей. М.: 2007. С.71-87.

непоследовательность его теории, выраженную в самых различных ее аспектах¹⁹³.

Пожалуй, наиболее интересным представляется вопрос о том, может ли теория Дубровского решить так называемую «трудную проблему» сознания, заключающуюся в поиске причин возникновения феноменального опыта. Сам философ отвечает на данный вопрос утвердительно: по его мнению, «трудная проблема» уже давно решена в рамках его подхода. Отвечая на вопрос Дэвида Чалмерса «Почему информационные процессы не идут в темноте, и почему они непременно должны сопровождаться субъективными переживаниями?», Дубровский отмечает, что подобная постановка вопроса предполагает отрицание специфической функциональной особенности и каузальной действенности ментальных качеств. Однако эти качества, по утверждению философа, являются необходимым инструментом выполнения и согласования организмом его интегральных функций, в числе которых поддержание своей целостности и управление. Ментальные представления являются важной частью моторных и когнитивных функций, и, так как субъективная реальность и ее кодовое воплощение представляют собой одновременные и однопричинные явления, то между собой они находятся в отношении однозначного соответствия; поэтому Дубровский отрицает эпифеноменализм, ведь феноменальный опыт, с его точки зрения, играет далеко не последнюю роль в функционировании сложных живых организмов, а потому не может считаться «номологическим бездельником». Отсюда, по мнению философа, следует, что существа, у которых информационные процессы происходят в темноте (то есть бессознательные зомби, чья функциональная организация полностью совпадает с нашей), не могут существовать в реальном мире.

Таким образом, будучи сторонником функционального подхода к толкованию информации, Дубровский отрицает наличие концептуального разрыва между описанием феноменальных и организационно-функциональных

¹⁹³ Дубровский Д.И. Новое открытие сознания? (по поводу книги Джона Серла «Открывая сознания заново») // Сознание, мозг, искусственный интеллект: сб. статей. М.: 2007. С. 37.

свойств организма, так как полагает, что субъективные состояния также являются функциональными состояниями. Как и все упомянутые ранее представители информационных подходов к толкованию сознания, российский философ отмечает, что и психофизиологическая проблема и тесно связанная с ней «трудная проблема» сознания возникают в силу категориальной разобщенности языков, на которых описываются нейрофизиологические и ментальные явления. Дубровский считает, что для соединения данных языков требуются релевантные понятийные структуры (философ именует их «концептуальными мостами»), которые вполне могут быть основаны на информационном терминологическом аппарате: *«Понятие информации, – пишет философ, – с одной стороны, способно выражать основное гуманитаристское содержание (смысл, ценность, интенциональность и т.д.), а, с другой стороны, в силу кодовой воплощенности информации, оно допускает “физикалистские” описания (пространственные, энергийные, субстратные и др.)»*¹⁹⁴. Именно концептуальный словарь теории информации представляется Дубровскому тем самым источником нейтральных терминов, способным стать основой для нового языка науки и философии.

Д.И. Дубровский создал достаточно последовательную информационную теорию, в рамках которой предпринял попытку наметить пути решений главных философских проблем, связанных с объяснением природы сознания. Однако, далеко не все описанные философом решения представляются успешными, и остаются сомнения, что оформленная в подобном виде информационная теория действительно способна пролить свет на «трудную проблему» сознания: так, остается не до конца понятным, почему функционально-обусловленная «информация об информации» предстает в виде феноменальных переживаний. Как отмечает российский философ Д.В. Иванов, *«единственный способ объяснить появление сознания, или субъективной реальности, – это наделить эго-систему особым онтологическим статусом, выделяющим ее из множества*

¹⁹⁴ Дубровский Д. И. Зачем субъективная реальность, или «почему информационные процессы не идут в темноте?» (Ответ Д. Чалмерсу) // Сознание, мозг, искусственный интеллект: сб. статей. М.: 2007. С. 159.

*иных физических или функциональных процессов»*¹⁹⁵. Однако такое решение явно не следует считать удовлетворительным.

Кроме того, предложенное Дубровским доказательство каузальной релевантности сознательных состояний подводит нас к вопросу о свойствах информационных структур, подвергнутых разным уровням обработки или кодового преобразования. Вопрос о том, почему именно обработанная информация является каузально релевантной, в то время как пребывающие «в темноте» непреобразованные информационные структуры не обладают должной каузальной силой, остается открытым (в случае, если необработанная информация также обладает релевантной каузальной силой, необходимой для управления организмом, возникает вопрос, зачем нам в принципе нужны феноменальные переживания). По этой причине представляется, что предложенная философом концепция информационной причинности нуждается в некоторой доработке.

Возможно, что последовательный концептуальный анализ, дополненный методологией экспериментальной философии, способствует развитию предложенной российским философом теории. Представляется, что используемый в ранних работах Дубровского кибернетико-ориентированный категориальный аппарат несколько затрудняет сопоставление положений его информационной теории с достижениями современных естественных наук, оперирующих несколько иной терминологией; тем не менее, этот фактор не стоит расценивать как очевидный недостаток концепции Дубровского, чей информационный подход к объяснению природы сознания уже более 50 лет остается весьма актуальным и конкурентоспособным.

¹⁹⁵ Иванов Д.В. Природа феноменального сознания. М.: Книжный дом «ЛИБРОКОМ». 2013. С. 88.

Заключение главы

Начиная с 1960-х годов информационный концептуальный аппарат стал играть важную роль в теориях многих известных аналитических философов. Среди наиболее показательных информационных теорий сознания, оказавших впоследствии значительное влияние на труды многих современных исследователей, можно выделить теорию Фреда Дретске, предпринявшего попытку объяснить сознание с помощью концепции информационно-обусловленного репрезентативизма; теорию Дэниела Деннета, который, вдохновившись настроением дезавуировать все существующие метафизические проблемы, связанные с вопросами о природе феноменальных состояний, активно использовал в своих философских построениях информационно-вычислительную метафору; кибернетическую теорию Кеннета Сейра, признавшего возможной онтологическую первичность информационных пространств не только по отношению к ментальным качествам, но и по отношению к физическим структурам; концепцию отечественного философа Д.И. Дубровского, оформившего последовательный подход к объяснению природы сознания, основанный на функциональной трактовке информационных структур.

Несмотря на то, что сторонники информационных теорий сознания признавали тесную связь своих концепций как с идеями классического компьютеризма, так и с некоторыми положениями популярного в биологических науках коннекционизма, они, однако, отдавали предпочтение более абстрактному описанию процессов, лежащих в основе работы сознания: понятие «вычисление», трактуемое ранее более-менее формализовано (например, когда имелось в виду вычисление в стиле машины Тьюринга), стало более размытым, что привело к концептуальному сближению данного понятия с менее проясненным понятием «информационная обработка» (это также стало причиной, по которой в современных философских текстах нередко встречается выражение «информационно-вычислительные теории сознания»),

подчеркивающее отсутствие четких границ между «информационными» и «вычислительными» подходами к объяснению сознания).

Рассмотрев ключевые положения теорий упомянутых выше философов, легко заметить, что основным поводом их обращения к информационному терминологическому аппарату стала потребность в поиске нейтральной категории, способной сократить разрыв между научным объяснением сознания в функциональных и нейробиологических терминах и феноменологией субъективных переживаний: отсылка к понятию «информация», которое, как нетрудно заметить, упомянутые выше философы толковали по-разному, являлась не только своеобразной данью моде 1960-1970-х годов на вычислительную и кибернетическую интерпретацию различных процессов и явлений (именно на эти годы пришлось становление теорий рассмотренных в данной главе философов), но и результатом признания необходимости поиска концептуального фундамента, на основе которого можно было бы объединить разнородные теоретические представления о природе сознания (которое, к слову, данные исследователи также интерпретировали по-разному). Искренняя вера в «магию» информационного объяснения побудила этих и некоторых других философов свести все нерешенные проблемы, возникающие на почве философии сознания, к таинственному процессу обработки информации, природа которого, по признанию самих философов, является весьма абстрактной и не до конца проясненной. Однако подобная вера была присуща далеко не всем исследователям того времени, вследствие чего уже вскоре после своего появления в рамках аналитической философии информационные теории

Глава 3. Основные аргументы против информационных теорий сознания

Хотя информационная интерпретация сознания стала весьма популярной как в когнитивных науках, так и в аналитической философии, ряд исследователей отрицает способность информационной теории пролить свет на природу субъективного опыта. Более того, некоторые критики, чьи аргументы будут разобраны далее в настоящей главе, считают подобную интерпретацию не просто временной и бесполезной, но даже вредной для философии в силу абстрактности используемых в ее рамках понятий и категорий. По их мнению, важнейшим условием применения информационной терминологии для описания сложных ментальных процессов является однозначность используемых в рамках подобного описания концептуальных структур; данное условие, как показывает история когнитивных наук и философии сознания, нередко нарушалось. Смещение синтаксического и семантико-прагматических аспектов толкования понятия «информация» вкупе с активным применением компьютерной метафоры в науке послужило размыванию границ между используемыми исследователями категориями, что, в свою очередь, крайне затруднило дальнейшее развитие многих информационных теорий сознания. Несмотря на активные попытки их сторонников концептуально оформить используемый ими терминологический аппарат, многие из рассмотренных выше теорий все еще сталкиваются с рядом характерных трудностей.

Далее в настоящей главе будут разобраны основные группы аргументов, выдвигаемых против информационных теорий сознания, а также проанализированы возможные пути решения обозначенных в них проблем.

3.1. Некоторые общие возражения против информационных и вычислительных моделей сознания

Как уже было отмечено, понятия «вычисление» и «обработка информации» стали практически взаимозаменяемыми в литературе о мозге и сознании, хотя утверждение «вычисление есть обработка информации» ставится многими учеными под сомнение¹⁹⁶. Так, один из современных представителей вычислительного подхода, американский философ итальянского происхождения Гвалтеро Пичинини считает, что распространенное предположение о том, что мышление имеет вычислительный характер в силу своей деятельности, связанной с обработкой информации, не совсем верно: использование данных категорий в рамках одной теории, полагает философ, делает подобную теорию весьма неоднозначной и затрудняет ее сопоставление с более формально-разработанными научными концепциями¹⁹⁷. Современный исследователь Лучано Флориди, занимающийся вопросами, связанными с философским толкованием информации, в одной из своих статей отмечает, что «вычисление» является лишь одним из процессов, в которые может быть вовлечена информация¹⁹⁸, в то время как философ Михаэль Раскорла определяет информационные теории сознания как частный случай вычислительных

¹⁹⁶ Gallistel C.R., King, A. *Memory and the Computational Brain*. Malden: Wiley-Blackwell, 2009; Lizier J., Flecker B., Williams P. *Towards a Synergy-based Account of Measuring Information Modification* // *Proceedings of the 2013 IEEE Symposium on Artificial Life (ALIFE)*, Singapore, 2013. P. 43–51; Milkowski M. *Explaining the Computational Mind*. Cambridge: MIT Press, 2013.

¹⁹⁷ По мнению Пичинини, основная трудность состоит в уже упомянутой выше размытости границ между подходами к толкованию информации. По этой причине философ предпринял попытку классифицировать основные подходы к интерпретации информационных структур и связанные с ними концепции сознания: согласно типологии Пичинини, информация может быть трактована не-семантически (то есть пониматься как уменьшение неопределенности) или семантически (как уменьшение неопределенности, связанной с тем или иным положением дел). Семантические подходы, в свою очередь, делятся на натурализирующие информацию (Ф. Дретске, Д. А. Фодор, Р. Г. Миликан) и ненатурализирующие (Л. Флориди). В зависимости от того, какая из перечисленных трактовок используется, вычислительные теории сознания и мышления можно поделить на такие же группы соответственно (Piccinini G., Scarantino A. *Information processing, computation and cognition* // *Journal of Biological Physics*, 2011. № 37(1). P. 1-38; Piccinini G., Scarantino A. *Computation vs. Information processing: Why their Difference Matters to Cognitive Science* // *Studies in History and Philosophy of Science*. 2010. №41. P. 237–246).

¹⁹⁸ Floridi L. *Trends in the Philosophy of Information* // *Handbook of philosophy of information* (ed. by P. Adriaans and V. Benthem). Amsterdam; London, North-Holland. 2008. P. 113-131.

теорий¹⁹⁹. Некоторые исследователи полагают, подобно Дэвиду Марру, что вычислительное и информационное описания ментальных переживаний, опирающиеся на концептуальный аппарат логико-математической теории, по меньшей мере дополняют друг друга, а источником рассматриваемых терминологических затруднений являются не сами подходы, а приверженность исследователей к догматическим «измам»²⁰⁰.

Таким образом, среди современных исследователей нет четкой договоренности о том, допустима ли трактовка понятий «вычисление» и «информационная обработка» не только как взаимосвязанных, но и как взаимозаменяемых. Безусловно, появление последовательной теории количественного измерения информации было тесно связано с развитием теории вычислимости, главную роль в которой играют понятия универсальности и инвариантности: так, теория Тьюринга предсказывает, что все универсальные модели вычислений, по-видимому, обладают одинаковой силой, и все возможные меры информации определены для данных универсальных моделей (что означает их инвариантность). С формальной точки зрения, детерминированные системы обработки информации (то есть локальные, последовательные системы, манипулирующие дискретными объектами по заданным правилам) не создают новую информацию, а лишь преобразуют ее, представляя в том или ином виде. Поэтому любое вычисление, представляющее собой манипуляцию символами в соответствии с заданным алгоритмом, можно также толковать в информационных терминах, то есть как процесс обработки информации, определяемый некоторой предзаданной инструкционной информацией. Разумеется, подобная трактовка, привязывающая нас к конкретной интерпретации вычислительного процесса, является несколько натянутой, так как далеко не все вычислительные программы можно однозначно представить как набор инструкционных данных; более того, не совсем ясно, как

¹⁹⁹ Rescorla M. The Computational Theory of Mind // The Stanford Encyclopedia of Philosophy (N. Zalta ed.), 2019. [Электронный ресурс]. URL: <https://plato.stanford.edu/entries/computational-mind/#AltConCom> (дата последнего обращения: 04.12.2020).

²⁰⁰ См. например: Faries F., Chemero A. Dynamic Information Processing // The Routledge Handbook of the Computational Mind (ed. by M. Sprevak, M. Colombo). NY, 2019. P. 134.

следует интерпретировать своеобразную двойственность информации, присущую подобным самореферентным вычислительным системам, оперирующим как инструкционными, так и неинструкционными данными. Тем не менее, некоторые исследователи склоняются к мысли, что возможность подобной интерпретации понятий информационной обработки и вычисления говорит об их теоретической взаимозаменяемости²⁰¹. Однако большинство ученых и философов предпочитают оставлять вопрос о связи процессов вычисления и обработки информации открытым; по этой причине, любая теория сознания, оперирующая понятиями «вычисление» и «обработка информации» в идеале должна продемонстрировать, как эти понятия концептуально связаны между собой. С историко-философской точки зрения решение данной проблемы призвано привнести ясность в вопросе «Являются ли все вычислительные теории сознания информационными (или все информационные теории вычислительными)?»²⁰².

Что касается более конкретных популярных критических замечаний, то большинство из них направлены непосредственно против классической версии компьютеризации и призваны показать, что символическое и алгоритмическое описание когнитивных процессов крайне утрирует наши представления о мышлении. Еще Алан Тьюринг, предвидя, что его представления о «мыслящих машинах» вызовут бурю протестов, парировал некоторым возможным возражениям, среди которых он выделил так называемое теологическое возражение (мышление – это функция бессмертной души, которую Бог вложил в человека, но не в машину), возражения, согласно которым машины не могут обладать различными «чисто человеческим» способностями (например, получать удовольствие, создавать произведения искусства,

²⁰¹ Adriaans P., Pet van Emde B. Computation, Information, and the Arrow of Time // *Computability in Context: Computation and Logic in the Real World* (ed. by B. Cooper, A. Sorbi). London: Imperial College Press, 2011. P. 1–17.

²⁰² Представленные далее в данной главе аргументы направлены равным образом как против вычислительных, так и против информационных моделей сознания, так как большинство авторов приведенных ниже аргументов не постулирует особой разницы между этими моделями. Поэтому в целях упрощения изложения материала в данной части исследования понятия «вычисление» и «обработка информации» будут употребляться как взаимозаменяемые, из чего, однако, не следует, что вопрос о связи этих понятий будет опущен в дальнейшем.

анализировать собственные поступки, вести себя «неформально», поступать в соответствии со своей интуицией и т.п.), аргумент о недискретном характере работы нервной системы и т.д. На данные и многие другие смежные аргументы Тьюрингу удалось легко ответить²⁰³.

Помимо вышеперечисленных возражений, а также уже упомянутых в Главе 1 аргументов систематичности и продуктивности, долгое время внимание исследователей привлекали аргумент темпоральности и аргумент воплощенного познания. Первый сводится к утверждению о том, что классический вычислительный подход не позволяет адекватно учитывать временные аспекты мышления: время, которое требуется условной вычислительной системе наподобие машины Тьюринга на реализацию некоторого вычисления не обговаривается, и, следовательно, любой вычислительный процесс может происходить в течение сколь угодно длительного временного интервала. Однако современные представители вычислительных теорий не видят здесь особой проблемы, ведь вполне возможно дополнить любую вычислительную модель описанием длительности каждой стадии, тем самым обеспечивая теорию подробными временными предсказаниями²⁰⁴. Авторство второго аргумента принадлежит достаточно разнородному кругу специалистов (как правило, сторонникам энактивистских когнитивных теорий, о которых было упомянуто в Главе 1), чьи доводы сводились, во-первых, к указанию на невозможность моделирования умственной деятельности с помощью абстрактных символических манипуляций, и, во-вторых, к отрицанию репрезентативизма по отношению к ментальным состояниям, которые, с их точки зрения, следует описывать как результат тесного взаимодействия субъекта с окружающим миром. Данный аргумент долгое время являлся достаточно серьезным, так как ставил под сомнение базовые положения ранних версий вычислительных когнитивных концепций. Однако представляется, что подобные аргументы могут угрожать далеко не всем теориям сознания, в рамках которых

²⁰³ Turing A. M. Computing Machinery and Intelligence. *Mind*, 1959. №49. P. 442-450.

²⁰⁴ Piccinini G. The Resilience of Computationalism // *Philosophy of Science*, 2010. №77. P. 852-861.

субъективные переживания трактуются как обладающие информационной природой: во-первых, далеко не все подобные теории оперируют строгими вычислительными формализмами (и не все основаны на положениях репрезентативизма); во-вторых, описываемые в их рамках когнитивные агенты могут обладать сенсорными входами и моторными выходами и, как следствие, вступать в активное информационное взаимодействие с окружающей средой.

В науке и философии некоторое время был также популярен так называемый аргумент «вычислительной неполноты», согласно которому теорема неполноты Курта Гёделя демонстрирует, что вычислительные способности человеческого мозга сильно превосходят возможности любой условной машины Тьюринга. К подобной идее склоняется, в частности, британский физик Роджер Пенроуз, автор известной книги «Новый ум короля. О компьютерах, мышлении и законах физики»²⁰⁵: Пенроуз полагает, что сознание человека не является полностью алгоритмическим, так как некоторые низкоуровневые физические процессы, определяющие его свойства, не могут быть описаны с помощью единой формальной символической системы, и для интерпретации ментальной деятельности следует разработать концепцию, опирающуюся на законы квантовой механики. Впрочем, идеи Пенроуза и других сторонников подобного аргумента²⁰⁶ были подвержены серьезной критике со стороны многих философов и логиков, утверждающих, что подобные доводы содержат не только безосновательные спекулятивные рассуждения, но и грубые теоретические ошибки²⁰⁷. Сегодня среди аналитических философов существует консенсус в отношении силы данного аргумента: сама по себе теорема Гёделя не дает оснований полагать, что некоторые когнитивные способности человека не

²⁰⁵ Penrose R. *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford: Oxford University Press, 1989.

²⁰⁶ К сторонникам аргумента неполноты можно отнести также британского философа Джона Лукаса (Lucas J.R. *Minds, Machines, and Gödel* // *Philosophy*. 1961. №36. P. 112–137.) и известного критика искусственного интеллекта Хьюберта Дрейфуса (Dreyfus H. *What Computers Can't Do*. New York. Harper&Row. 1972).

²⁰⁷ Lewis D. *Lucas against Mechanism* // *Philosophy*. 1969. №44. P. 231–243; Lewis D. *Lucas Against Mechanism II* // *Canadian Journal of Philosophy*. 1979. № 9. P. 373–376; Chalmers D. *Minds, Machines, and Mathematics* // *Psyche*. 1996. №2. P. 11–20; Feferman, S. *Penrose's Gödelian Argument* // *Psyche*. 1996. №2. P. 21–32; Putnam H. *The Best of All Possible Brains?* // *The New York Times*. 1994. №20. P.7-17; Shapiro S. *Truth, Mechanism, and Penrose's New Argument* // *Journal of Philosophical Logic*. 2003. №32. P.19–42.

могут быть смоделированы с помощью обращения к вычислительному описанию. Тем не менее, вопрос об объеме и пределах вычислительного моделирования остается сегодня открытым, и можно ожидать, что эта тема будет еще долго оставаться предметом научных исследований.

Как уже было отмечено ранее, начиная с 1980-х годов многие приверженцы классического компьютеризма стали отказываться от выбранной ими объяснительной стратегии в пользу коннекционистских или информационно-динамических моделей, которые, однако, также не лишены собственных концептуальных изъянов. Некоторые философы, как мы выяснили в Главе 2, и вовсе обходились без постулирования некоторого определенного вычислительного формализма, лежащего в основе их теории сознания, и использовали информационные категории лишь в качестве тематически нейтральных концептуальных конструктов. Критические возражения против подобных теорий продолжали появляться в философской и научной среде и были призваны продемонстрировать несостоятельность информационной интерпретации природы ментального опыта в целом. К наиболее важным из подобных аргументов принято относить аргумент тривиальности, аргумент о субъективном толковании информационных процессов, аргумент «Китайская комната» и аргумент о невозможности редукции феноменального сознания к структурно-функциональным информационным свойствам.

Рассмотрим подробнее данные аргументы.

3.2. Аргумент тривиальности и вычислительной реализации

Аргумент тривиальности базируется на допущении способности всякой физической системы выполнять некоторые вычисления или обрабатывать информационные сигналы, из чего следует, что всякая физическая система может быть описана как обладающая той или иной степенью разумности и ментальности. Данный аргумент основан на одноименном возражении против

функционализма, которое одним из первых сформулировал австралийский философ Ян Хинкфус, а затем его довод был представлен американским мыслителем Вильямом Лайканом в статье 1981 года «Форма, функция и чувство»: с точки зрения философов, даже родниковая вода, находящаяся в простом пластиковом ведре, на микроуровне может случайно реализовать какую-либо когнитивную программу в течение короткого промежутка времени. А если практически любой физический объект при любых условиях обладает подобной молекулярной активностью, то сторонники вычислительного функционализма вынуждены признавать положения панпсихизма, что, однако, кажется философам абсурдом²⁰⁸.

Позже схожее рассуждение встретится в многих работах Джона Сёрла, утверждающего, что даже штукатурку на стенах наших домов можно интерпретировать как вычислительную систему²⁰⁹. В книге «Открывая сознание заново» философ пишет:

«Трудно понять, как избежать следующих выводов: 1) для любого объекта существует некоторое описание этого объекта, такое, что, согласно данному описанию, этот объект является цифровым компьютером; 2) для любой программы и для любого достаточно сложного объекта существует некоторое описание объекта, согласно которому он реализует данную программу. Так, например, стена за моей спиной прямо сейчас реализует программу Wordstar, потому что существует некоторый паттерн движения молекул, который изоморфен формальной структуре Wordstar. Но если стена реализует Wordstar, и если это достаточно большая стена, то она реализует любую программу, включая любую программу, реализованную в мозге»²¹⁰.

²⁰⁸ Lycan W.G. Form, Function, and Feel // The Journal of Philosophy. 1981. №78. P. 45.

²⁰⁹ Searle J. The Problem of Consciousness // Consciousness and Language. Cambridge: Cambridge University Press. 2002. P.7-17.

²¹⁰ Searle J. The Rediscovery of the Mind. Cambridge, MA: MIT Press. 1992. P. 208-209.

Конечно, приведенные аргументы могут с первого взгляда показаться не очень убедительными. Во-первых, указанная возможность реализации вычисления происходит только в течение короткого временного промежутка и обусловлена какой-то счастливой случайностью. И хотя Серл отмечает, что тривиальная реализация будет происходить чаще и в течение более длительного периода, данное утверждение им никак не обосновывается. Во-вторых, неясно, какие системы можно считать «достаточно сложными»: если определять подобные системы как физические системы большого размера (что, по-видимому, подразумевает Серл), то почему мы должны полагать, что в них существует достаточно моделей физической активности, соответствующих любым вычислениям? Очевидно, что подобный аргумент не совсем логичен.

После отказа от своей теории машинного функционализма рассуждение о тривиальности приводит и Хилари Патнэм: в книге 1988 года «Репрезентация и реальность» философ поставил своей целью продемонстрировать, как возможно сопоставление состояний некоторой физической системы X с состояниями формального вычисления Y , а также как дать характеристику класса физических систем, которые уязвимы перед тривиальным объяснением. Философ полагает, что подобные физические системы являются «открытыми» (то есть необособленными) и находящимися в каузальном взаимодействии со своим окружением.

Патнэм рассуждает над примером с камнем: философ условно делит фазовые состояния камня на четыре части, соответствующие некоторым последовательным временным интервалам, и называет их r_1 , r_2 , r_3 и r_4 . Тогда, претерпевая многократные изменения в своем физическом состоянии, камень проходит следующие стадии трансформации своих состояний: $r_1 \rightarrow r_2 \rightarrow r_3 \rightarrow r_4$. Но, как замечает Патнэм, это не единственная возможная последовательность, ведь камень также претерпевает следующие изменения: $r_1 \vee r_3 \rightarrow r_2 \vee r_4 \rightarrow r_1 \vee r_3 \rightarrow r_2 \vee r_4$. Иными словами, наряду с перемещением через четыре соседних области своих фазовых состояний, камень также колеблется между двумя несвязанными областями своих фазовых состояний ($r_1 \vee r_3$ и $r_2 \vee r_4$), которые философ

сопоставляет с вычислительным состоянием А и вычислительным состоянием В соответственно²¹¹. В этом случае имеет место изоморфизм между физическими состояниями камня и формальными состояниями какого-либо условного конечного автомата, отчего можно сказать, что камень является конечным автоматом. То же самое рассуждение применимо и к другим открытым физическим системам²¹².

Философы, критикующие данный аргумент Патнэма, как правило, ссылаются на тот факт, что некоторые дизъюнкции фазовых состояний системы не могут служить кандидатами для реализации вычислительного состояния, и проблема заключается в трудности определения дизъюнкций, которые действительно являются подобными кандидатами. Кроме того, критики Патнэма утверждают, что большинство вычислений происходит над входной и выходной символической информацией, а аргумент философа применим только к конечным автоматам, не оперирующим подобной входящей и выходящей информацией, так как философ предполагает, что, хотя физически определенные входы и выходы будут частично ограничивать реализацию вычислений, эти вычисления все еще будут подвержены его аргументу²¹³. Наконец, еще одно возражение против аргумента Патнэма заключается в убеждении, что условный компьютер не должен разрушаться из-за произвольно малых физических изменений, однако нет никакой гарантии, что вычислительные состояния камня, описанные Патнэмом, сохранятся при небольших изменениях его физических качеств.

Аргумент тривиальности Патнэма лег в основу другого схожего аргумента, предложенного в 1996 году Дэвидом Чалмерсом. Австралийский философ

²¹¹ Отметим, что философ не видел ничего странного в отождествлении типа физического состояния системы с дизъюнкцией областей ее фазовых состояний: так, в случае компьютера, его физические состояния, реализующие вычислительные состояния, представляют собой разнообразные конфигурации электрических сигналов, которые могут возникать сразу в нескольких электронных компонентах.

²¹² Putnam H. Representation and Reality. Cambridge, MA. 1988. P. 84-87.

²¹³ Отдельный ответ на данное возражение против аргумента Патнэма, был предложен философом Питером Годфри-Смитом, который полагает, что входы и выходы вычислений редко определяются физически, а скорее описываются как абстрактные состояния в рамках абстрактного вычислительного формализма (например, в виде цифр, символов, строк и т.д.). Эти абстрактные состояния в принципе могут быть реализованы любым типом физического состояния (электрические сигналы, колебания в воздухе, повороты металлического винтика). Как абстрактные состояния, вычислительные входы и выходы кажутся столь же уязвимыми перед аргументом тривиальности Патнэма, как и внутренние вычислительные состояния (Godfrey-Smith P. Triviality Argument against Functionalism // Philosophical Studies, 2009. № 145. P. 273-295).

утверждает, что каждая физическая система, обладающая «часами» (компонентом физической системы, который проходит через последовательность физических состояний в течение некоторого временного интервала) и «циферблатом» (физический компонент системы с произвольным числом физических состояний, который, будучи помещен в одно из этих состояний, остается в нем в течение некоторого интервала времени), реализует любое вычислительное состояние конечного автомата²¹⁴. Философ полагает, что почти все открытые физические системы имеют своеобразные «часы» и «циферблат», а если по какой-либо причине физическая система не имеет «часов» или «циферблата», то их можно легко добавить. Чалмерс отмечает, что столь простое достижение тривиальных вычислений демонстрирует нам несостоятельность нашего понимания вычислительной реализации.

Проблема вычислительной реализации или имплементации (computational implementation) сводится к описанию условий, при которых физическая система выполняет те или иные вычисления (или выполняет обработку синтаксической информации); критерий подобной реализации может рассматриваться в качестве своеобразного «моста» между абстрактным математическим описанием вычислительного процесса и конкретной сферой его физических воплощений. Очевидно, что до тех пор, пока не будут определены условия, при которых физическая система способна к реализации вычислений, любым теориям, оперирующим информационной или вычислительной объяснительной стратегией, будут грозить приведенные аргументы тривиальности. При этом также очевидно, что искомое объяснение вычислительной реализации должно

²¹⁴ Стратегия Чалмерса состоит в том, чтобы определить соответствие между каждым формальным состоянием условного автомата и дизъюнкцией физических состояний $[X, Y]$ реализующей системы (где X соответствует номеру состояния часов, а Y – номеру состояния циферблата) и показать, что соответствующие физические состояния находятся в правильных контрфактических отношениях друг к другу. Так, если система запускается в физическом состоянии $[1, Y]$, она будет последовательно переходить к состояниям $[2, Y]$, $[3, Y]$ и т.д., по мере продвижения часов. Предположим, что система начинает свой фактический запуск в состоянии набора номера 1. Тогда начальное состояние автомата может быть представлено как $[1, 1]$, а последующие абстрактные состояния - как $[2, 1]$, $[3, 1]$ и т.д. Если некоторые состояния автомата не возникли в конце данного процесса, то данное состояние можно выбрать в качестве нового начального состояния автомата и сопоставить ему значение $[1, 2]$. Так можно продолжать до тех пор, пока не будут рассмотрены все нереализованные состояния автомата (Chalmers D.J. Does a Rock Implement Every Finite-state Automaton // *Syntese*, 1996. №108. P. 309-333).

оперировать терминами, отличными от вычислительных и информационных и не должно ставить истинность своих суждений в зависимость от человеческих убеждений.

Многие философы являются приверженцами подхода, согласно которому достаточным условием для реализации вычисления является существование изоморфизма между физическими состояниями рассматриваемой системы и абстрактными вычислительными состояниями. Современный философ Марк Спревак формулирует этот закон следующим образом:

«Физическая система X реализует формальное вычисление Y , если существует соответствие f , которое сопоставляет физические состояния X абстрактным состояниям формального вычисления Y , так что для каждой ступенчатой эволюции $S \rightarrow S'$ формализма Y выполняется следующее условие: если X находится в физическом состоянии s , где $f(s) = S$, то X войдет в физическое состояние s' такое, что $f(s') = S'$ »²¹⁵.

Этот принцип простого соответствия (simple mapping account), безусловно, не исключает возможности тривиальной вычислительной реализации, а потому требует существенной концептуальной доработки. Среди возможных вариантов развития данного принципа в современной философии преобладают каузальный, семантический, телеологический (функциональный) и антиреалистский подходы²¹⁶.

Согласно каузальному подходу, параметром, имеющим значение для реализации вычислительных процессов, является структура каузальных связей в рамках некоторой физической системы. Подобного принципа придерживаются

²¹⁵ Sprevak M. Triviality Arguments About Computational Implementation // The Routledge Handbook of The Computational Mind (ed. by M. Sprevak, M. Colombo). NY: Taylor&francis Group, 2019. P. 177.

²¹⁶ Помимо перечисленных вариантов существует еще несколько возможных способов развития закона соответствия, которые, однако, не столь распространены среди современных сторонников вычислительных подходов (к примеру, уже не раз упомянутый нами философ Гвалтеро Пичинини придерживается так называемого «механистического» представления о природе вычислительной реализации (Piccinini G. Physical Computation: A Mechanistic Account. Oxford University Press. 2015)). Некоторые философы и вовсе полагают, что закон простого соответствия не нуждается в дополнении, и вполне совместим с определенными вычислительными формализмами (Schweizer P. Triviality Arguments Reconsidered. Minds & Machines. 2019. №29. P. 287–308).

американские философы Нед Блок²¹⁷, Джон Хугеланд²¹⁸, Джон Коупланд²¹⁹, а также Дэвид Чалмерс, концепция которого будет подробнее изложена в Главе 4 настоящего исследования.

Согласно семантическому подходу, только физические системы, манипулирующие репрезентациями, реализуют вычисления. Так, еще Джерри Фодор утверждал: «*Не бывает вычислений без репрезентаций*» («*There is no computation without representation*»²²⁰). Помимо Фодора представления о том, что вычислительная идентичность физической системы определяется тем, какими репрезентациями она манипулирует и как она это делает, придерживаются многие другие философы²²¹ (положения данного принципа частично легли в основу продолжительной дискуссии вокруг мысленного эксперимента Джона Серла «Китайская комната», о котором речь пойдет ниже).

Телеологический подход основывается на утверждении о том, что вычисления и информационную обработку реализуют только физические системы с определенными реляционными функциями, связанными с использованием прагматической информации и зависящими от истории развития системы, ее отношений с окружающей средой, и условными целями, с которой она была создана. Как нетрудно заметить, подобного подхода придерживаются многие рассмотренные выше философы, являющиеся сторонниками информационных теорий сознания. Представляется, что данный подход предполагает существование достаточно обширного набора естественных телеологических функций, необходимых для обоснования всех вычислительных процессов, известных современной когнитивной науке и

²¹⁷ Block N. Searle's Arguments Against Cognitive Science // Views Into the Chinese Room: New Essays on Searle and Artificial Intelligence (eds. Preston, Bishop). Oxford University Press, 2003. P. 70-79.

²¹⁸ Haugeland J. Syntax, Semantics, Physics // Views Into the Chinese Room: New Essays on Searle and Artificial Intelligence (eds. Preston and Bishop), 2002. P. 379-392.

²¹⁹ Copeland J. What is Computation? // Synthese, 1996. №108. P. 335-359.

²²⁰ Fodor J.A. The Mind-Body Problem // Scientific American, 1981. №244. P. 118.

²²¹ Shagrir O. Brains as Analog-Model Computers // Studies in History and Philosophy of Science, 2010. №41. P. 271-279; Ladyman J. What Does it Mean to Say that a Physical System Implements a Computation? // Theoretical Computer Science, 2009. №410. P. 376-383; Rescorla M. A Theory of Computational Implementation // Synthese, 2014. №191. P. 1277-1307; Sprevak M. Computation, Individuation and the Received View on Representation // Studies in History and Philosophy of Science, 2010. №41. P. 260-270.

информатике (хотя некоторые исследователи, напротив, полагают, что подобный набор естественных функций весьма ограничен²²²).

Наконец, антиреалистский принцип подразумевает, что вычислительная и информационная идентичность некоторой физической системы определяется сугубо субъективными запросами условного наблюдателя. Джон Серл и Хилари Патнэм пытались продемонстрировать, что любая вычислительная реализация, не ограниченная человеческими интересами и ценностями, тривиальна: камни и стены не реализуют вычислений, несмотря на возможное наличие изоморфизма между их физическими состояниями и некоторыми абстрактными информационно-процессуальными схемами, потому что человек не может использовать данное соответствие для выполнения необходимой ему вычислительной работы. Учитывая, что вычислительное объяснение в когнитивной науке расценивается как способ понимания сознания в объективных и не зависящих от разума терминах, возможность подобной трактовки вычислительной реализации признается серьезным аргументом против рассматриваемой установки (подробнее данный аргумент будет рассмотрен далее).

Разумеется, не следует отрицать возможности применения плюралистического подхода к определению критериев информационной и вычислительной реализации, согласно которому ограничивающий реализацию фактор может варьироваться в зависимости от того или иного контекста. Однако необходимо учитывать, что в условиях подобной интерпретации мы будем вынуждены признать, что в некоторых случаях, в зависимости от принимаемой исследователями позиции, одна и та же физическая система может одновременно выполнять и не выполнять вычисления или обработку информации.

Таким образом, дискуссии о выборе возможного ограничительного принципа, способного пролить свет на понимание необходимых критериев, которым должна удовлетворять та или иная физическая система для реализации

²²² Shea N. Naturalising Representational Content // *Philosophy Compass*, 2013. №8. P. 496-509.

вычислений и обработки информации, будут, скорее всего, продолжаться еще долго. Некоторые философы, однако, не находят аргумент тривиальности достаточно весомым: хотя утверждение о том, что каждая физическая система способна к вычислению, приводит нас к панкомпьютационализму, эта позиция сама по себе не является, с их точки зрения, противоречивой, и, следовательно, может рассматриваться в качестве одного из возможных следствий из вычислительной установки²²³.

3.3. Вычисление и обработка информации как субъективные процессы

Как уже было отмечено выше, согласно антиреалистской установке по отношению к вычислительной реализации или информационной обработке, данные процессы не являются естественными и объективными и тесно связаны с суждениями условного наблюдателя или пользователя, которые, в свою очередь, основаны на принципах, принятых в современной ему практике. Подобная установка, безусловно, подрывает веру в возможность объективного описания когнитивных процессов в информационных терминах.

Как ни парадоксально, утверждение о том, что вычислительные процессы напрямую зависят от соответствующего толкования субъекта, признавалась многими сторонниками как классического компьютеризма, так и многих других информационных теорий сознания (вспомним, что еще Зенон Пилишин полагал, что в силу тезиса множественной реализуемости возможно говорить о потенциально бесконечном наборе физических состояний, которые, при тех или иных условиях, могут лежать в основе вычислительных процессов, отчего в любую вычислительную теорию необходимо вводить особую функцию наблюдателя, способного интерпретировать те или иные изменения в физических состояниях системы как вычисления; схожей позиции

²²³ К примеру, Чалмерс считает, что панкомпьютационализм должен настораживать исследователей только в том случае, если он постулирует способность каждой физической системы к осуществлению любого (а не только тривиального) вычисления (Chalmers D. A Computational Foundation for the Study of Cognition // The Journal of Cognitive Science, 2011. №12. P. 323–357).

придерживается и Дэниел Деннет²²⁴). Однако воспринимать подобные доводы в качестве действительной угрозы всем вычислительным подходам к объяснению сознания, претендующих на вполне научное и объективное описание работы мышления, философское сообщество стало лишь после последовательной критики данных подходов Джоном Серлом.

В книге «Открывая сознание заново», американский философ утверждает, что описание процесса как вычислительного всегда осуществляется «снаружи»: *«Вычислительные состояния не выявляются внутри физических свойств, они приписываются физическим свойствам»*²²⁵. Описание информационно-вычислительных процессов может выполнить только условный «гомункул», обладающий интенциональными состояниями, в то время как ни компьютер, ни мозг не осуществляют вычисления сами по себе. Тот же тезис Серл продолжает развивать и своих последующих работах. Так, в статье 2002 года «Проблема сознания» он пишет: *«Вычисление существует только относительно некоторого агента или наблюдателя, который дает вычислительную интерпретацию некоторому явлению. Это очевидный момент»*²²⁶. Подобный довод философ также использует в своей более поздней статье 2013 года «Может ли информация объяснить сознание?», в которой он попытался опровергнуть любой подход, основанный на информационной интерпретации природы сознания:

«Субъективные сущности существуют лишь когда они воспринимаются человеческим или животным субъектом.

²²⁴ Размышления философа об абстрактном характере вычислительных процессов представлено в его интервью 2001 года: «Идея вычисления - это туманная идея, и ошибочно думать, что у нас есть ясное и единое понимание вычисления. Даже ученые-компьютерщики имеют лишь смутное представление о том, что они на самом деле подразумевают под вычислениями; это одна из тех вещей, которые мы узнаем, только когда видим [...] Паттерны возникают везде, где есть вычислительные устройства, где есть организмы, обрабатывающие информацию. Они сохраняют и восстанавливают паттерны, а также поддерживают их в рабочем состоянии. И это действительно фундаментальная, новая особенность Вселенной. [...] Некоторые процессы, очевидно, являются вычислительными. Другие, очевидно, не являются. Где же граница между этими процессами? Это зависит от того, кто смотрит». (The Computational View. A Talk with Daniel C. Dennett. 2001. [Электронный ресурс] // Edge.com. URL: https://www.edge.org/conversation/daniel_c_dennett-the-computational-perspective.

²²⁵ Серл Д. Открывая сознание заново. М., 2002. С. 195.

²²⁶ Searle J. The Problem of Consciousness // Consciousness and Language. Cambridge: Cambridge University Press, 2002. P.17.

Онтологически объективные сущности существуют независимо от всякого опыта. Поэтому боль, щекотка, зуд, подозрения и впечатления являются онтологически субъективными, в то время как горы, молекулы и тектонические плиты – онтологически объективными. [...] Это различие лежит в основе другого различия – между теми свойствами мира, которые существуют независимо от человека, и теми, для чьего существования необходим человек. Я называю это различием между независимыми от наблюдателя свойствами и свойствами, зависимыми от наблюдателя»²²⁷.

По мнению философа, сознание представляет собой онтологически субъективную сущность, и для его объяснения следует опираться на явления, чье содержание независимо от субъективных интерпретаций, ведь любое зависящее от интерпретации явление зависит и от сознания интерпретатора. Философ утверждает, что информация – это явление, полностью зависящее от некоторого условного наблюдателя, уже обладающего ментальностью. Поэтому, чтобы избежать объяснительного замкнутого круга, теория должна оперировать концепциями, в рамках которых информационные феномены трактуются как объективно существующие, что, однако, представляется философу невозможным.

Конечно, далеко не все философы признают доводы Серла. К примеру, сторонники структурно-каузального подхода к толкованию информационно-вычислительной реализации полагают, что информация является организационным паттерном материи и энергии, существующим независимо от условного наблюдателя, заинтересованного в «уменьшении неопределенности». С их точки зрения, если трактовать изоморфное соответствие между некоторыми физическими состояниями системы и абстрактными вычислительными состояниями как структурно-каузальное, то информационная и вычислительная

²²⁷ Searle J.R., Can Information Theory Explain Consciousness? [Электронный ресурс] // Hhimwich. 2013. URL: <https://www.hhimwich.com/files/can-information-theory-explain-consciousness.docx> (дата обращения 28.03.2019).

идентичность данной системы всегда остается неизменной, так как конечный физический объект всегда способен хранить и обрабатывать конечное количество информации в силу конечности своего энергетического потенциала²²⁸. Однако подобное возражение скорее говорит в пользу простого структурного реализма, нежели информационного. Бинарные коды, с помощью которых в науке описываются информационные процессы, назначаются извне в зависимости от некоторых принимаемых учеными условий, и, как выразился философ Джон Коупланд, «Если информационные структуры не присущи физике, и мозг не является “цифровым компьютером”, то это подразумевает, что ни одна сущность не является “цифровым компьютером”»²²⁹. Однако некоторые философы полагают, что структурный реализм также подразумевает и информационный реализм, что избавляет информационные теории от антиреалистской трактовки²³⁰.

Безусловно, признание структурно-каузального подхода к толкованию условий информационно-вычислительной реализации не является единственным возможным способом спасти информационные теории сознания от упомянутого Серлом концептуального замкнутого круга. К примеру, Д.И.Дубровский полагает, что главная ошибка аргумента Серла заключается в том, что тот не признает информационной обусловленности интенциональных состояний, что, с точки зрения российского философа, противоречит его теории интенциональности. Доводы Дубровского состоят в следующем: Серл полагает, что только субъект, уже обладающий интенциональностью и, соответственно, сознанием, может воспринимать информацию, в то время как, с точки зрения самого Дубровского, субъект всегда вынужден сначала оперировать некоторой информацией (своеобразной «инструкционной информацией»), которая, при должной обработке, порождает интенциональность, позволяющую ему иметь сознательные переживания²³¹ (вспомним, что похожий довод приводил и Кеннет

²²⁸ Fresco N. Objective Computation Versus Subjective Computation // Erkenntnis, 2015. № 80 (5). P. 1031-1053.

²²⁹ Copeland J. What is Computation? // Synthese, 1996. №108. P. 336.

²³⁰ Floridi L. A Defence of Informational Structural Realism. // Synthese. 2008. №161(2). P. 219-253.

²³¹ Дубровский Д.И. Новое открытие сознания? (по поводу книги Джона Серла «Открывая сознания заново») // Сознание, мозг, искусственный интеллект: сб. статей. М.: 2007. С. 66.

Сейр, когда утверждал, что информация является онтологически первичной по отношению к сознанию и всем его проявлениям). Так как многим исследователям представляется очевидным тот факт, что в ходе эволюционного процесса организмы научились обрабатывать информацию в целях управления раньше, чем у них появились разум и сознание, то схожей стратегии опровержения антиреалистских доводов относительно статуса информационных процессов придерживаются практически все сторонники функционально-телеологического подхода к определению критериев информационно-вычислительной реализации.

Подобные рассуждения о сущности информации представляются прямым следствием уже не раз упомянутого в рамках настоящего исследования смешения аспектов толкования данного понятия, произошедшего в науках в результате когнитивной революции 1950-1960-х годов. Споры о степени распространённости информации в мире и ее онтологическом статусе актуальны до сих пор, а убеждение в том, что высокоразвитые биологические организмы отличаются от всех прочих физических систем своей деятельностью по обработке информации «особого рода» (семантической или прагматической) может быть следствием слишком буквального толкования полюбившейся многим биологам и философам информационной метафоры, о которой упоминала историк науки Лили Кей. Трудности, связанные с приписыванием синтаксическим информационным структурам семантических и прагматических свойств, также нашли непосредственное отражение и в спорах о проблеме ментальных содержаний, которая будет описана далее.

3.4. «Китайская комната» и проблема семантики

Сформулированный Джоном Серлом в 1980 году в работе «Сознание, мозг и программы», мысленный эксперимент «Китайская комната»²³² стал одним из

²³² Searle J. Minds, Brains, and Programs // Behavioral and Brain Sciences, 1980. №3. P. 417–457.

самых известных аргументов в аналитической философии, направленных против использования компьютерной метафоры и, как следствие, возможности создания искусственного интеллекта. Кратко этот мысленный эксперимент можно представить следующим образом:

«Представьте себе незнающего китайского языка человека, запертого в комнате, полной коробок с китайскими символами (база данных) вместе с книгой инструкций по использованию символов (программа). Представьте, что люди за пределами комнаты посылают ему другие китайские символы, которые являются вопросами на китайском языке (вход). И представьте себе, что, следуя инструкциям в программе, человек в комнате может выдать китайские символы, которые являются правильными ответами на вопросы (выход). Программа позволяет человеку в комнате пройти тест Тьюринга на понимание китайского языка, но при этом сам человек не понимает ни слова по-китайски»²³³.

Основной довод Сёрла базируется на утверждении, что реализуемые искусственными вычислительными системами программы состоят из формальных синтаксических операций, чей ход не зависит от значений вычислительных символов, в то время как мыслительная деятельность человека всегда оперирует семантическими структурами.

В своей работе 1984 года «Разум, мозг и наука» Сёрл сформулировал трехпосылочный аргумент против вычислительных теорий сознания:

- «1) Программы являются чисто формальными (синтаксическими).*
- 2) Человеческий разум имеет ментальное содержание (семантику).*

²³³ Данная формулировка взята из: Searle J. The Chinese Room // The MIT Encyclopedia of the Cognitive Sciences (eds. R.A. Wilson, F. Keil), Cambridge, MA: MIT Press, 1999. P. 115.

3) Синтаксис сам по себе не является ни конститутивным, ни достаточным для семантического содержания.

4) Следовательно, сами по себе программы не являются конститутивными или достаточными для сознания»²³⁴.

Согласно мысленному эксперименту «Китайская комната», служащему доказательством третьей посылки, формальные символы сами по себе никогда не могут быть достаточными для наличия ментального содержания, потому что символы, по определению, не имеют никакого значения, кроме тех случаев, когда кто-то вне системы дает им семантическую интерпретацию. Однако большинство критических замечаний, направленных против выводов Серла, сводятся к отрицанию первой посылки: как полагают некоторые философы, информационные и вычислительные системы представляют собой сложные каузальные структуры, чья работа не сводится к оперированию исключительно синтаксическими символами, подобно тому, как это происходит, например, в формальной логике. Так, Деннет находит очевидным, что написанная программистами синтаксическая компьютерная программа, которая «просто лежит на полке», не способна к вычислениям и, тем более, к генерированию каких-либо ментальных состояний; для этого программа, описывающая тот или иной вычислительный процесс, обязательно должна быть запущена²³⁵.

Сегодня известно множество возможных способов опровергнуть доводы Серла. Так, в Стенфордской философской энциклопедии упоминаются следующие критические возражения: во-первых, некоторые исследователи признают, что сам человек в Китайской комнате не понимает китайского языка, но запуск программы может спровоцировать понимание системы в целом (подобной объяснительной стратегии придерживаются Нед Блок, Дэниел Деннет, Дуглас Хофштадер, Джон Хугеланд); во-вторых, вероятно допущение, что описанная Серлом комната может расцениваться как своеобразный

²³⁴ Searle J. *Minds, Brains and Science*. Cambridge, MA: Harvard University Press, 1984.

²³⁵ Dennet D. *Fast Thinking / The Intentional Stance*. Cambridge, MA: MIT Press, 1987. P. 327.

Подобного мнения придерживаются также философы Дэниел Кол (Cole D. *Thought and Thought Experiments // Philosophical Studies*. 1984. №45. P. 431–444) и Маргарет Боден (Boden M.A. *Escaping from the Chinese Room // The Philosophy of Artificial Intelligence* (ed. by. M. A. Boden). NY: Oxford University Press, 1990.).

компьютер, встроенный в тело робота и взаимодействующий с физическим миром посредством сенсоров (так называемый «аргумент робота» Фодора и Деннета), или система, которая последовательно копирует организационное устройство и работу мозга («аргумент симулятора мозга», которого придерживаются супруги Черчленд и Зенон Пилишин); в-третьих, некоторые критики утверждают, что запертый в комнате человек на самом деле прекрасно понимает китайский, ведь все зависит от того, каким образом мы определяем, что конкретно значит «понимать» («аргумент интуиции» Марка Спревака); наконец, некоторые философы утверждают, что если «Китайскую комнату» нельзя охарактеризовать как «понимающую» систему, то и про обычных людей, осуществляющих схожие действия, нельзя сказать, что они что-либо понимают («аргумент других умов» Томаса Хоргана)²³⁶.

Аргумент Серла, затронул вопрос о возможности описания семантических структур в терминах синтаксических манипуляций. Изначально отсутствие интереса к проблеме содержания, наблюдавшееся в среде первых специалистов, занимавшихся изучением и разработкой вычислительной техники, было связано не только с приверженностью данных исследователей к идеям операционализма, но и с тем, что первые представители классического компьютеризма включали в свои функционально-вычислительные теории семантические идиомы. Многие философы и ученые уповали на возможность выявить своеобразный закон соответствия, способный описать связь между физическими состояниями некоторого компьютера с логико-математическим описанием вычислительных состояний таким образом, чтобы сохранить определенные семантические отношения между данными описаниями; это означает, что вычислительные состояния описывали и индивидуализировали лишь уже существующие, предзаданные семантические свойства механизма (так называемую «инструкционную информацию»). Разумеется, подобный подход не показался многим философам убедительным, так как он не предоставлял теорию

²³⁶ Cole D. The Chinese Room Argument. [Электронный ресурс] // The Stanford Encyclopedia of Philosophy (ed. Edward N. Zalta). 2020. URL: <https://plato.stanford.edu/archives/win2020/entries/chinese-room/> (дата обращения: 16.12.2020).

семантического содержания: его основная идея заключалась лишь в приписывании вычислительным состояниям семантического контента, но не в определении, в силу каких свойств или условий данные состояния приобретают этот контент.

Одним из первых философов-компьютеристов, серьезно подошедших к проблеме ментального содержания, был Гилберт Харман. Будучи знакомым с работами Патнэма и Фодора, Харман решил объединить вычислительный функционализм с некоторыми идеями функционально-ролевой семантики Уилфрида Селларса²³⁷ для создания натуралистической теории семантически-индивидуализированных ментальных состояний²³⁸. Однако в рассуждениях Хармана был замкнутый круг, так как по его теории функциональные отношения ментального состояния индивидуализировались вычислительными отношениями, которые, в свою очередь, предполагали наличие семантических качеств. По этой причине теория нуждалась в разработке несемантического взгляда на вычислительный процесс.

Проект вычислительного репрезентативизма Джерри Фодора, а также информационные теории ментальных содержаний, представленные Фредом Дретске и его последователями, тоже были призваны объяснить, в частности, каким образом те или иные функциональные системы, оперирующие информационными структурами (будь то компьютеры или живые организмы), могут обладать внутренними семантическими состояниями. Как мы уже убедились, данные концепции сталкиваются с рядом характерных трудностей, помимо которых, однако, можно также выделить следующие проблемы:

1) Подобные теории не могут объяснить, как появляются значения у некоторых абстрактных объектов (например, у терминов, обозначающих

²³⁷ В теории Селларса ментальные состояния интерпретировались по аналогии с лингвистическими предложениями, а их содержание определялось отношением отдельных мыслей к лингвистическим и поведенческим стимулам, реакциям и другим мыслям, то есть конституировалось их функциональными и инференциальными отношениями (Sellars W. *Meaning As Functional Classification* // *Syntese*, 1974. № 27. P. 417-37.). Однако сам Селларс не указал механизм, который мог бы физически реализовать эти отношения. Вычислительный функционализм, в свою очередь, описывал механизм, который, казалось бы, обладает свойствами, необходимыми для реализации соответствующих функциональных отношений.

²³⁸ Harman G. *Wide Functionalism* // *Cognition and Representation* (ed. by S. Shiffer, S. Steele). Boulder: Westview. 1988.

математические и логические отношения или пустые термины вроде «кентавр»). Сторонники информационных теорий ментальных содержаний обычно объясняют это тем, что описанные ими условия являются достаточными для наличия смысла, но не необходимыми, так что представления абстрактных объектов получают свое значение каким-то иным способом. Другая стратегия защиты теории от подобных замечаний заключается в обращении к онтологии каузальных концепций ментальных содержаний, которая основана прежде всего на номологических отношениях. Можно сказать, что существует «закон кентавра», обеспечивающий появление соответствующего ментального представления, даже если на самом деле кентавров не существует. Эта стратегия, однако, не работает, когда мы говорим об индивидуальных ментальных представлениях²³⁹.

2) Хотя ряд информационных теорий преуспевает в объяснении того, почему некоторое ментальное состояние R репрезентирует условное свойство S_1 , а не, к примеру, свойство S_2 , они все еще не объясняют, почему данное состояние вообще репрезентирует некие свойства. (Данную проблему исследователи обычно именуют «метасемантической» проблемой ментальных содержаний²⁴⁰).

3) Хотя некоторые вычислительные системы связаны с окружающей средой, вопрос о том, является ли что-то информационно-вычислительным механизмом, зависит не только от его отношения со своим окружением. Следовательно, принятие информационных теорий ментальных содержаний само по себе не является поводом для интерпретации ментальных состояний в информационных и вычислительных терминах.

4) Некоторые философы считают, что большая часть ментальных содержаний каузально зависит от внутренних семантических состояний. Так, сам Джон Серл пользовался интерналистской трактовкой интенциональных качеств сознания и настаивал на независимости ментальных репрезентаций от свойств

²³⁹ Adams F., Aizawa K. Causal Theories of Mental Content. [Электронный ресурс] // The Stanford Encyclopedia of Philosophy (ed. Edward N. Zalta). 2017. URL: <https://plato.stanford.edu/archives/sum2017/entries/content-causal/> (дата обращения: 12.12.2019).

²⁴⁰ Artiga M., Sebastian M.A. Informational Theories of Content and Mental Representation // Review of Philosophy and Psychology. 2020. №11(3). P. 613-627.

объектов внешнего мира. В статье «Внутренняя интенциональность» философ отстаивает данное положение, приводя следующий пример:

«Если бы я был мозгом в чане, я мог бы иметь точно такие же ментальные состояния, как и сейчас; просто большинство из них были бы ложными [...] Вот что я имею в виду, когда говорю о том, что работа мозга каузально достаточна для наличия интенциональности, и что именно работа мозга, а не воздействие внешнего мира имеет значение для содержания наших интенциональных состояний»²⁴¹.

Серл считает, что информационные теории ментальных содержаний, использующие более широкое, экстерналистское понимание семантики, не способны предоставить решение сформулированной им проблемы: с точки зрения философа, как слово «собака», написанное на листке бумаги, не может обладать собственным внутренним семантическим содержанием, так и символы и информация, которыми манипулирует условный компьютер, не могут иметь внутреннего интенционального содержания, даже будучи каузально связанными с некоторыми объектами окружающей среды.

Одним из самых популярных ответов на представленное Серлом возражение является следующий аргумент: возможно, что сами по себе символы, которыми оперирует компьютер, не имеют внутреннего содержания, однако есть вероятность, что семантическое значение появляется в результате взаимодействия данных символов друг с другом. В этом случае содержание будет внешним по отношению к отдельным символам, но внутренним по отношению к вычислительной системе²⁴². А если паттерны взаимодействий между вычислительными элементами могут обладать внутренним содержанием, то вполне вероятно, что аргумент Серла не работает с некоторыми нейросетевыми вычислительными моделями.

²⁴¹ Searle J.R. Intrinsic Intentionality // Behavioral and Brain Sciences, 1980. №3. P. 452.

²⁴² Подобный аргумент одним из первых сформулировал Нэд Блок (Block N. Advertisement for a semantics for psychology // Midwest Studies in Philosophy (P. French ed.). University of Minnesota Press, 1986. №10. P. 615-678).

Трудности, возникшие на фоне решения проблемы ментального контента, побудили некоторых сторонников вычислительных теорий и вовсе отказаться от приписывания информационно-вычислительным состояниям каких-либо семантических свойств. Так, американский философ Стивен Ститч в середине 1980-х годов сформулировал синтаксическую теорию сознания: по мнению философа, для того, чтобы некоторая система осуществляла вычисления, необходимо наличие соответствия между ее релевантными для поведения физическими состояниями и классом синтаксических типов, заданным условной грамматикой, которая определяла бы, как сложные типы могут быть сформированы из примитивных типов. Согласно Стичу, мозг, понятый в этом смысле, является вычислительным механизмом²⁴³. Основная проблема подобного «семантического элиминативизма» заключается в том, что не вполне ясно, понимает ли компьютер или другая условная вычислительная система синтаксис или синтаксические операции: в самом компьютерном устройстве нет такого отдельного механизма, который был бы способен к обнаружению и интерпретации подобных операций, и лишь сторонний наблюдатель может описывать функциональные состояния компьютера как синтаксические.

Как уже было показано в Главе 2, попытку решить семантический вопрос предпринял Деннет с помощью своей версии интерпретационной семантики. Подобный подход, безусловно, скорее растворяет, чем решает проблему ментального содержания: очевидно, что в рамках концепции, предложенной американским философом, достаточно трудно объяснить разницу между «вычислительными» и «невычислительными» системами, равно как и индивидуализировать вычислительные механизмы и информационные состояния без помощи интенциональной установки. Кроме того, интерпретационная семантика, является недостаточной для характеристики ментального содержания как реального свойства психики – безусловно, это может не беспокоить тех, кто, подобно Деннету, изначально не верит в

²⁴³ Stich S. From Folk Psychology to Cognitive Science. Cambridge, MA: MIT Press, 1983.

реальность сознательного опыта, однако вынуждает других исследователей искать иные теории ментального контента.

Рассмотренные выше варианты, разумеется, не исчерпывают возможные пути развития подходов к интерпретации природы ментальных содержаний в рамках информационных и вычислительных концепций; тем не менее, каждая из существующих сегодня теорий ментального контента имеет ряд собственных нареканий. По этой причине некоторые исследователи предлагают создавать такие версии информационных теорий сознания, которые оставались бы нейтральными по отношению к описанию интенциональных качеств ментальных состояний и были бы сформулированы без использования семантического языка (как выразился философ Джон Хугеланд, *«Если будет синтаксис, то семантика позаботится о себе сама»*²⁴⁴). Такая стратегия, однако, не является удовлетворительным решением проблемы содержания, которая, безусловно, заслуживает дальнейшего изучения со стороны философов, психолингвистов и когнитивистов.

Отсутствие удовлетворительного решения по обозначенному вопросу побудило некоторых представителей аналитической философии частично перенаправить вектор своих исследований в сторону объяснения влияния феноменального опыта на интенциональность. Предположив, что приватные ощущения и субъективные ментальные переживания появляются еще до закрепления за ними концептуальной системы значений, выраженной в языке и коммуникации, такие философы, как Терри Хорган, Джон Гинсон, Дэвид Питт и др., сочли необходимым предварительно объяснить природу феноменального сознания, а затем приступить к проекту натурализации интенциональных состояний.

К подобному проекту, однако, исследователи вернулись лишь относительно недавно, когда начали активно разрабатываться теории, совмещающие в себе как концепции телеосемантической, информационно-каузальной и феноменальной интенциональности, призванные объяснить

²⁴⁴ Haugeland, J. *Artificial intelligence: The Very Idea*. MIT Press. 1985. P. 118.

появление простейших, доконцептуальных ментальных репрезентаций, так и некоторые идеи интерпретационной или функционально-ролевой (инференциалисткой) семантики²⁴⁵, лежащие в основе объяснения сложных ментальных представлений. Подобные теории, основанные на плюралистическом подходе к описанию природы ментальных содержаний, по мнению современных исследователей, способны пролить свет на проблему, столь насущную для философии сознания²⁴⁶. Однако, в силу того, что многие из данных теорий так или иначе опираются на некоторые положения концепции феноменальной интенциональности, перед ними по-прежнему стоит задача объяснения природы субъективных переживаний.

3.5. Критика функционализма и «провал в объяснении»

В силу того, что большинство информационных подходов к решению проблемы сознания основаны на функционалистской парадигме, все они так или иначе сталкиваются со многими присущими ей трудностями. Как уже было отмечено в Главе 1, американский философ Нэд Блок, один из самых последовательных критиков функционализма, отмечает, что объяснительная стратегия данной концепции представляется чрезмерно «либеральной», так как в ее рамках психические свойства могут приписываться объектам, которые на самом деле ими не обладают. Этот уже знакомый нам довод, демонстрирующий уязвимость функционализма перед аргументом тривиальности, Блок

²⁴⁵ Отметим, что в последние годы философы, занимающиеся проблемой интенциональности, предпочитают не использовать понятие «семантика», по преимуществу применяемое в лингвистике. По этой причине в современных исследованиях чаще встречаются понятия «интерпретативизм» и «инференциализм».

²⁴⁶ Речь идет, в первую очередь, о теориях, представленных в работах таких исследователей, как Карен Неандер (Neander K. A Mark of the Mental: A Defence of Informational Teleosemantics. Cambridge, USA: MIT Press. 2017), Анджела Менделовичи (Mendelovici A. The Phenomenal Basis of Intentionality. New York, USA: Oxford University Press. 2018), Николас Шиа (Shea N. Representation in Cognitive Science. Oxford University Press. 2018), Робби Уильямс (Williams J.R.G. The Metaphysics of Representation. Oxford University Press. 2020) и Адам Паутц (Pautz A. Consciousness meets Lewisian interpretation theory: A multistage account of intentionality. Oxford Studies in Philosophy of Mind (ed. by U. Kriegel). 2021). О достоинствах данных теорий упоминает, в частности, Дэвид Чалмерс в одной из своих последних статей (Chalmers D.J. Inferentialism, Australian style. 2021. [Электронный ресурс] // Philpapers.org. URL: <https://philpapers.org/archive/CHAIAS-8.pdf> (дата обращения: 27.10.2021).

подкрепляет мысленным экспериментом «Китайская нация». Данный эксперимент был впервые описан философом в его статье 1978 года «Проблемы с функционализмом»^{247,248}: Блок предлагает представить, что каждому китайскому гражданину предоставлен список радио каналов, и в заранее установленное время они должны связаться по радио с тем, кто находится в их списке. Когда любой гражданин получает сигнал, он связывается с тем, кто был в его списке, а те, в свою очередь, связываются с другими. Списки вызовов выстроены так, чтобы паттерны сигналов соответствовали паттернам нейронной активации, соответствующим тем или иным ментальным состояниям. Вызов по радио выполняет ту же функцию, что и нейронный сигнал, но при этом, отмечает Блок, вряд ли можно утверждать, что совокупность людей, совершающих подобные вызовы, испытывает схожие ментальные переживания (если переложить данный довод на язык вычислительной установки, то получается, что мы не можем с уверенностью утверждать, что условный компьютер, находящийся в схожем с мозгом вычислительном состоянии, находится в том же ментальном состоянии, что и обладатель мозга).

Действительно, проблема приписывания функциональным состояниям (даже если они трактуются в информационных или вычислительных терминах) ментальных качеств представляет собой большую угрозу для функционализма, в рамках которого сознательные состояния описываются исключительно в реляционных и структурно-каузальных терминах. Как полагает Блок и ряд других философов, подобные характеристики не способны объяснить появление феноменальных состояний или квалиа.

В другом важном антифункционалистском аргументе, так называемом «аргументе знания», в разных формах представленном философами Томасом Нагелем в его мысленном эксперименте «Летучая мышь» и Фрэнком Джексоном в «Ученом Мэри», также показано, что даже обладая всеми знаниями о функциональной и физической организации некоторого носителя ментального

²⁴⁷ Block N. Troubles with Functionalism // Minnesota Studies in the Philosophy of Science, №9, 1978. P. 261-325.

²⁴⁸ Стоит отметить, что похожий мысленный эксперимент был впервые описан ранее, в произведении 1961 года советского писателя-фантаста А.П. Мицкевича (А. Днепров) «Игра».

состояния, мы не можем знать «каково это быть» в данном состоянии. Аналогичным образом, условный компьютер, в котором хранится вся информация о некотором феноменальном состоянии человека, вряд ли может знать, о том, каково это переживать данное состояние, ведь оно выражается в терминах субъективных переживаний, а не в терминах, выражающих факты, доступные стороннему наблюдателю, обладающему лишь точкой зрения «от третьего лица».

Также широко известен аргумент инвертированного спектра американского философа Сидни Шумейкера, согласно которому возможно представить человека, находящегося в определенном функциональном состоянии, соответствующем некоторому феноменальному переживанию (например, переживанию зеленого цвета), но при этом испытывающего иное феноменальное переживание (к примеру, переживание красного цвета), что, тем не менее, никак не отражается на его поведении²⁴⁹. Данный аргумент призван показать, что рассматриваемые функциональные и реляционные характеристики (равно как и информационно-вычислительные) не могут служить основаниями различия между феноменальными переживаниями с изоморфными каузальными паттернами (вспомним доводы Патнэма, отказавшегося от своей теории «психологического компьютеризма» по этим же причинам). Схожее возражение заключается и в «аргументе отсутствующих квалиа», согласно которому возможно существование существа, чьи психические состояния вообще не имеют качественного характера. Тем не менее, некоторые философы отмечают, что, хотя функциональные дубликаты человека с инвертированными квалиа в принципе возможны, то вероятность существования дубликатов с отсутствующими квалиа может быть оспорена, так как подобные сценарии являются контрпримерами лишь к «грубым» функционалистским теориям и обусловлены лишь силой воображения философов²⁵⁰. Следует также упомянуть «аргумент зомби», наиболее известную версию которого представил Дэвид

²⁴⁹ Shoemaker S. The Inverted Spectrum // Journal of Philosophy. 1982. № 79. P. 357-381.

²⁵⁰ Levin J. Functionalism and the Argument from Conceivability // Canadian Journal of Philosophy, 1985. №11. P. 85-104.

Чалмерс. Аргумент австралийского философа состоит в том, что возможно представить абсолютного функционального двойника человека, который, однако, не обладает феноменальным сознанием; подобная представимость, в свою очередь, означает реальную метафизическую возможность существования подобного «зомби» (как известно, философ основывается на положениях теории «двумерной семантики» Сола Крипке), из чего следует, что любое редукционистское описание природы сознания ошибочно.

Безусловно, предложенный аргумент до сих пор расценивается многими философами как спорный, однако, вслед за философом Джозефом Левином можно утверждать, что даже если представимость зомби не должна побуждать нас к полному отказу от положений функционализма, аргумент «отсутствующих квалиа» все равно демонстрирует, что в функциональном объяснении ментального опыта существует значимый пробел: так как из функционально-структурного описания работы нашей когнитивной системы не следует, что данная система способна к феноменальным переживаниям, то перед соответствующими теориями встает, как минимум, серьезная эпистемологическая проблема. Конечно, можно предположить, что на сегодняшний день зомби и инвертированные квалиа представимы, но, учитывая рост эмпирических знаний, ученые вскоре разуверятся в теоретической возможности подобных явлений. Однако некоторые современные философы, такие как Колин Макгинн, отрицают, что мы в принципе способны сформулировать удовлетворительную эмпирическую теорию сознания в силу ограниченности наших когнитивных возможностей²⁵¹.

Одной из возможных стратегий защиты функционального объяснения сознания является принятие позиции своеобразного элиминативизма, согласно которому убежденность в существовании квалиа вызвана иллюзией или заблуждением, – помимо Деннета подобной позиции придерживается Кит

²⁵¹ McGinn C. How Not To Solve the Mind-Body Problem // Physicalism and Its Discontents (ed. by C. Gillett, B. Loewer). Cambridge University Press. 2001. P. 284-307.

Франкиш²⁵² и некоторые другие современные философы и когнитивисты. Однако очевидно, что для тех исследователей, которые признают реальность квалиа, описанные выше аргументы являются существенной преградой на пути к построению последовательной функционалистской теории сознания. Тем не менее, далеко не все философы, признавшие возражения против строгого (редуктивного) функционализма, отказались также и от информационной интерпретации природы сознания, которая получила свое развитие даже после принятия вызова «трудной проблемы».

Заключение главы

Помимо внутренних концептуальных трудностей, присущих отдельным версиям информационных теорий сознания, многие из них сталкиваются с рядом проблем, характерных для подобных теорий в целом. К ним относится проблема тривиальности, согласно которой любая физическая система может рассматриваться как вычисляющая или обрабатывающая синтаксическую информацию, а также вытекающая из нее проблема информационно-вычислительной реализации. Несмотря на то, что большинство сторонников информационных подходов к объяснению природы сознания признают данные проблемы, между современными философами до сих пор нет консенсуса, касающегося выбора релевантного принципа, с помощью которого можно было бы провести границу между «вычисляющими» и «не-вычисляющими» системами. Аналогично, нет единого мнения и о том, как следует решать проблему синтаксиса и семантики, представленную в знаменитом аргументе «Китайская комната» Джона Серла: несмотря на наличие достаточно широкого спектра теорий, призванных объяснить появление семантических содержаний у информационных систем, каждая из них имеет ряд собственных нареканий. Более того, в ответ на критику своего мысленного эксперимента, Серл

²⁵² Frankish K. Illusionism as a Theory of Consciousness // *Journal of Consciousness Studies*. 2016. №23 (11-12). P. 11-39.

продолжил развитие аргумента против информационных теорий сознания, утверждая, что описание тех или иных структур как информационных полностью зависит от практической установки условного наблюдателя. Возможность антиреалистской трактовки основополагающих для информационных теорий категорий требует от философов и других исследователей поиска доказательств объективного течения процессов вычисления и информационной обработки, независимых от изначального существования сознания. Наконец, одним из самых важных аргументов против информационных теорий сознания, основанных на функционалистской парадигме, является возражение, согласно которому феноменальные качества сознания невозможно редуцировать к функциональным и вычислительным свойствам: такие известные философские аргументы, как «аргумент знания», «аргумент зомби», «аргумент инвертированного спектра» и т.д., демонстрируют, что функциональная установка непременно сталкивается с «провалом в объяснении». И несмотря на то, что многие упомянутые выше сторонники информационных теорий сознания использовали понятия «вычисление» или «информация» как раз для того, чтобы найти тематически нейтральный язык для описания как функциональных, так и субъективных качеств ментальных процессов, их теории все еще не дают последовательного ответа на вопрос «Почему информационные процессы, происходящие в нашей когнитивной системе, сопровождаются характерными феноменальными переживаниями?».

Таким образом, любой информационной теории, претендующей на решение проблемы сознания, необходимо ответить на вопросы, которые можно условно поделить на две группы: в первую группу входят вопросы, касающиеся трактовки основополагающих информационных и вычислительных категорий, а также связанных с ними понятий информационного синтаксиса и семантики, вычислительной реализации, обработки информации, и т.д.; во вторую группу входит вопрос о связи информационных процессов с феноменальным опытом. Попытку развития последовательной теории, отвечающей на вопросы из обеих представленных групп предприняли некоторые современные аналитические

философы, в числе которых Дэвид Чалмерс, чья концепция будет подробнее изложена в следующей главе.

Глава 4. Информационные теории сознания после «трудной проблемы»: «минимальный компьютеризм» и двухаспектный принцип информации

Представленные выше проблемы, с которыми обыкновенно сталкивается большинство информационных теорий сознания, побудили современных философов искать не только новые способы толкования основополагающих для подобных теорий категорий, но и новые варианты объяснения природы возможной связи, существующей между информационными и феноменальными состояниями. Так, уже упомянутый выше философ Дэвид Чалмерс предпринял попытку построить модель последовательной теории сознания посредством выдвижения ряда подходов и принципов, призванных прояснить суть информационных и вычислительных процессов, объяснить их связь с феноменальными переживаниями и выявить эпистемологическую роль информационного объяснения природы сознания в целом. Далее в настоящей главе будут подробно рассмотрены данные подходы и принципы, среди которых особое место занимают так называемый «минимальный компьютеризм», двухаспектный принцип информации и концептуальный структурный реализм.

Ввиду того, что некоторые идеи, выдвинутые австралийским философом, были восприняты многими современными исследователями, в данной главе будут также проанализированы теории сознания таких аналитических философов, как Грегг Розенберг, Риза Мали, Сэм Колман и др., и кратко разобраны информационные концепции сознания, авторами которых являются представители когнитивных наук. Подобный анализ призван наглядно продемонстрировать современное состояние информационных теорий сознания и поможет точнее сформулировать необходимые условия дальнейшего развития подобных теорий не только в философии, но и в естествознании.

4.1. «Минимальный компьютеризм» Дэвида Чалмерса

Несмотря на то, что мировую известность Чалмерсу принесло его выступление на Туссанской конференции в 1994 году и последующая за ним публикация книги «Сознающий ум» в 1996 году, австралийский философ выдвинул ряд немаловажных философских идей, позже нашедших свое отражение в его более зрелых теориях, еще до середины 1990-х годов.

Получив образование математика, в 1989 году Чалмерс поступил в университет Индианы, где стал изучать философию сознания под руководством американского когнитивиста и информатика Дугласа Хофштадера, оказавшего существенное влияние на становление взглядов юного философа. Еще будучи студентом математического отделения в Оксфорде, Чалмерс написал статью «Разум, паттерн и информация»²⁵³, в которой, как он позже отметил, представлены некоторые идеи его будущей теории сознания, основанной на двухаспектном принципе информации. В данной статье философ отмечает, что любое исследование структурной организации мозга необходимым образом должно учитывать тот уровень теоретической абстракции, с которым мы подходим к данному исследованию. Любая абстракция ведет к обнаружению определенного организационного паттерна в структуре, неотделимого от информации, воплощенной в данном паттерне. Таким образом, основное утверждение Чалмерса заключалось в том, что структура разума представляет собой паттерн организации мозга, а любое ментальное состояние является информацией, содержащейся в данном паттерне; при этом представление о структуре разума как о паттерне соответствует описанию от третьего лица, а представление о ментальном состоянии как информации соответствуют описанию от первого лица. Описывая мозг как систему по обработке паттернов (brain as a pattern processor), Чалмерс отмечает, что искусственные системы, обрабатывающие схожие паттерны, также могут обладать сознанием,

²⁵³ Chalmers D.J. Mind, Pattern, Information. 1988 [Электронный ресурс] // Персональный сайт Д. Чалмерса. URL: <http://consc.net/papers/oxford2.pdf> (дата обращения: 10.04.2020).

независимо от субстрата, из которого состоят данные системы. По этой причине его концепция может быть вполне совместима с некоторыми положениями функционализма или компьютеризации, несмотря на то, что эти позиции, с точки зрения Чалмерса, основываются лишь на объяснении «от третьего лица». Тем не менее, философ полагает, что лучше всего его идеям подходит двуаспектная интерпретация, в рамках которой информация означает «каково это быть паттерном».

Уже в этой ранней статье Чалмерс затрагивает некоторые ключевые вопросы, встающие перед информационными теориями сознания, рассмотренными выше: философ говорит о необходимой связи информации с организационным паттерном и о соотнесенности ментальных состояний и информации (подобно тому, как Д.И. Дубровский утверждает о необходимой связи информации с ее кодовым воплощением и о возможной интерпретации явлений субъективной реальности как информации), а также о необходимости принятия реализма по отношению к абстрактным паттернам, несмотря на угрозу своеобразного «платонизма». Безусловно, в рамках данной прототеории Чалмерс еще не смог дать последовательного ответа на вопрос о том, как лучше следует толковать понятия «информация» и «паттерн», почему информация необходимым образом представляется нам в виде сознательных переживаний, а также где лежат границы подобного структурно-информационного объяснения. Тем не менее, философ усматривал в своей прототеории большой объяснительный потенциал, а потому продолжил развитие некоторых ее положений в последующих работах.

Во время обучения под руководством Хофштадера и работы в его исследовательской лаборатории по проблемам искусственного интеллекта, Чалмерс опубликовал значительное количество статей, связанных с проблемой описания мозга как вычислительной системы²⁵⁴. В 1992 году он также принял

²⁵⁴ Chalmers. D.J. Analog vs. Digital Computation. 1989. [Электронный ресурс] // Персональный сайт Д. Чалмерса. URL: <http://consc.net/notes/analog.html>; Chalmers. D.J. Computing the Thinkable // Behavioral and Brain Sciences, 1990. №13. P. 658-659; Chalmers. D.J. Syntactic Transformations on Distributed Representations // Connection Science, 1990. №2. P. 53-62; Chalmers. D.J., French R. M., Hofstadter D. High-Level Perception,

участие в e-mail-дискуссии, посвященной вопросу «Что такое вычисление?», организованной венгерским ученым-когнитивистом Стефаном Харнадом, по результатам которой написал статью «Вычислительные основания исследований мышления», окончательная публикация которой состоялась лишь в 2012 году²⁵⁵. В данной статье Чалмерс описывает свою версию вычислительного подхода к объяснению работы сознания, которую называет «минимальным компьютеризмом».

В начале статьи философ формулирует два тезиса, лежащих, по его мнению, в основе многих современных когнитивных теорий, и являющихся базовыми положениями отстаиваемой им позиции «минимального компьютеризма»: это тезис вычислительной достаточности (computational sufficiency), согласно которому обладание правильной вычислительной структурой достаточно для обладания разумом и ментальностью, и тезис вычислительного объяснения (computational explanation), в рамках которого утверждается, что вычислительное описание способно выступать в качестве основания для объяснения различных когнитивных и поведенческих процессов. Для того, чтобы доказать эти положения, необходимо прояснить, как абстрактное понятие «вычисление» связано с конкретными физическими качествами той или иной системы, то есть определить критерии, которым должна удовлетворять некоторая физическая система, чтобы осуществлять вычисления. Чалмерс полагает, что система реализует то или иное вычисление в том случае, если ее каузальная структура отражает формальную структуру данного вычисления (философ представляет данное положение, ссылаясь на метафору комбинаторного автомата²⁵⁶, чей формализм, с точки зрения Чалмерса, достаточен для описания различных типов вычисления). Таким образом, в

Representation, and Analogy: A Critique of Artificial Intelligence Methodology // Journal of Experimental and Theoretical Artificial Intelligence. 1992. №4(3). P. 185 – 211.

²⁵⁵ Chalmers. D.J. Computational Foundation for a Study of Cognition // Journal of Cognitive Science, 2012. №12(4). P. 323-357.

²⁵⁶ Комбинаторный автомат отличается от обычного конечного автомата только тем, что его внутреннее состояние описывается не единственным состоянием S (как например в случае с ленточными квадратами в машине Тьюринга), а вектором $[S^1, S^2, S^3, \dots]$, элементы которого можно рассматривать как компоненты общего состояния.

рамках подобной структурно-каузальной трактовки вычисление является всего лишь абстрактным описанием каузальной организации некоторой физической системы.

Чалмерс утверждает, что некоторые свойства системы, основанные на ее каузальной топологии, являются организационными инвариантами: если с системой происходят изменения, при которых сохраняется ее внутренняя каузальная структура (перемещение системы в пространстве, замена некоторых ее частей на функционально эквивалентные и т.д.), то сохраняются и данные свойства. Подобными свойствами являются, как полагает философ, и свойства мышления, которые определяются с точки зрения их каузальной роли (так называемые психологические свойства, к которым относятся убеждения, восприятие, память и т.д.). Феноменальные свойства, которые, казалось бы, не играют какой-либо каузальной роли, Чалмерс также предлагает рассматривать как организационные инварианты. Философ подкрепляет данную интуицию с помощью мысленного эксперимента «исчезающих квалиа»: с его точки зрения, в случае постепенной замены мозговых нейронов на функционально-эквивалентные кремниевые чипы, каузальная топология мозговых структур останется неизменной, и когнитивный агент, вероятно, не заметит каких-либо существенных изменений в своем поведении и переживаниях.

Очевидно, что подобный аргумент не может помочь в демонстрации логической и метафизической невозможности инвертированных или отсутствующих квалиа, как и доказать правомерность сильного функционализма, утверждающего, что сознание полностью зависит от функциональной организации когнитивной системы²⁵⁷. Однако, по мнению Чалмерса, данный аргумент преуспевает в установлении более слабой формы функционализма, которую философ называет нередуктивным функционализмом: в рамках данной версии функциональное устройство

²⁵⁷ Более подробно взаимосвязь аргумента «исчезающих квалиа» (в некоторых переводах используется выражение «блекнущие квалиа») с аргументом «отсутствующих» и «измененных» квалиа Чалмерс анализирует в своей статье 1995 года «Отсутствующие квалиа, исчезающие квалиа, танцующие квалиа» (Chalmers D.J. *Absent Qualia, Fading Qualia, Dancing Qualia // Conscious Experience*. (ed. by T. Metzinger). Imprint Academic. 1995. P. 309-328).

системы определяет структуру сознательных переживаний, которые, тем не менее, не обязательно сводятся к подобному устройству. Не являясь фундаментальной теорией сознания, эта форма функционализма потенциально совместима как с физикализмом, так и с определенными версиями теории дуализма свойств, нейтральным монизмом и т.д.

Таким образом, обозначенные выше положения о вычислительной достаточности и вычислительном объяснении доказываются с помощью ссылки на фиксируемость инвариантных свойств в рамках некоторой вычислительной структуры: так как организационно-инвариантные свойства системы изоморфны ее каузальной структуре, которая, в свою очередь, отражает формальную структуру некоторого вычисления, то и они изоморфны формальной структуре данного вычисления.

Представляется, что предложенный Чалмерсом подход решает сразу несколько проблем, характерных для вычислительных моделей сознания. Во-первых, философ отвечает на проблему вычислительной тривиальности: в рамках его концепции можно утверждать, что любая физическая система реализует некоторое вычисление (и, возможно, не одно), однако нельзя утверждать, что любая физическая система реализует любое вычисление; отсюда следует, что условная стена, о которой в свое время упоминал Джон Серл, не может реализовать программу «Wordstar», хотя вполне вероятно, что она реализует некоторые тривиальные вычисления. Подход Чалмерса также подразумевает, что, так как вычисления изоморфны некоторой объективно существующей каузальной структуре, то они не могут зависеть от сознания условного наблюдателя.

Во-вторых, философ частично отвечает на проблему содержания: реализация вычисления, с его точки зрения, является чем-то большим, чем простой набор синтаксических правил, «лежащих на полке»; сам по себе синтаксис, по утверждению философа, может быть недостаточен для появления семантики, в то время как правильно подобранная каузальная структура является вполне достаточной, так как и формальные вычисления, и их семантическое

содержание должны зависеть от лежащей в их основе каузальной структуры. Чалмерс также добавляет, что в случае мысленного эксперимента «Китайская комната», классический системный ответ можно переформулировать с помощью уже упомянутого аргумента «исчезающих квалиа»: данный ход призван показать, что мозг, каузальная организация нейронов которого сопоставима с каузальной структурой «Китайской комнаты», будет осуществлять схожие вычисления и будет разумным.

В силу того, что некоторые критики расценивают предлагаемую австралийским философом интерпретацию вычислительных процессов как достаточно абстрактную и потому неудовлетворительную²⁵⁸, сам Чалмерс отмечает, что «минимальный компьютеризм» представляет собой скорее методологическую и эпистемологическую программу, нежели полноценную теорию, а потому не связан с каким-то конкретным вычислительным формализмом и совместим с такими подходами, как классический компьютеризм, коннекционизм, теория динамических систем и т.п. Однако, несмотря на то, что представленный философом подход достаточно успешно обходит некоторые ключевые проблемы, встающие перед информационными и вычислительными теориями сознания (проблемы, которые в заключении Главы 3 мы условно поместили в первую группу), он не дает ответа на вопрос о том, почему некоторым вычислительным состояниям физической системы могут соответствовать феноменальные свойства. Иными словами, все еще остается нерешенной проблема, связанная с необходимостью определения релевантной теории сознания, призванной стать основанием концепции нередуктивного функционализма.

²⁵⁸ Более подробно их замечания изложены в статье Чалмерса 2012 года «Разновидности вычислений» (Chalmers D.J. The Varieties of Computation: A Reply // The Journal of Cognitive Science. 2012. №13. P. 213–248).

4.2. Двухаспектный принцип информации и «трудная проблема» сознания

Вопрос «Почему существуют квалиа?» стал центральным для докторской диссертации Чалмерса, в которой философ продолжил развитие некоторых своих ранних идей, касающихся интерпретации субъективных ментальных состояний и их связи с определенными структурно-организационными качествами нашей когнитивной системы. Текст диссертации Чалмерса послужил основой для его книги «Сознающий ум», увидевшей свет в 1996 году, в которой философ обозначил деление на «легкие» и «трудные» проблемы сознания, ставшее подспорьем для формулировки знаменитой «трудной проблемы» сознания²⁵⁹. Доказывая отсутствие логической супервентности ментальных свойств на физических, что, в свою очередь, показывало несостоятельность всех редукционистских подходов к объяснению сознания, Чалмерс последовательно разбирает многие из упомянутых в предыдущей главе мысленные эксперименты, и приходит к выводу о необходимости поиска релевантного психофизического закона, призванного объяснить причину существования квалиа. Так философ приходит к своей теории «натуралистического дуализма», которую предложил развивать в рамках двухаспектного подхода.

Как уже было ранее отмечено, Чалмерс начал симпатизировать данному подходу еще при создании своей прототеории в 1989 году, когда описал связь мозга и сознания как связь между «паттерном» и «информацией». Дальнейшее развитие эта идея получила в его статье 1990 года «Сознание и познание»²⁶⁰, где описываются достоинства двухаспектизма, согласно которому феноменальные и

²⁵⁹ Строго говоря, первая формулировка «трудной проблемы» сознания появляется уже в статье Чалмерса 1995 года «Лицом к лицу с проблемой сознания»: «Почему, когда наши когнитивные системы участвуют в обработке визуальной и аудиальной информации, мы получаем визуальный или аудиальный опыт: качество глубокого синего, ощущение ноты “до”? Как мы можем объяснить, почему существует нечто такое, что похоже на то, что мы называем “вынашивать мысленный образ” или “испытывать эмоции”? Широко распространено мнение, что опыт возникает на физической основе, но у нас нет хорошего объяснения тому, почему и как это так возникает. Почему физические процессы вообще должны способствовать появлению богатой внутренней жизни? С объективной точки зрения это кажется безосновательным, но, тем не менее, это так. Если какая-либо проблема и расценивается как проблема сознания, то именно эта» (Chalmers D. J. Facing up to the Problem of Consciousness // Journal of Consciousness Studies. 1995. №2(3). P. 202).

²⁶⁰ Chalmers. D.J. Consciousness and Cognition. 1990. [Электронный ресурс] // Персональный сайт Д. Чалмерса. URL: <http://consc.net/papers/c-and-c.html> (дата обращения: 11.04.2020).

физические явления представляют собой аспекты некоторой единой фундаментальной субстанции, а потому обладают равным метафизическим статусом и не могут быть сведены друг к другу²⁶¹.

В данной статье Чалмерс отмечает, что двуаспектные концепции хороши тем, что сочетают в себе как идеи своеобразного эпифеноменализма, так и идеи теории тождества в духе Уллина Плейса, Герберта Фейгла и Джона Смарта, избегая при этом некоторых присущих данным подходам недостатков. Так, можно сказать, что в рамках подобной «теории тождества», признающей субъективные состояния всерьез, последние, будучи лишь аспектом физически-реализованных функциональных состояний мозга, являются «как бы каузально действенными». Философ полагает, что такой подход выгоден тем, что не слишком расширяет нашу онтологию, особенно если принимать в качестве базовых аспектов такие «естественные» категории, как «паттерн» и «информация», где информация – это «то, на что похож паттерн изнутри» (философ иронично добавляет, что и сам не вполне понимает, что это означает, однако данная формулировка кажется ему неплохой). И хотя для описания паттернов состояний «от третьего лица», возможно использовать некоторые объяснительные модели функционализма или компьютеризма, для построения «моста» от данных паттернов к феноменальным состояниям, необходимо принимать двуаспектизм²⁶².

²⁶¹ Подобные двуаспектные теории стали особенно популярны среди современных аналитических философов, в числе которых Томас Нагель, Грегг Розенберг, Макс Вэлманс, Гален Стросон и др. Ряд исследователей утверждает, что двуаспектный подход имплицитно содержится в трудах таких мыслителей, как Джордано Бруно, Бенедикта Спинозы, Иоганна Готфрида Гердера, Артура Шопенгауэра, Эрнста Маха и др. (Skrbina D. Dualism, Dual-Aspectism and the Mind // Contemporary Dualism: A Defence (ed. A. Lavazza, H. Robinson). NY: Routledge, 2014. P. 220-245).

²⁶² Говоря о необходимости построения подобного «моста», Чалмерс упоминает проблему «Ум-Ум» («the Mind-Mind problem»), сформулированную американским лингвистом Реем Джекендоффом. Данная проблема заключается в различии между двумя уровнями описания физических состояний, «вычислительным» и «феноменологическим», соответствие между которыми еще не доказано. Джекендофф отмечает, что, хотя некоторые специалисты и полагают, будто феноменальный опыт магическим образом появляется при достижении некоторой вычислительной системой определенного уровня организационной сложности, у нас, тем не менее, нет никаких гарантий возникновения сознательного опыта из вычислительной осведомленности. Ученый выдвигает «теорию промежуточного уровня» («Intermediate-Level Theory»), призванную объяснить переход от подобной осведомленности к феноменальному переживанию (Jackendoff R. Consciousness and the Computational Mind. Cambridge, Massachusetts: MIT Press, 1987).

В «Сознающем уме» Чалмерс продолжает развитие двуаспектного принципа и представляет его следующим образом:

«Любое феноменальное состояние реализует какое-то информационное состояние, которое реализуется также в когнитивной системе мозга. И наоборот, по крайней мере некоторые физически реализованные информационные пространства таковы, что при любой физической реализации информационного состояния в этом пространстве данное состояние получает также и феноменальную реализацию»²⁶³.

Отношение реализации предполагает конститутивные отношения между свойствами разных порядков, что означает, что свойства второго порядка «наследуют» каузальные силы свойств первого порядка. Подобная трактовка, таким образом, способна решить проблему ментальной каузальности, связанную с угрозой эпифеноменализма.

Для определения информационного пространства Чалмерс пользуется уже известным нам слоганом МакКея-Бейтсона «различие, производящее различие», и указывает на то, что *«информационное пространство, связанное с физическим объектом, всегда является сообразным каузальному пути некоторого объекта и пространству возможных последствий в конце этого пути»²⁶⁴*. Подобная трактовка информации является вполне совместимой с интерпретацией процессов вычисления, которую философ представил в рамках своей концепции «минимального компьютеризма»: по Чалмерсу, и вычисление, и информация являются сообразными каузальной структуре некоторой физической системы (при этом данная сообразность предполагает не полное совпадение информационных и вычислительных структур, а скорее их корреляцию и изоморфизм, – так философ ответил на столь насущный для большинства информационных теорий сознания вопрос о соотношении категорий «вычисление» и «информация»). Так как и формально-синтаксические

²⁶³ Чалмерс Д. Сознательный ум. В поисках фундаментальной теории. Пер. В.В. Васильева. М.: УРСС: Книжный дом «ЛИБРИКОМ». 2013. С. 357.

²⁶⁴ Там же. С. 358.

вычислительные операции, и их семантическое содержание должны зависеть от лежащей в их основе каузальной структуры, то получается, что и под «различием, производящим различие», может одновременно пониматься и синтаксическая, и семантическая информация²⁶⁵.

Как отмечает философ, преимущество понятия «информация» заключается в том, что в информационных терминах можно легко истолковать высокоуровневые принципы структурной когерентности, постулирующие систематическую связь между сознанием и осведомленностью, а также уже известный нам принцип организационной инвариантности, констатирующий зависимость сознательного опыта от функциональной организации физической системы. Более того, Чалмерс полагает, что так как ментальные переживания представляют собой простейшие структуры различия, лежащие в основе любого информационного пространства, то даже сложные сознательные состояния относятся к подобным пространствам, соответствующим их внутренней комбинаторной структуре:

«Естественно предположить, что система, которая может непосредственно знать о своей локализации в информационном пространстве, не имея доступа к какому-либо дополнительному знанию, будет попросту маркировать свои состояния как изначально и элементарно различные, различающиеся по своему “качеству”. И, разумеется, следует ожидать, что эти различия

²⁶⁵ Это также объясняет, почему принцип МакКея-Бейтсона, который, к примеру, Деннет в своей книге «От бактерии к Баху и обратно» использует для определения семантической информации, Чалмерс также применяет и для описания синтаксических аспектов информации, когда упоминает теорию Клода Шеннона в «Сознающем уме».

Таким образом, достаточно распространенное мнение о том, что австралийский философ является сторонником сугубо синтаксического подхода к толкованию информации, не совсем верно: философ отмечает, что даже в концепции Шеннона предполагается, что количественная характеристика информации во многом зависит от свойств отправителя, что подразумевает необходимое присутствие некоторых семантических и прагматических составляющих в информационном синтаксисе (на этом, как мы помним, в свое время настаивал и Фред Дретске). Поэтому можно утверждать, что используемая Чалмерсом «каузальная» интерпретация является своего рода универсальной, так как она потенциально применима в рамках синтаксического, семантического и прагматического подхода к толкованию информационных пространств. Оставляя подобное неформальное определение информации, Чалмерс, по его собственному признанию, оставляет некоторые «возможности для маневра».

Разумеется, ряд исследователей находит подобное определение весьма абстрактным и, как следствие, непригодным в силу того, что оно способствует дальнейшему размытию концептуальных границ понятия «информация», начавшемуся в середине прошлого века (например: Harms W.F. The use of information Theory in Epistemology // Philosophy of Science. 1998. №65. P. 472-501).

будут “непосредственно” воздействовать на систему: она оказывается в этих состояниях, которые, в свою очередь, оказываются непосредственно доступными для организации последующих процессов; и, к примеру, в знании этой системы о том, в каком состоянии она находится, нет ничего выводного [...] Можно было бы возразить, сказав, что данная система могла бы быть сконструирована таким образом, что она имела бы доступ к указанной информации на уровне “догадок”, примерно как у субъектов со слепым зрением [...] Но это устройство, скорее всего, было бы неэффективным»²⁶⁶.

Объяснение того факта, что некоторые физические системы обладают доступом к собственным информационным состояниям, безусловно, во многом напоминает простое функциональное объяснение, однако Чалмерс не хочет ставить здесь точку, ведь подобная интерпретация феноменальных суждений еще не доказывает верность информационного объяснения самой природы сознания; более того, философ также не вполне уверен в выборе релевантной онтологической трактовки категории информации, как и в ответе на вопрос о степени распространенности информационных пространств в мире. Именно по этой причине Чалмерс больше не сопоставляет предложенный им двааспектный принцип с идеями теории тождества, а развивает его в духе нейтрального монизма Бертрана Рассела, описанного британским философом в его книге 1927 года «Анализ материи»²⁶⁷. Развитие двааспектного принципа информации как

²⁶⁶ Чалмерс Д. Сознательный ум. В поисках фундаментальной теории. Пер. В.В. Васильева. М.: УРСС: Книжный дом «ЛИБРИКОМ», 2013. С. 363.

²⁶⁷ Как известно, Рассел предположил, что основой как внешних, физических, так и внутренних, феноменальных, свойств, являются некие нейтральные по своему характеру «чтойности» («quiddities»), благодаря которым все объекты обладают своим «что». Большинство современных философов, среди которых и сам Чалмерс, предлагает развивать двааспектизм именно в такую наиболее очевидную форму монизма, как расселианский нейтральный монизм, разделяющий с двааспектизмом центральное утверждение о существовании некой фундаментальной реальности, которая не является сама по себе ни ментальной, ни физической (отметим, что некоторые философы полагают, что имеет смысл говорить о двааспектизме как об отдельной форме монистической теории - например: Skrbina D. Dualism, Dual-Aspectism and the Mind // Contemporary Dualism: A Defence. Ed. A. Lavazza, H. Robinson. NY: Routledge, 2014. P. 227-228). В своей статье 2013 года, посвященной панпсихизму, Чалмерс отмечает, что панпсихизм не обязательно должен быть расселианского типа; однако, по мнению философа, не-расселианский панпсихизм, согласно которому существуют микрофеноменальные свойства, не играющие

возможного фундаментального закона, способствующего нередуктивному объяснению природы феноменального опыта, с одной стороны, дает философу основания верить в потенциальную разрешимость загадки сознания, но, с другой стороны, ведет к некоторым контринтуитивным следствиям.

4.3. Неограниченный двухаспектный принцип информации и панпсихизм

Итак, развитие двухаспектного принципа может идти двумя разнонаправленными путями: первый путь соответствует признанию повсеместного распространения информационных пространств и, соответственно, феноменального опыта, в то время как второй заключается в поиске некоторых ограничений, накладываемых на свойства информационных структур, с которыми могут быть связаны феноменальные качества. Для Чалмерса не очевидно, что неограниченный двухаспектный принцип ошибочен, и потому философ предлагает всерьез рассмотреть подобное развитие его концепции, ведущее, как нетрудно догадаться, к панпсихизму.

Панпсихизм, как полагает Чалмерс, может быть либо конститутивным, по которому макроопыт полностью или частично конституируется микроопытом, либо эмерджентным, предполагающим сильную эмерджентность макроопыта на микроопыте²⁶⁸. Сам философ склоняется к конститутивному расселианскому панпротопсихизму А-типа²⁶⁹, учитывающему возможность того, что простейшие

микрофизической роли, сталкивается с очевидными проблемами ментальной каузальности. По этой причине Чалмерс рассматривает именно расселианский панпсихизм.

²⁶⁸ Сильная эмерджентность предполагает, что высокоуровневые свойства возникают из свойств более низкого уровня, но истины, касающиеся данных свойств, являются невыводимыми из истин о низкоуровневых свойствах (иначе говоря, они не являются концептуально или метафизически обусловлены истинами о свойствах более низкого порядка). В то же время, свойства высокого уровня слабо эмерджентны на свойствах низкого уровня, когда высокоуровневые свойства возникают из низкоуровневых, но истины, касающиеся данных свойств, являются неожиданными, учитывая те принципы, которые регулируют область свойств низкого уровня (Chalmers D. J. Strong and weak emergence // *The Re-Emergence of Emergence: The Emergentist Hypothesis From Science to Religion* (ed. by P. Davies, P. Clayton). Oxford University Press. 2006. P. 244).

²⁶⁹ Согласно конститутивному панпсихизму типа А, существует априорный закон, связывающий микрофеноменальный и макрофеноменальный опыт, в то время как конститутивный панпсихизм типа В утверждает, что существует апостериорная необходимая связь микрофеноменальных и макрофеноменальных качеств (Chalmers. D.J. Panpsychism and Panprotopsychism. 2013. [Электронный ресурс])

системы не испытывают полноценных ментальных переживаний, обладая лишь протофеноменальными свойствами. Подобная версия теории, с точки зрения философа, не только позволяет избежать некоторых проблем, характерных для физикализма и дуализма, но и не постулирует эмерджентизма. Однако перед подобным вариантом панпсихизма стоит так называемая проблема комбинации, сводящаяся к вопросу о том, каким именно образом протофеноменальные свойства, конституируют переживания более высокого уровня.

Проблема комбинации, в свою очередь, делится на три подпроблемы: проблему комбинации субъекта, проблему комбинации квалитативного опыта и смежную с ней «проблему палитры», а также проблему структурной комбинации. Наиболее вероятными решениями данных проблем Чалмерс считает так называемые «комбинаторные» ответы, среди которых возможны различные стратегии²⁷⁰.

Так, в качестве возможных решений проблемы субъекта философ предлагает рассматривать, к примеру, дефляционистский подход, отрицающий субъекта опыта как метафизически базовую сущность. Крайней формой подобного подхода является элиминативизм по отношению к субъекту опыта, по которому феноменальные свойства присущи неким условным «феноменальным потокам»; однако здесь также можно задаться вопросом о принципах объединения феноменальных свойств в единый поток. Более слабая форма дефляционизма признает существование субъектов опыта и отрицает лишь их

// Персональный сайт Д. Чалмерса. URL: <http://consc.net/papers/panpsychism.pdf> (дата обращения: 12.04.2020)).

²⁷⁰ Конечно, возможны и некомбинаторные способы решения обозначенных проблем, не подразумевающие, что микро-опыт конституирует макро-опыт. Помимо эмерджентного решения, философ сюда относит панпсихизм тождества (identity panpsychism), по которому макро-переживания тождественны микро-переживаниям (то есть макро-опыт уже присутствует на фундаментальном уровне, и никакая комбинация не требуется), и автономный панпсихизм (autonomous panpsychism), утверждающий, что макро-переживания автономны от микро-переживаний, поскольку они не конституируются, не возникают и не тождественны микро-переживаниям. Отдельно стоит упомянуть космопсихизм, в рамках которого признается существование единого макро-сознания вселенского масштаба (данная версия панпсихизма, как правило, сталкивается с проблемой «декомбинации», заключающейся в описании условий, при которых внутри данного сознания обособляются другие формы феноменального опыта, подобные человеческим) (Chalmers D.J. The Combination Problem for Panpsychism. 2017. [Электронный ресурс] // Персональный сайт Д. Чалмерса. URL: <http://consc.net/papers/combination.pdf> (дата последнего обращения: 12.04.2020)).

метафизическую фундаментальность, подразумевая под этим их производность, но такое решение, очевидно, не избавляет нас от проблемы комбинации.

Еще одной возможной стратегией решения проблемы комбинации субъекта является «комбинаторный инфузионизм», согласно которому комбинированное макрофеноменальное состояние «вытесняет» исходные микрофеноменальные состояния. Данный подход, однако, сталкивается со следующей проблемой: если мы трактуем подобную «инфузию» как синхронический процесс, то возникает вопрос о количестве субъектов, одновременно конституирующих одного макросубъекта; если же «инфузия» происходит в течение какого-то времени, то отношения между микро- и макросубъектами будут скорее каузальными, нежели конститутивными, а потому панпсихизм такого рода будет скорее эмерджентным, нежели конститутивным.

Существуют и другие возможные варианты решения проблемы субъекта, среди которых концепция феноменальной связки (*phenomenal bonding*), постулирующая существование определенных феноменальных отношений между микросубъектами; дефляционизм по отношению к осведомленности, согласно одной из версий которого феноменальные состояния, существующие повсеместно, осознаются только системами, обладающими определенной каузальной или функциональной организацией; наконец, панквалитизм, сталкивающийся, в свою очередь, с проблемой комбинацией качеств.

Проблема комбинации качеств, сводящаяся к вопросу о том, как микрокачества формируют макрокачества, а также о том, как из данного небольшого набора («палитры») микро- (или прото-) качеств появляются такие разнообразные макрофеноменальные качества. Проблема комбинации качеств состоит в следующем: если одна и та же сущность одновременно осознает «красноту» и «белизну», то вполне вероятно, что она осознает и «розовость», хотя остается непонятным, каким именно образом происходит комбинация этих качеств. Возможно, что осознание двух различных качеств двумя различными сущностями в единой сложной системе может привести к осознанию

совершенно нового качества, но данное предположение работает, только если мы принимаем дефляционизм по отношению к осведомленности. Что касается проблемы «палитры», то она, как полагает Чалмерс, имеет два основных класса решений – так называемое решение «малой палитры» и «большой палитры». Согласно первому, все макро-качества могут быть созданы только из нескольких микро-качеств, если мы найдем правильные базовые микро-качества; согласно второму, полный спектр макро-качеств включен в число микро-качеств, то есть существуют микро-качества, связанные с различными цветами, звуками, запахами, вкусами и так далее. С одной стороны, второе решение полностью устраняет необходимость в решении проблемы комбинации качеств, сводя ее к вопросу о том, как макросубъекты могут наследовать качества от микросубъектов; с другой стороны, подобное решение сталкивается с проблемой ментальной каузальности, ведь в рамках расселианского панпсихизма, микро-качества каузально эффективны в силу того, что служат в качестве «чтойностей» для микрофизических свойств, и, учитывая, что существует всего несколько фундаментальных микрофизических свойств и одна «чтойность» для каждого из них, только несколько микро-качеств могут быть каузально релевантными, в то время как остальные будут эпифеноменальными. Поэтому сам Чалмерс склоняется к первой стратегии решения.

Наконец, проблема структурной комбинации сводится философом к вопросу о том, почему макрофеноменальная структура не является изоморфной макрофизической структуре, которая, в свою очередь, конституируется микрофизической структурой (к примеру, макрофеноменальная структура нашего зрительного поля сильно отличается от макрофизической структуры соответствующих нейронных соединений в мозге). Наиболее очевидным ответом на подобный аргумент Чалмерс считает отсылку к изоморфизму, существующему между макрофеноменальной структурой ментального опыта и определенной информационной структурой в головном мозге: философу кажется очевидным, что принципы феноменальной композиции с большей вероятностью отражают структуру информационных пространств, нежели

какую-либо макрофизическую структуру. Разумеется, философ не пытается сформулировать законы феноменального отображения информационных структур, так как не находит их необходимыми. (Возможность подобного объяснения проблемы структурной комбинации во многом является причиной, по которой многие современные сторонники теорий конститутивного панпсихизма нередко прибегают именно к информационному концептуальному аппарату при описании своих идей²⁷¹).

Итак, сформулировав основные проблемы, встающие перед конститутивным панпсихизмом, а также возможные пути их решения, Чалмерс приходит к выводу о том, что наиболее перспективными представляются решения в духе аргумента феноменальной связки и дефляционизма по отношению к осведомленности, подкрепленные принципами информационной структурной организации. Некоторые стратегии решения проблемы комбинации в той или иной степени постулируют определенные ограничения, которые должны быть наложены на каузальную или функциональную организацию систем-обладателей макрофеноменального опыта. Принятие панпсихизма, таким образом, в некоторых случаях возвращает нас к более интуитивно правдоподобным версиям, связанным с поиском критериев, ограничивающих виды информационных структур, с которыми связан феноменальный опыт.

4.4. Ограниченный двааспектный принцип информации

Описывая основные положения двааспектного принципа информации в «Сознающем уме», Чалмерс не раз отмечает, что не будет пытаться формулировать однозначные и строгие критерии физической реализации

²⁷¹ См. например: Seager W. Consciousness, Information, and Panpsychism // *Journal of Consciousness Studies*. 1995. №2(3). P.272-288; Brüntrup G. Panpsychism and Structural Realism. *The Mental as Fundamental*. New Perspectives on Panpsychism (ed. by M. Blamauer). 2011. P.15-35; Roelofs L. Dennettian Panpsychism: Multiple Drafts, All of Them Conscious // *Acta Analytica*. 2021. P.1-18.

информации и будет использовать лишь упомянутый им принцип МакКея-Бейтсона. Тем не менее, философ также замечает, что двухаспектный принцип не обязательно должен выполняться для любого физического объекта:

«[...] мы должны показать, что наличное физическое информационное пространство относится к числу тех пространств, к которым применим двухаспектный принцип. Для этого нам потребовалась бы более детализированная версия этого принципа, надлежащим образом ограничивающая задействуемые информационные пространства, так, чтобы он хотя бы включал информацию, доступную для глобального контроля в привычным нам случаях»²⁷².

Чалмерсу кажется интуитивно понятным, что с феноменальными переживаниями могут быть связаны далеко не все информационные структуры, получившие физическую реализацию, а только некоторые из них, обладающие набором конкретных качеств. Так, приводя свой знаменитый пример с термостатом, с функционированием которого, как полагает философ, связано некоторое каноническое информационное пространство, Чалмерс утверждает, что другой объект, как, например, камень, вряд ли имеет какие-либо ментальные переживания²⁷³. Философ задается вопросом о том, требуется ли для появления сознания «деятельная» причинность:

«Каузальность, фундирующая опыт в мозге, кажется по большей части деятельной, так как релевантная информация постоянно обрабатывается, нейроны посылают импульсы и т.д. С другой

²⁷² Чалмерс Д. Сознательный ум. В поисках фундаментальной теории. Пер. В.В. Васильева. М.: УРСС: Книжный дом «ЛИБРИКОМ», 2013. С. 359.

²⁷³ Подобные рассуждения Чалмерса легко вписываются в более широкий контекст дискуссий о выборе между уже упомянутыми в данной работе «атрибутивным» и «функциональным» подходами к толкованию информации: можно провести некоторые параллели между рассуждениями сторонников «неограниченного» подхода и доводами сторонников атрибутивного подхода, равно как и параллели между «ограниченными» версиями двухаспектизма, которые будут рассмотрены далее, и функционально-ориентированными интерпретациями информационных пространств. Данное замечание иллюстрирует тезис Д.И. Дубровского о том, что основные спорные вопросы, касающиеся интерпретации информации, остаются в силе на протяжении уже более пятидесяти лет (Бодякин В.И., Дубровский Д.И., Колин К.К., Лекторский В.А., Мелик-Гайказян И.В., Пружинин Б.И., Урсул А.Д. Информационный подход в междисциплинарной перспективе. (материалы «круглого стола») // Вопросы философии. 2010. №2. С. 84-113).

стороны, не исключено, что различение деятельной и пассивной каузальности не может быть проведено на фундаментальном уровне, и в таком случае их можно было бы рассматривать как равноценные. Я не знаю ответа на этот вопрос, но интуитивно кажется, что для опыта требуется какая-то деятельность»²⁷⁴.

Кроме того, философ отмечает, что большинство доступных индивиду процессов преобразования информационных сигналов протекает неосознанно; в рамках неограниченного двухаспектного принципа это можно было бы объяснить, сославшись на реализацию бессознательной информации в качественном поле иных мозговых подсистем («Я не мог бы ожидать прямой доступности для меня переживаний других систем, так же как я не жду прямой доступности для меня переживаний других людей»²⁷⁵, – пишет Чалмерс). Однако возможность подобной пролиферации опыта в пределах одного мозга не вполне обнадеживает философа.

По этой причине Чалмерс не преминул рассмотреть альтернативные варианты развития двухаспектного принципа, связанные с ограничением класса информационных структур, имеющих некоторые феноменальные корреляты, и указал на наиболее приемлемые критерии ограничения данного принципа. Он упоминает, во-первых, критерий информационной обработки (хотя, как отмечает Чалмерс, само выражение «информационная обработка» слишком неопределенное, в силу чего оно вряд ли может помочь в выявлении конкретного фундаментального психофизического закона). Во-вторых, это мог бы быть критерий усиления (amplification) информации, подразумевающий, что информация получает феноменальную реализацию только в том случае, если она оказывает значительное влияние на каузальную топологию некоторой физической системы (понятие «усиление» также кажется философу размытым, так как существует немало примеров систем, способных к амплификации внутренней информации, но вряд ли обладающих сознанием). Наконец,

²⁷⁴ Чалмерс Д. Сознание и ум. В поисках фундаментальной теории. Пер. В.В. Васильева. М.: УРСС: Книжный дом «ЛИБРИКОМ», 2013. С. 359.

²⁷⁵ Там же. С. 374.

возможно наложить ограничение на тип каузальных отношений, реализованных в системе. И хотя сам философ не уверен в выборе релевантного ограничительного критерия, он не отрицает вероятности его существования.

В современной науке и философии существует достаточно много стратегий, призванных так или иначе ограничить область информационных состояний, с которыми может быть ассоциирован феноменальный опыт: среди них присутствуют как концепции, подразумевающие некоторый тип «усложнения» информационных состояний, так и теории, пересматривающие основания каузальных отношений, с которыми коррелируют те или иные информационные состояния. Рассмотрим их подробнее.

4.4.1. Феноменальный опыт как результат преобразования информационных структур

Как нетрудно заметить, большинство аналитических философских теорий, описывающих природу феноменального сознания в информационных терминах, подразумевают, что участвующая в подобном процессе информация должна быть необходимым образом «преобразована» или «обработана»: потребность в постулировании подобного рода критерия появления сознания напрямую связана с необходимостью избежать выводов о вычислительной тривиальности и ограничить круг субъектов опыта, испытывающих полноценные сознательные переживания. Вспомним, например, какую роль отводят туманному выражению «обработка информации» Дретске, Сейр, Дубровский и многие другие рассмотренные выше философы. Вера во всесильность «информационной обработки» присуща и многим представителям естественнонаучных дисциплин; так, большинство современных теорий, описывающих различные аспекты информационных процессов, часто упоминают те или иные операции, связанные с обработкой информации, ее амплификацией, интеграцией и т.д.

Так, одной из наиболее популярных сегодня нейробиологических теорий, призванных объяснить возникновение феноменальных переживаний, является Теория интегрированной информации итальянского ученого Джулио Тонони. Согласно данной теории, сложная феноменальная структура сознания обусловлена не только его информативностью, обеспечивающей индивидуальность каждого отдельного сознательного переживания, но и его интегрированностью и исключительностью²⁷⁶. Интеграция определяется несводимостью сознательного опыта к простой сумме его компонентов (так, согласно Тонони, переживание слова «SONO», не редуцируемо к переживанию его частей «SO» и «NO» по отдельности, как и переживание зеленого круга не редуцируемо к отдельным переживаниям геометрической фигуры и цвета). Исключительность же обуславливается невозможностью возникновения различных одновременных сознательных состояний: в каждый конкретный момент времени имеет место лишь одно феноменальное переживание, а не суперпозиция множества «полупереживаний»; каждый качественный опыт имеет определенные границы (что-то поддается осознанию, что-то - нет) и конкретные пространственные и темпоральные свойства^{277, 278}.

Опираясь на данные аксиомы, Тонони и его исследовательская группа заявляют о совпадении феноменальных состояний и интегрированной информации в рамках целостной физической системы. Из этого предположения следует, что степень «сознательности» может быть дифференцированной, ведь различные виды системной организации влияют на уровень информационной интеграции; кроме того, Тонони полагает, что даже простейшие системы вполне

²⁷⁶ Стоит отметить, что о связи интеграции информационных состояний и сознания говорят многие современные ученые, среди которых Бернард Баарс (Baars B. J. A Cognitive Theory of Consciousness. NY: Cambridge University Press, 1988), Томас Метцингер (Metzinger T. Being No One: The Self-Model Theory of Subjectivity. Cambridge, MA, MIT Press, 2003), Дэвид Гамез (Gamez D. Information Integration Based Predictions About the Conscious States of a Spiking Neural Network. Consciousness and Cognition, 2010. №19(1). P. 294-310) и др. Однако далеко не все эти ученые утверждают, что сознание – это и есть интегрированная информация.

²⁷⁷ Oizumi M., Albantakis L., Tononi G. From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0. // PLoS Computational Biology. 2014, №10. P. 23-57.

²⁷⁸ Более подробно основные положения теории интегрированной информации описаны в статье: Танюшина А.А. Философские проблемы теории интегрированной информации Д. Тонони // Вестник Московского государственного университета. Серия 7. Философия. 2020. №1. С. 13-31.

могут обладать минимальным уровнем сознания. Таким образом, если рассматривать описанные нейробиологом принципы в качестве некоторого фундаментального закона природы, то выходит, что макроопыт эмерджентен на физических структурах, отчего теория интегрированной информации принимает вид эмерджентного панпсихизма. При этом Тонони о своей концепции пишет следующее:

«Как относится моя теория с панпсихизмом, согласно которому всё во Вселенной обладает некоторым родом сознания? Разумеется, теория интегрированной информации подразумевает, что многие сущности, состоящие из особых функциональных механизмов и делающие выбор между данными им альтернативами, обладают некоторым уровнем сознания. Однако, в отличие от традиционного панпсихизма, теории интегрированной информации не приписывает сознание всем объектам без разбору»²⁷⁹.

Стоит отметить, что далеко не все исследователи считают, что Тонони выбрал верный ограничительный принцип, ведь предложенная им теория не может ответить на вопрос «Почему именно вследствие информационной интеграции система может испытывать феноменальное переживание?». Ряд ученых предлагает иные способы подсчета меры интеграции информации в сложных системах²⁸⁰, в то время как другие выдвигают альтернативные концепции, которые могут похвастаться такой же предсказательной силой, как и теория Тонони, но при этом используют иные ограничительные критерии для определения информационных пространств, с которыми может быть связано сознание. К примеру, итальянский нейробиолог Микаэль Черулло описывает условную теорию (так называемая Теория скоординированных циркулярных сообщений – Circular Coordinated Message Theory), в рамках которой появление

²⁷⁹ Tononi G. Consciousness as Integrated Information: A provisional manifesto // Biological Bulletin. 2008, №7. P. 236.

²⁸⁰ См. например: Alexander I., Gamez D. Informational Theories of Consciousness: A Review and Extension // Advances in Experimental Medicine and Biology. 2011. №718. P. 139-147.

ментального опыта обеспечивается рекурсивной динамикой информации по петлям обратной связи²⁸¹, а японские исследователи Риота Канай и Йен Ю представили так называемую Теорию нетривиальной информационной замкнутости (Non-Trivial Informational Closure), чьи положения во многом схожи с положениями теории Тонони, хоть и основываются на других математических формализмах²⁸².

Еще одним возможным критерием может быть уровень амплификации, или усиления информации. Его предлагает, в частности, бельгийская исследовательница Лиан Габора, согласно концепции которой, в целях развития двуаспектного принципа информации можно предположить, что усиление информации в рамках единой системы может способствовать переходу ее элементов из «потенциально феноменальных» состояний в «актуально феноменальные». Согласно Габора, это может происходить в два этапа: во-первых, информация должна быть проинтегрирована, будучи заключенной в замкнутой рекурсивной петле; во-вторых, данная интегрированная информация должна быть подвергнута соответствующей динамической дивергенции и конвергенции. Усиление информации происходит за счет замкнутости самой системы: по мере того как отдельные ее элементы превращаются в замкнутую структуру, они порождают не просто новые процессы обработки информации, но именно те процессы, которые могут быть использованы уже существующей системой. Такое разделение на релевантную и нерелевантную информацию обуславливает, по мнению исследовательницы, исчезновение микроопытов из пространства феноменального поля макросубъекта²⁸³.

Проблема всех теорий, подразумевающих, что для появления феноменальных качеств необходимо своеобразное «усложнение» информационных структур той или иной системы, заключается в том, что

²⁸¹ Cerullo M. Integrated Information theory: A promising but ultimately incomplete theory of consciousness // Journal of Consciousness Studies. 2011, №18. P. 45-58.

²⁸² Biehl M., Chang A.Y.C, Kanai R., Yu Y. Information Closure Theory of Consciousness // Front Psychol. 2020. №11. P. 1504-1522.

²⁸³ Gabora L. Amplifying phenomenal information: Toward a fundamental theory of consciousness. // Journal of Consciousness Studies, 2002. №9(8). P. 3-29.

ограниченный подобным образом двааспектный принцип информации будет предполагать существование сильных эмерджентных свойств, обеспечивающих появление информационных пространств, с которыми связан феноменальный опыт. Сами по себе эмерджентные формы панпсихизма имеют некоторое преимущество перед конститутивными формами, так как они не сталкиваются с описанной выше проблемой комбинации: эмерджентные феноменальные свойства можно трактовать как фундаментальные, а не конституированные более простыми прото-свойствами. Однако по этой причине данные свойства не наследуют каузально-релевантные свойства прото-свойств, и сталкиваются либо с угрозой эпифеноменализма, либо с необходимостью постулирования каузальных сил особого рода, выходящих за рамки принципа каузальной замкнутости физического, существование которых еще не доказано современной наукой.

4.4.2. Феноменальный опыт, ассоциированный с особой каузальной деятельностью

Помимо самого Чалмерса идею о том, что для появления высокоуровневого сознания необходима некая активная «деятельная» причинность, допускают многие философы. Такой подход к объяснению ментального опыта, как нам уже известно, использует большинство сторонников информационных теорий сознания, склоняющихся к функциональной трактовке информационных пространств, а не атрибутивной. Однако уточнение понимания каузальных коррелятов информационных состояний в рамках непосредственно двааспектной теории было предложено относительно небольшим количеством философов. Один из проектов развития прототеории Чалмерса в данном направлении принадлежит австралийскому философу Риза Мали,

предложившему использовать исключительно прагматический способ толкования информационных пространств в рамках двуаспектного принципа²⁸⁴.

Мали полагает, что используемый многими философами формально-синтаксический способ толкования информации, имеющий дело с математической мерой количества информации, не вполне оправдан: его отношение к непосредственно информационным состояниям можно сравнить с отношениями между «массой» и «материей»: в то время как понятие «масса» является всего лишь научной абстракцией, материя реально представлена в мире. И даже когда Чалмерс вместо сугубо синтаксического толкования информации использует определение, данное Бейтсоном, Мали задается следующим вопросом: если понимание физически реализованной информации не идет дальше каузальной интерпретации, формализованной с помощью выражения «различие, которое создает различие», то почему мы вообще должны использовать понятие информации? Можно легко заменить двуаспектный принцип информации, на «двуаспектный принцип каузальности», согласно которому каждое каузальное действие реализует физические и феноменальные свойства, и данная замена не лишит теорию ее объяснительного потенциала. С точки зрения философа, описываемое Бейтсоном различие должно быть значимым, а значимость определяется целенаправленностью и прагматикой некоторой системы (вспомним схожие рассуждения Д.И. Дубровского о ложности атрибутивных подходов, в рамках которых понятие «информация»

²⁸⁴ Философ ссылается на концепцию прагматической информации физика Хуана Роедерера, согласно которой в природе существуют два базовых типа каузальных взаимодействий: силовые и информационные. Ключевое различие между данными типами взаимодействий состоит в том, что первый сводится к энергетическому обмену между основными взаимодействующими элементами, в то время как при информационном взаимодействии энергия, направленная на информационную трансформацию получателя информации, не передается непосредственно самим отправителем, а приходит от некоторого иного, внешнего источника; по этой причине все информационные системы являются, по Роедереру, открытыми системами. Кроме того, процесс информационного взаимодействия должен быть способен повторяться, что, в свою очередь, требует от системы-получателя наличия своеобразной памяти, призванной обеспечить соответствие ее организационных паттернов изменениям. Иными словами, однозначное соответствие изменения паттерна обеспечивается после установления некоторого информационного кода между отправителем и получателем. В случае отсутствия подобного кода взаимодействие не может считаться информационным (Maleeh R. Pragmatic Information as a Unifying Biological Concept // *Information*, 2014. №5. P. 451-478).

сближается с физическими категориями, вследствие чего его трактовка сильно размывается).

Впрочем, прагматическую информацию философ также трактует в терминах каузальности, отмечая при этом, что мы должны отличать каузальность в биологических системах от каузальности в искусственных системах. Подлинные информационные взаимодействия происходят, с точки зрения Мали, исключительно в живых системах: именно для живых организмов информация играет активную роль, так как информационные взаимодействия, в которых форма или паттерн являются управляющими факторами, необходимы для поддержания продолжительного неустойчивого термодинамического равновесия с окружающей средой (философ говорит в первую очередь о биомолекулярных информационных процессах, управляющих метаболизмом, ростом, размножением и дифференцированием клеток, а также о нейронных информационных процессах, управляющих поведением и интеллектом). Таким образом, Мали выступает против утверждения Чалмерса, согласно которому информация существует везде, где представлены каузальные отношения; с точки зрения философа, каузально-обусловленная информация обнаруживается лишь в живых системах, в то время как физическая каузация подчиняется исключительно правилам силовых взаимодействий.

Очевидно, что ограничивая область применения двуаспектного принципа лишь областью прагматической каузации, Мали несколько противоречит первоначальному предположению Чалмерса о возможном движении вниз по шкале информационной и ментальной сложности, ведь, трактуя информацию прагматически, мы непременно сталкиваемся с необходимостью определения естественной границы между силовыми и информационными отношениями (то есть между «живыми» и «неживыми» системами). Чалмерс, однако, отмечал, что, каким бы не был релевантный ограничительный критерий для его двуаспектной прототеории, он все же может быть совместим с утверждением о наличии информации и связанных с ней феноменальных переживаний у простых систем вроде термостата. Кроме того, даже если мы ограничиваем область

субъектов феноменального опыта живыми организмами, перед теорией все равно стоит своеобразная «проблема комбинации»: вопрос о том, как биологические клетки, являющиеся, по Мали, информационными агентами, образуют макросубъекта опыта наподобие человека, является все еще открытым.

Наконец, стоит также отметить, что подход Мали не может дать ответа на самый трудный вопрос: почему именно прагматическая информация необходимым образом должна быть реализована как физически, так и феноменально? Очевидно, что концепция философа сводится к постулированию определенного психофизического закона, связывающего между собой эти свойства в рамках обозначенного Чалмерсом принципа, однако выдвигаемый им критерий может быть легко оспорен. Впрочем, и сам Мали не до конца уверен в верности выбранного им «прагматического» критерия, хотя и полагает, что его применение намного выгоднее для двуаспектной теории, чем принятие ее неограниченной версии.

Некоторые философы предлагают иные, более радикальные способы трактовки каузальных пространств, с которыми может быть ассоциирован информационно-обусловленный феноменальный опыт. Так, если Мали находит необходимым всего лишь выделить из всех известных нам каузальных отношений те, которые следует ассоциировать с реализацией информации, то ряд философов предлагает пересмотреть саму структуру каузальных отношений.

В качестве примера можно рассмотреть подход американского философа Грегга Розенберга, представленный в его книге 2004 года «Место для сознания»²⁸⁵. Розенберг предполагает, что наиболее естественным ограничительным критерием, служащим для индивидуации субъектов опыта, является именно такая фундаментальная категория, как каузальность. Данное предположение, однако, разрывается между постулированием повсеместного распространения каузальных отношений и, как следствие, существованием единого субъекта вселенского масштаба, и признанием неограниченного

²⁸⁵ Rosenberg G.H. A Place for Consciousness: Probing the Deep Structure of the Natural World. Oxford University Press, 2004.

количества микросубъектов опыта, вопрос о способе образования которыми макросубъектов остается открытым. Розенберг предлагает проплыть между обозначенными «Сциллой и Харибдой панпсихизма» с помощью создания новой теории каузальных взаимодействий, которую он именуется «теорией каузальной значимости» («theory of causal significance»).

Философ предлагает описывать каузальную действенность как совокупность ограничений, которые некоторое событие накладывает на то, каким может быть мир, а не на эффектах, которые оно производит в мире; эта система ограничений подразумевает, что действительный мир возникает из пространства возможностей.

С помощью предлагаемой философом трактовки каузального действия нетрудно обосновать возможность существования информационной структуры, которую мы приписываем миру: трактуя информацию как уменьшение неопределенности, Розенберг отмечает, что только информационно-релевантные каузальные свойства индивида низшего уровня (микросубъекта) представляют данного индивида внутри индивида более высокого уровня (макросубъекта). По этой причине, индивиды, образованные несколькими низкоуровневыми индивидами, обладают лишь ограниченной информацией о совокупности образующих их каузальных отношений, ведь элементы некоторой системы всегда теряют количество степеней своей свободы, вступая в некоторые структурные отношения, определяющие организацию данной системы. Таким образом, можно говорить, что: а) феноменальный опыт индивида должен соответствовать совокупности каузальных свойств, представленных, в свою очередь, подобной информацией; б) данная информация может быть представлена также объективно (как уменьшение неопределенности возможных диспозиций каузальных отношений). Отсюда следует, что в рамках теории Розенберга выполняется своеобразный двуаспектный принцип информации. При этом философ идет дальше и утверждает, что общее количество опыта у индивидов более высокого уровня совпадает с количеством интегрированной информации о каузальных свойствах микроиндивидов; подобная трактовка явно

сближает концепцию Розенберга с теорией интегрированной информации Тонони, что, однако, замечает и сам философ²⁸⁶.

Итак, теория Розенберга представляет собой своеобразную версию эмерджентного панпсихизма: философ напрямую указывает на то, что каузальные свойства, принадлежащие индивидам на разных уровнях организации, являются сильными эмерджентными свойствами. Предложенная философом теория каузальной значимости представляет мир в качестве «наслоенных» друг на друга индивидов разного порядка, где детерминированные макроиндивиды могут существовать «поверх» микроиндивидов, находящихся в неопределенных состояниях, делая последних, в свою очередь, детерминированными. Такая трактовка нисходящей каузации, как утверждает сам Розенберг, позволяет его теории избежать присущей большинству эмерджентных теорий сознания проблемы ментальной каузальности: отмечая, что каузальное действие обеспечивает не только взаимодействия между индивидами одного уровня, но также и между индивидами разных уровней, философ указывает на способность представителей более высокого уровня каузально влиять на представителей низших уровней, детерминируя их потенциально возможные состояния²⁸⁷. В то же время, некоторые философы полагают, что такой вариант трактовки каузальных отношений предполагает скорее слабую эмерджентность, так как состояния индивидов разных уровней являются лишь разными аспектами единого каузального процесса и зависят исключительно от точки зрения, с которой мы его изучаем²⁸⁸. Вследствие этого теория Розенберга представляется неполной.

²⁸⁶ Rosenberg G.H. Causality and the combination problem // *Consciousness in the Physical World: Perspectives on Russellian Monism* (T. Alter, Y. Nagasawa, eds.). Oxford University Press, 2014. P. 244.

²⁸⁷ Сторонником подобной «многоуровневой» формы эмерджентного панпсихизма с нисходящей каузацией является также современный немецкий философ Годехард Брюнтруп (Brüntrup G. *Emergent Panpsychism // Panpsychism* (Brüntrup G., Jaskolla L. eds.). New York: Oxford University Press, 2016. P. 48–71).

²⁸⁸ Mørch H. H. *Panpsychism and Causation: A New Argument and a Solution to the Combination Problem*. Ph.D. thesis, University of Oslo, 2014. [Электронный ресурс] // Сайт philpapers.org. URL: <https://philpapers.org/archive/HASPAC-2.pdf>. (дата обращения: 26.04.2020).

Рассмотрев несколько наиболее показательных стратегий ограничения двухаспектного принципа информации, можно сделать вывод, что все они так или иначе сталкиваются с необходимостью постулирования эмерджентных системных свойств, полученных либо в результате преобразования и обработки соответствующих информационных структур, либо в результате появления особых видов каузальных отношений, ассоциированных с информационными пространствами. Разумеется, нельзя утверждать, что в силу описанной выше проблемы каузальной релевантности феноменальных свойств ограниченное развитие двухаспектного принципа является определенно неверным; однако, так как именно проблема ментальной каузальности во многом обусловила обращение многих современных философов к панпсихизму, возвращение к ней представляется большим шагом назад. Как отметил философ Вильям Сигер: *«Если панпсихизм [...] подразумевает эмерджентность, то почему бы не пойти теоретически более экономным путем, и допустить, что сознание возникает непосредственно из самой физической основы»*²⁸⁹. Действительно, с таким же успехом можно вернуться к положениям нередуктивного физикализма, которые представляются более правдоподобными (по крайней мере, научному сообществу).

Тем не менее, столкновение теории с проблемой комбинации или с необходимостью признания эмерджентизма еще не означает, что проект развития двухаспектного принципа информации является бесперспективным, и некоторые философы предлагают взглянуть на выдвигаемую Чалмерсом прототеорию под другим углом.

4.5. «Аспектизм» и дефляционизм по отношению к осведомленности

Как уже было отмечено выше, двухаспектная онтология с самого начала импонировала Чалмерсу, так как философ предполагал, что она является

²⁸⁹ Seager W. The emergence of consciousness // *Philosophic Exchange*, 2006. №36(1). P. 20.

достаточно простой с точки зрения накладываемой ею онтологических обязательств: она постулировала лишь одну фундаментальную реальность, которая обладает двумя аспектами – физическим и феноменальным. При первоначальной трактовке двуаспектного принципа Чалмерс говорил о том, что паттерн и информация – это лишь возможные точки зрения на подобную реальность, взгляды «снаружи» и «изнутри». Однако, подобная двуаспектная теория позволяет объяснить лишь наличие корреляции между определенными наблюдаемыми физическими свойствами и субъективными переживаниями, но не отвечает на вопрос о том, почему эти переживания в принципе существуют (как отмечает философ В.В. Васильев, именно неспособность двуаспектной теории ответить на «глубочайший вопрос» о том, почему вообще существуют квалиа, способствовала дальнейшему смещению акцентов в теории австралийского философа²⁹⁰). В своих последующих работах Чалмерс формулирует двуаспектный тезис иначе: так, в «Сознающем уме» он утверждает, что информация присутствует везде и может получить как физическую (объективную), так и феноменальную (субъективную) реализацию; вкупе с идеями неорасселовского монизма, подобная трактовка представляла ментальные сущности как необходимые для онтологии физической реальности. Постулирование отношений реализации, как отношений между свойствами разных порядков, превратило феноменальные «аспекты» в конститутивные свойства, что позволило избежать проблемы ментальной каузальности²⁹¹, но нивелировало остальные метафизические преимущества двуаспектного подхода, которые изначально отмечал философ. Вероятно, что именно допущение трактовки аспектов как свойств затрудняет дальнейшее развитие предложенной Чалмерсом прототеории, неограниченной версии которой угрожает проблема

²⁹⁰ Васильев В.В. Трудная проблема сознания. М.: Прогресс-традиция. 2009. С. 158.

²⁹¹ Стоит отметить, что некоторые философы считают, что конститутивный панпсихизм не решает проблему ментальной каузальности, так как в его рамках фактически утверждается, что нефеноменальные по определению протофеноменальные свойства конституируют феноменальные, в то время как сам Чалмерс с помощью аргумента представимости зомби показал, что подобное логическое следование невозможно (Кузнецов А.В. Аргумент против конститутивного панпсихизма // Вестник Томского государственного университета. Философия. Социология. Политология, № 4, 2016. С. 60-68).

комбинации, в то время как ограниченные версии сталкиваются с необходимостью постулирования сильных эмерджентных свойств²⁹².

Разумеется, теория, в рамках которой феноменальные состояния трактуются не как свойства, а как аспекты, тоже сталкивается с проблемой ментальной каузальности: можно сказать, что каузальной силой в рамках двуаспектной теории обладает лишь онтологически фундаментальная сущность, а физические и феноменальные явления, соответствующие ее аспектам, не выступают в роли ни полноценных каузальных агентов, ни «номологических бездельников». Связь аспектов в данном случае не обусловлена какими-либо естественными или психофизическими законами и относится к категории метафизической модальности: она является следствием того факта, что аспекты чего-либо являются метафизически взаимозависимыми по определению.

Исследователей, как правило, волнует проблема, связанная с интерпретацией онтологического статуса такой категории, как аспект, ведь из определения не ясно, что он представляет собой: некоторую сущность *sui generis* или же сущность, зависящую от условного наблюдателя. Некоторым философам кажется очевидным, что отказ от реалистской позиции по отношению к аспектам чреват весьма противоречивыми выводами, так как предполагает, что существование ментальных явлений возможно только при условии наличия сознательного наблюдателя, чье сознание, в свою очередь, обусловлено существованием этих ментальных явлений. Подобный замкнутый круг лишает двуаспектную онтологию всех ее теоретических преимуществ. Согласно реалистскому подходу, ментальные феномены существуют вне зависимости от сознательного наблюдателя, и данное утверждение вновь приводит нас к своеобразному панпсихизму. Так как оба подхода к толкованию аспектов имеют концептуальные трудности, некоторые философы предлагают трактовать

²⁹² Возникновение подобных проблем, связанных с дальнейшим развитием двуаспектного принципа не удивительно: еще Джон Серл в своей статье с говорящим названием «Почему я не дуалист свойств?» продемонстрировал, что все теории, постулирующие существование двух видов взаимно нередуцируемых свойств, неизбежно сталкиваются с проблемой «феноменальной связки» (созвучной с проблемой комбинации) и проблемой ментальной каузальности (Searle J.R. Why I Am Not a Property Dualist // Journal of Consciousness Studies, №9, 2002. P. 57-64).

сознательный опыт как комбинацию двух характеристик: непосредственно качественного переживания и осведомленности (субъективности), которая, в отличие от феноменального опыта, не является метафизически фундаментальной и скорее представляет собой организационное качество. К примеру, современный швейцарский аналитический философ Батист Ле Биан утверждает, что данное концептуальное деление помогает лучше понять, как возможно независимое существование данных явлений, и объяснить, почему лишь сложные системы, обладающие определенным уровнем физической организации, испытывают сознательные переживания²⁹³. Нетрудно заметить, что подобный подход во многом совпадает со стратегией признания дефляционизма по отношению к осведомленности, о которой было упомянуто в ходе рассуждения о возможных решениях проблемы комбинации (к которой склоняется и Чалмерс, которому кажется очевидным, что индивид, осведомленный о собственных феноменальных переживаниях, является субъектом опыта, и поэтому, если какой-то подход может показать, как мозг или организмы являются осведомленными о своих феноменальных состояниях, то он покажет, как мозг или организмы могут быть субъектами опыта).

Одним из сторонников такого подхода является американский философ Сэм Колман, утверждающий, что квалиа могут существовать независимо от воспринимающих их субъектов; в противном случае, сторонникам панпсихизма вряд ли удастся найти решение проблемы комбинации макросубъекта из микросубъектов. Философ утверждает, что так как понятие субъективности должно непременно предполагать некоторую феноменологическую точку зрения (о чем упоминал еще Томас Нагель в своей хрестоматийной статье «Каково это быть летучей мышью»), то под субъектом опыта мы должны подразумевать некую воспринимающую сущность. Колман предлагает следующий мысленный эксперимент, демонстрирующий взаимосвязь между точкой зрения субъекта и теми феноменальными качествами, которые входят в его качественное поле:

²⁹³ Le Bihan B. Aspects in Dual-Aspect Monism and Panpsychism: A Rejoinder to Benovsky // *Philosophical Investigations*, 2019. №42(2). P.186-201.

представим большое количество качественно одинаковых субъектов на «стартовой линии» существования, каждый из которых с течением времени проходит свой уникальный путь в окружающем мире, что, в свою очередь, приводит к различным модификациям качественного поля каждого. Таким образом, каждая точка зрения субъекта имеет доступ к качественно отличному массиву качества, по сравнению с массивами других субъектов, и именно факт существования подобных различных «взглядов» на мир и определяет характер особого набора качеств, которые каждый субъект испытывает в данный момент времени²⁹⁴.

Проблема конститутивного панпсихизма, полагает Колман, заключается в том, что точки зрения не подлежат комбинированию: в любой комбинации «ингредиенты» могут преобразовываться, но при этом они сохраняются; однако достаточно трудно представить себе, как все первоначальные микросубъекты опыта каким-то образом добавляют свою точку зрения получившемуся макросубъекту. Как полагает философ, у субъекта либо существует единая точка зрения, либо ее не существует вовсе; комбинированию могут подлежать лишь феноменальные качества, но не перспективы на них. Отсюда следует, что панпсихист, настаивающий на существовании микросубъектов опыта, должен принять позицию «инфузионизма», согласно которой микросубъекты «растворяются» в макросубъекте; но если микросубъекты «сливаются», то получившийся макросубъект уже не является с ними структурно-связанным. Таким образом, результирующий макросубъект может быть в лучшем случае лишь каузально связан с образующими его микросубъектами. Следовательно, макросубъект эмерджентен на микросубъектах²⁹⁵.

Итак, выходит, что основная идея конститутивного панпсихизма о существовании микросубъектов опыта противоречит идее о возможности

²⁹⁴ Coleman S. The Real Combination Problem: Panpsychism, Micro-Subjects, and Emergence // Erkenntnis, 2013. №1. P.16.

²⁹⁵ Безусловно, это признают и сами сторонники «инфузионизма»: например, философ Пол Хамфрис, предложивший идею «слияния» микросубъектов, не отрицает, что его теория скорее подразумевает некомбинаторное решение проблемы возникновения макросубъектов (Humphreys P. How properties emerge // Philosophy of Science. 1997. № 64 (1). P. 1-17).

структурного решения вопроса об их комбинации в макросубъекта; следовательно, полагает Колман, следует отказаться от какого-то из данных положений. Философ приходит к выводу о необходимости концептуального разграничения феноменальных и субъективных качеств, где под последними он подразумевает осведомленность о первых. Так как осведомленность не подлежит комбинации, единственный способ избежать постулирования сильной эмерджентности сводится к принятию нейтрального монизма, в рамках которого осведомленность является реляционным свойством. В нейтральном монизме качественные качества являются неизменным атрибутом фундаментальной реальности, некоторые части которой проявляют конфигурацию, обеспечивающую осведомленность об этих качествах, – согласно Колману, так и появляется сознание. Как отмечает философ, мозг – это фактически вид «самонаводящегося микроскопа», раскрывающего свой собственный внутренний характер.

Предложенный философом подход, предполагающий трактовку сознания непосредственно как точку зрения (или перспективу) на некоторую нейтральную фундаментальную реальность, уводит нас в некотором роде назад от панпсихизма к физикализму (сам Колман отмечает, что не будет ничего удивительного, если физикалисты найдут его теорию значительно менее отталкивающей, чем, например, конститутивный панпсихизм). Трудность заключается лишь в том, чтобы убедить исследователей в возможности независимого существования квалиа от воспринимающих их «субъектов». Это концептуальная возможность, является, как полагает философ, вполне реальной, так как обратное утверждение основывается исключительно на наших интуициях и не имеет никаких подтверждений (в конце концов, считает философ, если квалиа боли могут разбудить нас ночью, заставив себя осознать, то вероятно, что квалиа существовала до того, как человек осознал их). Таким образом, подход Колмана, являющийся совместимым с двуаспектным принципом, в рамках которого под аспектом понимается именно «точка зрения»

или «перспектива», не сильно противоречит традиционным функциональным подходам к объяснению сознания.

Безусловно, данная концепция сталкивается с некоторыми трудностями. К примеру, остается не вполне понятно, откуда появляется «субъективность» или осведомленность: очевидно, что ее вряд ли следует редуцировать к некоторым сугубо организационным качествам сложноорганизованной физической системы (в данном случае, согласно аргументу представимости, можно легко представить «панпсихистского зомби», обладающего феноменальными качествами, но не осознающего их). Более того, представляется, что подобного рода «аспектизм» сталкивается с насущной проблемой «картезианского театра», вследствие чего, как утверждает один из критиков Колмана, единственным возможным способом объяснить субъективность является принятие ее иллюзорности²⁹⁶; однако в данном случае теории будет выгоднее отречься от постулирования реального существования феноменальных качеств и принять вид простого элиминативного материализма иллюзионистского типа²⁹⁷. Таким образом, вопрос о том, как следует толковать природу субъективности и осведомленности является открытым.

Еще один возможный вариант интерпретации двуаспектного принципа предлагает английский философ Макс Велманс, автор нашумевшей в начале

²⁹⁶ Blamauer M. Panpsychism without Subjectivity? A Brief Commentary on Sam Coleman's «Mental Chemistry» and «The Real Combination Problem» // *Disputatio*, 2013. № 5 (37). P. 299-309.

²⁹⁷ Отказ от постулирования особых феноменальных качеств ментального опыта может иметь и другие теоретические следствия. Любопытные рассуждения на этот счет можно найти в работах российского философа Д.В. Иванова: опровергая такие известные аргументы в пользу существования функционально нередуцируемых феноменальных свойств, как аргумент знания и аргументы отсутствующих и инвертированных качеств, философ предлагает обратиться к интенционалистской интерпретации качественных переживаний. Ссылаясь на доводы Серла, Иванов отмечает, что основной чертой интенциональности является ее «аспектуальность»: «Существенной чертой интенциональных состояний является их аспектуальность. Она объясняет перспективный характер нашего опыта, то есть тот факт, что субъективные феномены всегда даны нам определенным образом [...] Сказать, что интенциональное состояние обладает аспектуальной формой, фактически означает указать на то, что это состояние обладает определенным содержанием, фиксирующим объект, на который оно направлено, в определенных аспектах» (Иванов Д.В. Природа феноменального сознания. М.: Книжный дом «ЛИБРОКОМ». 2013. С. 204-205). Чтобы построить натуралистическую концепцию интенциональности, философ предлагает обратиться к экстерналистской интерпретации ментальных содержаний. Решить проблему субъективности (возникающую, с точки зрения Иванова, в связи с критикой интенционализма сторонниками радикального энактивизма), предлагается посредством опоры на теорию концептуального содержания Джона Макдауэла и идеи сенсорно-моторного энактивизма Альва Ноэ (Иванов Д.В. Радикальный энактивизм и проблема субъективности // *Вопросы философии*. 2016. № 11. С. 60–69; Иванов Д.В. Проблема субъективности и прямой реализм Дж. Макдауэла // *Вопросы философии*. 2017. № 12. С. 121–130).

1990-х годов статьи «Является ли обработка информации человеком сознательной?»²⁹⁸, представляющей собой последовательный анализ последних экспериментальных данных о связи сознания с информационными процессами, происходящими в мозге. Оживленная дискуссия вокруг данной публикации побудила философа к написанию дополнительного текста, в котором он описывает связь феноменальных переживаний с их информационными коррелятами в терминах двухаспектного принципа информации (заметим, что основные идеи данного принципа философ сформулировал без опоры на тексты Чалмерса, ведь статья Вэлманса вышла на несколько лет раньше ключевых публикаций австралийского философа), отмечая при этом, что взгляды от первого и третьего лица на сознание можно рассматривать как взаимодополняющие и взаимно не редуцируемые²⁹⁹. Именно эти идеи позже лягут в основу его концепции рефлексивного монизма, в рамках которой феноменальные и физические аспекты сознания трактуются как комплементарные.

Философ приводит аналогию с принципом комплементарности волн и частиц в квантовой механике: в физике ни волновая, ни корпускулярная природа частиц не рассматривается как более фундаментальная³⁰⁰. Велманс отмечает, что воспринимаемый нами «физический мир» тоже является частью сознательного опыта, и это требует иного понимания того, как переживаемые нами явления соотносятся не только с коррелирующими с ними процессами в мозге, но и с окружающей средой. Рефлексивный монизм предлагает не отделять сознание от

²⁹⁸ Velmans M. Is human information processing conscious? // Behavioral and Brain Sciences. 1991. №14(4). P. 651-669.

²⁹⁹ Velmans M. Consciousness from a first-person perspective // Behavioral and Brain Sciences. 1991. №14 (4). P.702-726.

³⁰⁰ Velmans M. Reflexive monism // Journal of Consciousness Studies, 2008. № 15(2). P. 31-32.

Подобную параллель между корпускулярно-волновой природой частиц в физике и психофизической природой фундаментальной реальности проводит и немецкий философ Гаральд Атманшпрахер. Он отмечает, что, согласно оригинальной интерпретации двухаспектного монизма, впервые представленной известными швейцарскими учеными Карлом Густавом Юнгом и Вольфгангом Паули, комплементарная комбинация ментальных и физических аспектов образует единую реальность (Atmanspacher H. 20th Century Variants of Dual-Aspect Thinking // Mind&Matter, 2014. №12(2). P. 245-288). Поближе присмотреться к комплементарному двухаспектизму Паули предлагает и современный аналитический философ Вильям Сигер (Seager W. A New Idea Of Reality: Pauli on the Unity of Mind and Matter // Mind and Matter. 2011. № 9(1). P. 37-52). Тем не менее, данная интерпретация двухаспектного принципа не является особенно популярной в аналитической философии, хотя охотно принимается некоторыми философствующими учеными.

материи и не сводить его к состоянию мозга, а также не отделять наблюдателя от наблюдаемого; вместо этого он постулирует целостную психофизическую вселенную, неотъемлемой частью которой мы являемся и которую можно познать принципиально различными способами:

«Индивидуальные сознательные представления являются перспективами. То есть точный способ, посредством которого сущности, события и процессы преобразуются в переживания, зависит от местоположения в пространстве и времени данного наблюдателя и точного сочетания перцептивных, когнитивных, аффективных, социальных, культурных и исторических влияний, которые вовлечены в “построение” данного опыта. В этом смысле каждая сознательная конструкция является приватной, субъективной и уникальной. Содержимое сознания обеспечивает представление о более широкой Вселенной, придавая ей вид трехмерного феноменального мира [...]. В этом смысле существует одна Вселенная (вещь-сама-по-себе) с относительно дифференцированными частями в виде сознательных существ, подобных нам, каждое из которых обладает уникальным, сознательным взглядом на большую Вселенную, частью которой оно является. Поскольку мы есть части Вселенной, мы участвуем в рефлексивном процессе, посредством которого Вселенная переживает саму себя».

Подобные рассуждения в некотором смысле сближают теорию Велманса с энактивизмом³⁰¹, некоторыми идеями буддийской философии, а также со своеобразным идеализмом, что, однако, не способствует широкому признанию концепции рефлексивного монизма среди аналитических философов и ученых (хотя, как однажды иронично отметил Чалмерс в одной из своих статей: «Если

³⁰¹ В одной из своих статей Франсиско Варела особо отмечает достоинства предложенной Вэлмансом концепции (Varela F. J. Present time consciousness // Journal of Consciousness Studies. 1999. №6(2–3). P. 111–140).

вы начинаете как материалист, то затем становитесь дуалистом, затем панпсихистом и заканчиваете как идеалист»³⁰²).

Несмотря на то, что все рассмотренные выше варианты развития двуаспектного принципа информации сталкиваются с теми или иными концептуальными трудностями, они, как полагает сам Чалмерс, не умаляют эпистемологической значимости подобной прототеории. О том, как именно представленные подходы к толкованию информационных и вычислительных пространств (и, соответственно, связанного с ними сознания) соотносятся с вопросами, касающимися некоторых гносеологических и эпистемологических проблем, философ рассуждает в ряде своих статей, посвященных концепции структурного реализма.

4.6. Информационно-вычислительный подход Д. Чалмерса и концептуальный структурный реализм

Итак, стратегия решения «трудной проблемы» сознания, предложенная Чалмерсом и рядом других современных аналитических философов, несколько отличается от стратегий, представленных в рамках большинства рассмотренных ранее в рамках настоящего исследования информационных теорий сознания. Не отдавая предпочтения ни одной из популярных сегодня вычислительных моделей, Чалмерс предлагает собственный подход к толкованию вычислительных пространств, который, по его утверждению, потенциально совместим как с положениями классического компьютеризма, так и с моделями динамических систем, различными вариантами нейросетевой установки и т.д. Выдвигая двуаспектный принцип информации в качестве своеобразной основы для «минимального компьютеризма», способной пролить свет на форму естественного закона, обуславливающего появление

³⁰² Chalmers. D.J. Idealism and the Mind-Body Problem // [Электронный ресурс] // Персональный сайт Д. Чалмерса. URL: <http://consc.net/papers/idealism.pdf> (дата обращения: 12.05.2021).

феноменального опыта у вычислительных систем, Чалмерс указал, что подобный принцип обладает статусом «прототеории», то есть своеобразного онтологически нейтрального «трафарета» или схемы, в которые можно вписывать более конкретные законы и различные концепции сознания; действительно, как было показано выше, при том или ином толковании двуаспектный принцип информации может получить свое дальнейшее развитие в сторону различных метафизических теорий. И хотя сам философ долгое время отдавал предпочтение развитию двуаспектного принципа в духе конститутивного панпсихизма, он никогда не отрицал правомерность тех или иных версий теорий эмерджентизма и признания за феноменальными качествами статуса эпифеноменальных; в некоторых своих недавних статьях философ и вовсе предлагает поближе присмотреться к идеализму или иллюзионистским подходам³⁰³.

Подобная неоднозначность предложенной стратегии решения проблемы сознания нередко ставит многих исследователей в тупик: с одной стороны, присущий Чалмерсу плюрализм взглядов открыл дорогу широкому спектру философских концепций, многие из которых начали восприниматься всерьез исключительно благодаря работам австралийского философа; с другой стороны, подобный подход несколько дискредитирует выдвигаемые Чалмерсом идеи и заставляет усомниться в их надежности. Как уже было отмечено выше, одним из популярных замечаний к предложенной Чалмерсом интерпретации информационных и вычислительных состояний в структурно-каузальных терминах является довод о том, что она не приносит никакой эпистемической пользы для философских и научных теорий, так как совмещает в себе различные

³⁰³ Chalmers D.J. The Meta-Problem of Consciousness // *Journal of consciousness studies*. 2018. №25 (9-10). P. 6-61; Chalmers. D.J. Idealism and the Mind-Body Problem // [Электронный ресурс] // Персональный сайт Д. Чалмерса. URL: <http://consc.net/papers/idealism.pdf> (дата обращения: 12.05.2021).

Любопытно, что в некоторых своих статьях, посвященных философскому анализу компьютерных симуляций и виртуальной реальности, Чалмерс отмечает, что информационный подход к объяснению сознания может быть также совмещен со своеобразным видом картезианского дуализма (Chalmers D.J. *The Matrix As Metaphysics* // *Philosophers Explore the Matrix* (ed. by C. Grau). Oxford University Press. 2005. pp. 132-169). Некоторые исследователи также прослеживают тенденцию развития философских взглядов Чалмерса в сторону к интеракционизму (см. например: Васильев В.В. Трудная проблема сознания. М.: Прогресс-традиция. 2009. С. 182-184).

аспекты толкования понятия информации. Обращение Чалмерса к информационному терминологическому аппарату во многом обусловлено его первым образованием, влиянием его учителя Дугласа Хофштадера, а также рядом факторов, в той или иной степени оказавших влияние на всех сторонников информационных теорий сознания, среди которых основополагающими, как мы выяснили, являются потребность в совмещении концептуальных словарей философии и современной науки, а также необходимость в обнаружении тематически нейтральных понятий и категорий, применимых для описания природы ментального опыта. Подобные доводы могут натолкнуть на мысль, что информационное объяснение сознания в философии Чалмерса – это не более чем «абстрактный фокус» (выражение В.В. Васильева³⁰⁴), никак не способствующий развитию его теории, ведь, как и в случае с теорией Деннета, возможно предположить, что используемые австралийским философом понятия могут быть замещены другими, более приемлемыми концептуальными аналогиями.

Однако сам Чалмерс не столь категоричен в отношении к принимаемой им трактовке информационных и вычислительных состояний: философу представляется допустимым подходить к информационно-вычислительному подходу с позиции структурного реализма, в рамках которого информация, которой приписываются ментальные и физические аспекты, может рассматриваться в качестве категории, играющей некоторые фундаментальные структурные роли.

Как известно, методология структурного реализма признается многими учеными в качестве наиболее приемлемой и перспективной основы научного знания. Вопрос о том, как мы можем быть уверены в том, что известные нам факты соответствуют некоторой реальной структуре мира, является одним из главных вопросов философии, и попытки преодолеть так называемое «картезианское сомнение» были предприняты многими известными мыслителями. В своих работах по эпистемологии и онтологии Чалмерс также обращается к данному вопросу, анализируя широко известный аргумент

³⁰⁴ Васильев В.В. Трудная проблема сознания. М.: Прогресс-традиция. 2009. С. 171.

Декарта, который можно представить следующим образом: так как мы не можем достоверно отрицать существование условного «злого демона», управляющего всеми нашими чувственными переживаниями и нашим восприятием действительности, мы не можем точно знать что-либо об истинности наших эмпирических суждений о внешнем мире. Многие философы, в числе которых Джордж Эдвард Мур, Хилари Патнэм, Бертран Рассел и др., в основном пытались доказать ложность посылки, ссылаясь на аргументы здравого смысла, лучшего объяснения и т.д. В то же время Чалмерс предлагает поставить под сомнение вывод подобного аргумента, отмечая, что даже при осуществлении сценария со «злым демоном» наши убеждения о мире могут претендовать на истинность в силу своей эквивалентности некоторым структурным утверждениям.

В ряде своих книг и статей философ представляет версию структурного реализма, которую именует концептуальным структурным реализмом, отличающимся как от эпистемического, так и от онтического структурного реализма; согласно данной версии, наши научные теории и существующие в их рамках положения эквивалентны структурным теориям и структурным положениям³⁰⁵. Как известно, в рамках эпистемической трактовки структурного реализма утверждается, что все познаваемые научные истины соответствуют структурным истинам, то есть наши научные концепции описывают лишь наблюдаемые нами реляционные свойства, но ничего не говорят о реальной природе вещей. В то же время, согласно онтическому структурному реализму, фундаментальная реальность, описываемая в подобных концепциях, также обладает структурными свойствами. Можно сказать, что концептуальный структурный реализм подразумевает эпистемический структурный реализм, однако сам не следует из последнего, ведь некоторые научные положения могут быть структурными, не будучи при этом познаваемыми. Из концептуального структурного реализма не следует онтический структурный реализм, так как

³⁰⁵ Chalmers D. J. *Constructing the World*. Oxford University Press. 2012; Chalmers D.J. *Structuralism as a Response to Skepticism* // *Journal of Philosophy*. 2018. №115(12). P. 625-660; Chalmers D.J. *The Virtual and the Real*. // *Disputatio*. 2017. № 9(46). P. 309-352.

возможно существование неструктурного по своей природе свойства, выполняющего в научной теории структурную роль. Таким образом, концептуальный структурный реализм является реализмом по отношению к содержанию научных теорий, но ничего не утверждает об истинности данных теорий (поэтому, во избежание путаницы, Чалмерс предпочитает называть свою концепцию просто структурализмом).

Философ отмечает, что методология концептуального структурного реализма нашла свое частичное выражение в работах английского математика Фрэнка Рамсея, предложившего формальный способ представления научных теорий, элиминирующий при этом соответствующие теоретические термины³⁰⁶. Будучи дополненной условиями о существовании некоторых фундаментальных и естественных отношений, в числе которых номические и каузальные отношения³⁰⁷, подобная формализация, с точки зрения Чалмерса, вполне удовлетворяет положениям концептуального структурного реализма, так как соответствует следующему утверждению: существуют такие естественные свойства и отношения, что эти свойства и отношения находятся в определенных номических/каузальных связях друг с другом и с нашим опытом.

Используя данное положение, Чалмерс обращается к картезианскому аргументу и выдвигает следующее возражение против глобального скептицизма:

«1. Все физические утверждения P эквивалентны неким структурным утверждениям $S(P)$ ».

³⁰⁶ Впоследствии подобное логическое построение, также известное как предложение Рамсея, было развито американскими философами Рудольфом Карнапом и Дэвидом Льюисом. Карнап приводит следующий пример перевода обычного предложения в предложение Рамсея: допустим, некий объект X имеет массу 5 кг; тогда для формализации необходимо преобразовать теоретические термины класса и отношения в переменные, а также указать все кванторы существования. Тогда формула будет представлена конъюнкцией следующего вида:

$(\exists C1)(\exists C2)\dots (\exists R1)(\exists R2)\dots [C1\dots C2\dots R1\dots R2\dots; R1\dots R2\dots O1\dots O2\dots \text{ и } R1(X) = 5]$

где C – некий класс объектов, R - отношения веса, O – термины наблюдения. Для других теоретических терминов или законов требуется добавление еще больших переменных и кванторов существования (Карнап. Р. Философские основания физики: Введение в философию науки. М.: Издательство ЛКИ, 2008. С. 336).

³⁰⁷ На необходимость дополнения идей Рамсея в 1927 году обратил внимание математик Джеймс Ньюман, отметивший, что предложения Рамсея достаточно слабы для выражения содержания многих научных теорий, так как их структура является недостаточной для однозначного выделения каких-либо отношений, существующих в мире.

2. Для всех позитивных физических утверждений P , если $S(P)$ является истиной в случае, если мы не находимся в картезианской симуляции, то оно является истинным и в картезианской симуляции.

3. Для всех позитивных физических утверждений P , если P является истинным в случае не-симуляции, то оно истинно и в любой симуляции»³⁰⁸.

Идея о том, что структурные утверждения, существующие в созданной демоном «симуляции», допустимы и в случае «не-симуляции», является ключевой для данного аргумента: Чалмерс хочет доказать, что каузальная структура, присутствующая в «симуляции», является изоморфной каузальной структуре, присутствующей во «внешнем мире», которому принадлежит условный «злой демон». С точки зрения философа, каждому физическому объекту нашего мира должен соответствовать некий «виртуальный» объект, будь то некие паттерны репрезентаций в разуме демона или, например, информационная структура компьютера, симулирующего нашу реальность. Изоморфизм подобных каузальных отношений позволяет, как полагает философ, говорить о применимости дополненного метода Рамсея к описанию истинных утверждений о характеристиках «внешнего» мира.

Безусловно, утверждение об изоморфизме структурных отношений можно оспорить, и Чалмерс выделяет несколько потенциальных возражений против данного тезиса: во-первых, известные нам свойства физического мира могут быть менее естественными, чем их аналоги во «внешнем» мире, а потому могут быть вообще не связаны с какими-либо фундаментальными законами; во-вторых, каузальные связи между «виртуальными» объектами и физическими объектами могут быть сильно опосредованными, и, следовательно, симуляция будет иметь весьма опосредованную каузальную структуру; в-третьих, симуляция может быть встроена в мир, содержащий множество дополнительных фундаментальных законов и соответствующие им дополнительные структуры;

³⁰⁸ Chalmers D.J. Structuralism as a Response to Skepticism // Journal of Philosophy. 2018. №115(12). P. 639.

наконец, симуляция могла бы быть легко спроектирована кем-то, способным в любой момент изменить ее в любое время. Однако философ не считает, что представленные возражения способны как-либо опровергнуть вторую посылку его аргумента: так, учитывая, что в аргументе говорится лишь о позитивных утверждениях о физическом мире, то любое возражение, апеллирующее к существованию во «внешнем» мире некоторых дополнительных, более фундаментальных или опосредованных структурных свойств, не умаляет правомерности использования известных нам структурных утверждений.

Разумеется, предложенный Чалмерсом структурный подход не лишен внутренних концептуальных трудностей: в частности, остается не вполне понятно, как в его рамках можно объяснить такие фундаментальные физические категории, как пространство, время и масса (хотя в некоторых своих работах философ предлагает концепцию пространственно-временного феноменально функционализма, согласно которому подобные физические явления могут быть проанализированы в терминах их воздействия на наши ощущения пространства, времени, массы, цвета и т.д.³⁰⁹). Кроме того, концептуальный структурализм вряд ли может быть представлен в качестве сильного аргумента против скептических воззрений, касающихся вопросов о сознании других субъектов, оснований нашего познания и пр. Наконец, некоторые философы отмечают, что предложенный Чалмерсом структуралистский подход противоречит базовому положению его теории натуралистического дуализма, основанному на отрицании возможности редукционистского объяснения феноменального сознания в тематически-нейтральных (структурных) терминах³¹⁰. Тем не менее, философу кажется очевидным, что подобное объяснение является более

³⁰⁹ Chalmers D.J. Finding Space in a Nonspatial World // [Электронный ресурс]: Персональный сайт Д. Чалмерса. URL: <http://consc.net/papers/finding.pdf>.

³¹⁰ Подобный аргумент представил австралийский философ Дэниел Столджар в своей статье «Чалмерс против Чалмерса» (Stoljar D. Chalmers v. Chalmers // *Noûs*. 2020. №54(2). P. 469-487). При этом сам Чалмерс отмечает, что Столджар не совсем верно трактует основные положения его теории концептуального структурализма, которые при должной интерпретации являются вполне совместимыми с «аргументом зомби». Более того, так как аргумент Столджара нацелен на доказательство несовместимости структурализма Чалмерса с его дуалистическими представлениями о природе сознания, из подобных доводов не следует, что прочие теории сознания, отличные от теории дуализма свойств, являются также несовместимыми со структурным подходом (Chalmers D.J. Spatiotemporal Functionalism v. The Conceivability of Zombies // *Nous*. 2020. №54. P. 488-497).

приемлемым, чем принятие глобального скептицизма, и, вместе со стратегией лучшего объяснения, оно способно дать надежду на познание некоторых фундаментальных законов природы.

Концептуальный структурализм способен пролить свет на непростую проблему, связанную с возможностью зависимости физических свойств известной нам реальности от чьего-либо сознания: Чалмерс не считает, что из подобной возможности может следовать неверность структурных утверждений об известной нам реальности. К примеру, может оказаться, что некоторая форма панпсихизма верна, и фундаментальная физика реализуется на микрофизическом уровне сознанием микросубъектов; однако в этом случае в мире по-прежнему существовали бы физические сущности. Может даже оказаться, что истинным является определенный вид макроидеализма или даже космоидеализма, и физика Вселенной в целом реализуется структурой и динамикой некоторого единого сознания; но и в этом случае физические сущности также вполне имеют право на существование, даже если они фундированы некоторой ментальностью. Чалмерс отмечает, что сценарии, предполагающие, что реальность контролируется разумом некоего демона, можно разобрать по аналогии: в них физические сущности играют те же структурные роли, которые они должны играть, даже если данные сущности на самом деле являются ментальными. Таким образом, возможность этих сценариев не угрожает знанию обычных структурных утверждений: они могут лишь угрожать знаниям более общего, метафизического порядка.

Таким образом, развиваемый Чалмерсом информационный подход к описанию природы сознания, основанный на положениях нередуктивного функционализма, может быть совмещен со структуралистским подходом, концептуальная версия которого, как демонстрирует философ, устойчива к различным онтологическим интерпретациям сознания. Безусловно, можно задать вопрос о том, почему именно информационные отношения следует рассматривать в качестве базовых структурообразующих отношений, лежащих в основе искомой теории сознания. Самым простым вариантом решения вопроса

об онтологическом статусе информации Чалмерс считает признание своеобразного дефляционизма, по которому эту категорию следует рассматривать лишь как удобный теоретический конструкт или инструмент, с помощью которого мы строим наши структурные теории. Весьма вероятно, что фундаментальная реальность имеет совершенно иную природу, познание которой нам недоступно в силу ограниченности наших познавательных возможностей, – в этом случае мы можем попросту принять положения концептуального структурализма и сказать, что известные нам научные утверждения являются эквивалентными неким структурными утверждениям, которые, в свою очередь, могут быть изоморфны структурным утверждениям о природе фундаментальной реальности; при этом форма применяемых нами структурных утверждений может быть совершенно разной, а потому они могут быть описаны в информационных, каузальных или других удобных для нас терминах.

Несмотря на то, что дефляционизм по отношению к информационным пространствам является наиболее простым способом интерпретации подобных пространств, он не является единственным, так как некоторые ученые и философы не отрицают возможности толкования информации как онтологически фундаментальной категории.

4.7. Онтологический статус информационных пространств и информационный реализм

Хотя в «Сознающем уме» Чалмерс отмечает, что предположение о том, что информация является элементарной характеристикой физического мира, отнюдь не является необходимым для изучения искомых психофизических законов, объясняющих природу нашего сознания, философ, подобно Кеннету Сейру, не отрицает возможности рассмотрения информационных пространств как онтологически фундаментальных. Завершая свою книгу, Чалмерс отводит несколько страниц рассуждениям о возможных вариантах интерпретации

метафизики информации, отмечая, в частности, известную концепцию «Всё из бита» Джона Уилера, панкомпьютационалистскую теорию Эдварда Фредкина, гипотезу математической Вселенной Макса Тегмарка и прочие идеи так называемой «цифровой физики»³¹¹. В своих более поздних работах, посвященных рассуждениям об онтологии виртуальных и цифровых пространств, Чалмерс также предлагает лучше присмотреться к подобным гипотезам³¹².

Стоит отметить, что далеко не каждый подход, предполагающий, что информационные пространства являются онтологически фундаментальными по отношению к физике и сознанию, сводится к цифровым или панкомпьютационалистским гипотезам (которые, к слову, также не обязательно предполагают друг друга: так, Джон Уилер являлся сторонником первой, но ничего не говорил о возможности второй). К примеру, современный философ Лучано Флориди, специализирующийся на философии информации³¹³, отмечает,

³¹¹ Строго говоря, «цифровая физика» представляет собой совокупность теорий, утверждающих, что Вселенная является информационной и вычислимой. Одним из первых идею о вычислительном характере Вселенной предложил в 1969 году один из пионеров компьютеростроения, немецкий инженер Конрад Цузе (Zuse K. *Calculating Space*. Cambridge, MA: MIT Press, 1970); затем формальное представление данной идеи было разработано американским физиком Фредериком Кантором, а также такими физиками, как Сетом Ллойдом (Lloyd S. *Programming the Universe: A Quantum Computer Scientist Takes On the Cosmos*. Knopf, 2006) и Дэвидом Дойчем, отстаивающим панкомпьютационалистское представление о Вселенной как гигантском квантовом компьютере. Дойч, в частности, предложил дополнить классический тезис Чёрча-Тьюринга, согласно которому любая вычислительная машина, по мощности эквивалентная машине Тьюринга, способна вычислять всё, что может вычислять человек, если ей дано достаточно времени; согласно дополненному Дойчем тезису, универсальное компьютерное устройство способно моделировать любой конечный физический процесс (Deutsch D. *The Fabric of Reality*. New York: Allan Lane, 1997).

³¹² Chalmers D.J. *The Matrix As Metaphysics // Philosophers Explore the Matrix* (ed. by C. Grau). Oxford University Press. 2005. pp. 132-169; Chalmers D.J. *The Virtual and the Real. // Disputatio*. 2017. № 9(46). P. 309-352; Chalmers D.J. *The Virtual as the Digital. // Disputatio*. 2019. № 11(55). P. 453-486.

Особенно подробно данная тема будет рассматриваться в книге Чалмерса «Реальность+. Виртуальные миры и философские проблемы» («Reality+: Virtual Worlds and the Problems of Philosophy»), которая будет опубликована в издательстве «W.W. Norton Press» в январе 2022 года.

³¹³ Само понятие «философия информации» было введено самим Флориди в 1990-х годах, хотя он отмечает, что первым о необходимости выделения подобной области философского знания упоминал еще английский философ Аарон Сломан, автор книги «Компьютерная революция в философии» (Sloman A. *The Computer Revolution in Philosophy: Philosophy, science and models of mind*. Harvester press and Humanities press, 1978). Флориди определяет философию информации как область, занимающуюся изучением природы информации, а также развитием и применением методов информационной и вычислительной теорий к решению философских задач (Floridi L. *What is the Philosophy of Information? // Metaphilosophy*, 2002. №33(1-2). P. 123-145).

Стоит отметить, что среди отечественных философов исследованию проблем, так или иначе связанных с изучением природы информации, посвятили свои работы А.Д. Урсул (Урсул А.Д. *Природа информации*. М.: Политиздат, 1968; Урсул А.Д. *Проблема информации в современной науке*. М.: Наука, 1975), К.К. Колин (Колин К.К. *Философия информации и фундаментальные проблемы современной информатики // Alma mater* (Вестник высшей школы), 2010. №1. С. 29-35), А. В. Соколов (Соколов А. В. *Философия информации: учеб. пособие*. Челябинск: ЧГАКИ, 2011).

что следует различать «цифровую» и «информационную» онтологии³¹⁴. Философ полагает, что предзаданная система различий, являющаяся основанием любой вычислительной парадигмы, представляет собой лишь способ, с помощью которого реальность переживается и концептуализируется неким эпистемическим агентом, использующим тот или иной уровень абстракции, а потому не способна пролить свет на реальное положение дел. Под уровнем абстракции философ понимает некоторую эпистемическую позицию, которую условный наблюдатель принимает по отношению к некоторому наблюдаемому явлению: исследование каждого явления, характеризующегося неким набором данных, сводится к построению модели данного явления, несущей определенную информацию; чем ниже уровень абстракции, тем больше информации о явлении мы можем получить при использовании данного уровня абстракции. Отсюда следует, что каждое наблюдаемое явление потенциально несет в себе крайне большое количество информации для наблюдателя, и, хотя каждый уровень абстракции имеет право на существование, агенту требуется выбрать тот уровень абстракции, который принесет ему релевантное количество информации при данном эпистемическом запросе. Флориди полагает, что как только мы осознаем, что никакой познавательный процесс не может обходиться без использования того или иного уровня абстракции, мы также поймем, что тот факт, что некоторая система (пусть даже Вселенная) может быть смоделирована и представлена в цифровом виде, еще не означает, что внутренняя природа этой системы является цифровой; подобный довод вынуждает нас обратиться к онтологиям, постулирующим лишь реальность реляционных отношений, наподобие структурного реализма или информационного реализма, являющегося разновидностью первого.

Информационный структурный реализм постулирует существование независимых от наблюдателя информационных структур, модели которых являются, по Флориди, более объяснительно-, инструментально- и прогностически-успешными, чем многие другие используемые исследователями

³¹⁴ Floridi L. Against digital ontology // Synthese, 2009. №168 (1). P. 151-178.

эпистемические модели. Утверждение о том, что предельная природа реальности носит информационный характер означает, что имеет смысл принять такой уровень абстракции, при котором наши теории будут представлять реальность как независимую от разума и состоящую из структурных объектов, не являющихся в обязательном порядке материальными, а представляющих собой скорее когерентные наборы данных, то есть различия *de re*. Эпистемическая работа с подобными структурными объектами равносильна ограничению доступности одних конструкторов и выявлению других, зависящих от принимаемого уровня абстракции. Последние могут быть использованы теорией в качестве ее входных данных для получения информации, релевантной нашему запросу.

Безусловно, подобная концепция представляет собой версию эпистемического структурного реализма, и может показаться, что она ничего не говорит об онтологической фундаментальности информационных пространств; однако Флориди хочет показать, что в данном случае разговор об информационном реализме является вполне оправданным, и обоснование данного утверждения философ посвящает несколько работ³¹⁵.

Во-первых, Флориди пытается продемонстрировать, каким образом можно совместить эпистемический и онтический структурный реализмы с помощью упомянутого метода уровней абстракций. Эпистемический структурный реализм, согласно философу, подразумевает минималистский подход, по которому теория оправдана в принятии некоторого уровня абстракции, который онтологически связывает саму теорию с реалистической интерпретацией только структурных свойств системы, определенных ее моделью; это минимальное онтологическое обязательство первого порядка касается лишь знаний о структурных свойствах исследуемой системы, и кажется, что принятие какого-либо уровня абстракции, подразумевающего более высокую степень онтологического обязательства, является рискованным и подозрительным.

³¹⁵ Floridi L. Informational Realism. 2003. // [Электронный ресурс]: ReaearchGate.net. URL: https://www.researchgate.net/publication/262350693_Informational_Realism (дата обращения: 03.11.2021); Floridi L. A Defence of Informational Structural Realism. // Synthese, 2008. №161(2). P. 219-253.

Однако можно также заключить, что независимо от того, чем система и ее компоненты (т. е. объекты или отношения) являются сами по себе, они должны быть такими, чтобы позволить теории охватить их структурные и реляционные свойства. Отсюда следует вопрос о поиске релевантного уровня абстракции, который позволяет теории описать структурные свойства системы, определенные моделью в рамках данной теории: выходит, что приемлемый в данном случае уровень абстракции подразумевает, что сами отношения понимаются в структурных терминах. Можно сказать, что Флориди таким образом использует принцип «бритвы Оккама», согласно которому разумно допустить, что реляционные отношения являются структурными, поскольку это все, что нужно теории для обоснования ее онтологического обязательства первого порядка. Следовательно, заключает философ, эпистемический и онтический структурный реализмы в некоторой степени предполагают друг друга³¹⁶.

Безусловно, нам по-прежнему необходимо более определенное понимание природы подобных структурных объектов. Флориди считает, что наилучшим способом объяснения их природы является информационное описание, согласно которому они представляют собой наборы данных, трактуемых как безусловные различия и потому являющихся исключительно реляционными сущностями (которые, разумеется, являются непознаваемыми сами по себе, так как никто никогда не имеет дело с чистыми данными, а только с уже интерпретированными). Поэтому, говоря об «информационных объектах», мы подразумеваем под ними не «объекты» в буквальном смысле слова, а скорее

³¹⁶ Нетрудно заметить, что суть данного аргумента частично совпадает с содержанием аргумента структурного изоморфизма и лучшего объяснения Чалмерса, который он использовал против позиции глобального скептицизма.

Разумеется, Флориди отмечает, что устранение несогласованности между эпистемическим и онтическим структурным реализмом еще не доказывает верность последнего. Принятие реализма по отношению к структурам требует поиска их оснований (релятов), которые сами по себе не могут быть определены в реляционных терминах, не впадая при этом в некоторый порочный круг или бесконечный регресс. Однако, если мы рассмотрим некий совершенно недифференцированный объект X (а отношение различия является фундаментальным для всех структурных отношений), то он будет ненаблюдаемым при любом возможном уровне абстракции; это эквивалентно утверждению, что X не существует (Флориди предлагает называть это модифицированным принципом тождества неразличимых Лейбница). Отсюда следует, что реляты не обязательно логически предшествуют отношению различия, ведь одно немислимо без другого.

«точки пересечения» определенных отношений. Чтобы лучше понять эту интуицию, Флориди использует аналогию с шахматными фигурами. Так, идентичность пешки определяется не столько ее физическими характеристиками, сколько ее состояниями по отношению к другим шахматным фигурам и логическому пространству игровой доски:

«Для хорошего игрока пешка – это только заполнитель, обозначающий “информационный объект”. Это не материальная вещь, а набор типизированных переменных, если говорить в терминах уровней абстракции, или ментальная сущность, выражаясь словами Беркли, или, используя выражение Юма, сущность, образованная пучком свойств, чье существование и природа определяются различиями и номологическими отношениями, которые характеризуют игру в шахматы»³¹⁷.

Таким образом, будучи версией структурного реализма, информационный реализм поддерживает как минимальное онтологическое обязательство первого порядка, так и онтологическую установку второго порядка, которая, как показывает Флориди, является эпистемически оправданной. Как агенты познания, мы можем напоминать Серла, запертого в его Китайской комнате: мы получаем данные о мире с одной стороны и производим информацию с другой, и смысловые основания данной игры зависят исключительно от используемого нами уровня абстракции³¹⁸. Ровно так же информационные системы могут быть изучены с разных сторон, учитывая тот или иной принимаемый нами уровень абстракции; не исключено, полагает философ, что взгляд на подобную систему изнутри будет характеризовать ее как ментальную³¹⁹.

³¹⁷ Floridi L. Informational Realism. 2003. // [Электронный ресурс]: ResearchGate.net. URL: https://www.researchgate.net/publication/262350693_Informational_Realism (дата обращения: 03.11.2021).

³¹⁸ Можно провести параллель между соображениями Флориди и идеями об онтологическом анти-реализме Чалмерса, полагающего, что выбор той или иной онтологической концепции зависит исключительно от нашей эпистемологической установки (Chalmers D.J. Ontological anti-realism // *Metametaphysics: New Essays on the Foundations of Ontology* (D. J. Chalmers, D. Manley R., Wasserman (eds.). Oxford University Press. 2009. P. 77-130).

³¹⁹ Floridi L. The Informational Nature of Personal Identity // *Minds and Machines*, 2011. №21(4). P. 549-566.

Как справедливо отметил Флориди, без выбора релевантного уровня абстракции мы столкнемся с бесконечным набором различий, которое, вероятно, является онтологически фундаментальным основанием природы, но, тем не менее, приводит нас к уже упомянутой проблеме информационной или вычислительной тривиальности. По утверждению философа, возможность описания любого явления как информационного скрадывает особенности философии информации как отдельной области философии с отдельным предметом исследования, буквально делая ее синонимом философии в целом (вспомним схожие рассуждения Д.И. Дубровского). По этой причине следует признать, что критерием целесообразности информационного анализа некоторого явления X является не проверка, может ли X быть сформулирован в информационных терминах, а скорее вопрос о том, что станет с X , если подобное явление вовсе не будет обладать информационной природой. И если оправданность применения подобного анализа в науках о мозге и сознании признается многими учеными и философами, то необходимость обращения к информационному концептуальному аппарату для описания онтологической основы физического мира является на сегодняшний день весьма спорной.

Заключение главы

Знаменитое выступление Дэвида Чалмерса на Туссанской конференции в 1994 году и публикация его книги «Сознающий ум» в 1996 году оказали значительное влияние на многие современные теории сознания, в числе которых оказались и информационные концепции. Постулирование «трудной проблемы» побудило многих исследователей выйти за рамки строгого функционального объяснения и искать дополнительные психофизические законы, способные объяснить возникновение у некоторых информационных и вычислительных систем феноменального опыта. Сам Чалмерс, будучи знакомым со многими информационными теориями сознания, а также со всеми трудностями, с которыми сталкиваются подобные теории, представил собственную версию

вычислительного подхода, основанную на представлении о наличии изоморфизма между каузальной структурой рассматриваемой физической системы и формальной структурой некоторого вычисления; несмотря на то, что данная трактовка вычислительных систем не обеспечивает подход жесткими критериями вычислительной реализации, подобный «минимальный компьютеризм» дает вполне однозначные ответы на некоторые аргументы, направленные против вычислительной установки в целом. Однако, чтобы решить проблему, связанную с возникновением феноменального опыта, Чалмерс основывает «минимальный компьютеризм» на протозаконе, выраженном в двухаспектном принципе информации. Двухаспектный подход к толкованию информационных пространств, удовлетворяющий принципам структурной когерентности и организационной инвариантности, в свою очередь, может быть развит в сторону конститутивного или эмерджентного панпсихизма, комплиментарного аспектизма и других теорий, развиваемых современными философами, среди которых Грегг Розенберг, Риза Мали, Сэм Колеман, Макс Вэлманс и др. Этот факт доказывает потенциальную совместимость информационного подхода к объяснению природы сознания с различными метафизическими концепциями. Не являясь приверженцем какой-либо конкретной вычислительной модели и не отрицая полностью правомерность той или метафизической теории сознания, Чалмерс согласует положения «минимального компьютеризма» и двухаспектного принципа информации с более общим методологическим подходом, который он называет концептуальным структурным реализмом. Доказывая, что наши структурные утверждения о мире могут соответствовать некоторым фундаментальным структурным законам природы, Чалмерс оправдывает структурно-каузальный подход, лежащий в основе его информационно-вычислительной установки. Философ также не отрицает возможности толкования информационных пространств как онтологически фундаментальных. Тем не менее, хотя различные версии цифровых онтологий и информационного структурного реализма являются достаточно популярными среди ряда современных ученых и

философов, эпистемическая оправданность применения информационного концептуального аппарата для описания предельной природы физической реальности вызывает некоторые сомнения.

Заключение

Середина XX века стала переломным моментом в истории аналитической философии сознания: попытка дезавуировать метафизику посредством применения последовательного логического и лингвистического анализа, представленная, в первую очередь, в работах сторонников логического бихевиоризма, обернулась крахом их проекта. Невозможность объяснения субъективных ощущений и ментальных переживаний в терминах соответствующих высказываний и анализа поведенческих диспозиций заставила многих философов усомниться в правомерности выбранной ими стратегии и заняться поиском новых решений, способных пролить свет на проблему возникновения ментального опыта. Одной из ведущих теорий, пришедших на смену бихевиоризма, стала типовая теория тождества, в рамках которой каждый тип сознательных состояний отождествлялся с соответствующим типом физических состояний; данный подход к толкованию природы субъективных переживаний, однако, также столкнулся с рядом характерных трудностей. Своеобразной альтернативой данным философским концепциям стал функционализм, согласно которому каждому ментальному переживанию соответствует определенное функциональное состояние системы или живого организма, определяющее ментальные состояния посредством описания их каузального влияния на другие состояния организма.

Параллельно с данными процессами происходило становление целого ряда новых научных дисциплин, развивавшихся на стыке математических, инженерных и компьютерных наук; коммуникационная теория Шеннона-Уивера, абстрактные вычислительные машины Тьюринга, математические модели нейронных систем МакКаллока-Питтса – эти и многие другие научные достижения середины прошлого века оказали огромное влияние на различные междисциплинарные исследования, направленные, в частности, на изучение процессов мышления. Активное развитие кибернетики и теории искусственного интеллекта сыграли основополагающую роль в ходе так называемой

«когнитивной революции», ставшей толчком к развитию ряда научных дисциплин, объединенных общим названием «когнитивистика». Концептуальный и терминологический аппараты информационной теории, теории вычислимости и теории искусственного интеллекта легли в основу психолингвистики, нейробиологии, антропологии и, разумеется, философии сознания.

Объединение подобных концептуальных схем с философской программой функционализма привело к возникновению в 1960-х годах первых вычислительных теорий сознания, совокупность которых сегодня принято называть классическим компьютеризмом. Активное использование популярной тогда компьютерной метафоры послужило развитию представления, согласно которому мышление уподоблялось вычислениям, реализуемым абстрактной машиной Тьюринга, а ментальные состояния человека сопоставлялись с логическими состояниями данной машины. Появление различных вариантов теорий репрезентативизма способствовало дальнейшему развитию вычислительных подходов, трактующих мыслительные процессы в терминах формальных символическо-алгоритмических вычислений, осуществляемых над соответствующими ментальными репрезентациями. И хотя в 1980-х годах идеи классического компьютеризма стали отступать на второй план, уступая место различным версиям нейросетевых моделей, вычислительная установка продолжала оставаться одной из наиболее влиятельных философских и научных программ, нацеленных на исследование работы мозга.

Тем не менее, далеко не все философы того времени, активно использовавшие информационный терминологический аппарат, стали развивать свои теории в рамках какой-либо строго определенной вычислительной парадигмы. Так, американский философ Фред Дретске в своей теории репрезентативизма полностью заменил популярное в то время понятие «вычисление» на выражение «информационная обработка», совмещая математическую трактовку информации Шеннона с концепцией

натурализированной семантической информации Пола Грайса. Такие философы, как Кеннет Сейр и Д.И. Дубровский построили оригинальные информационные теории сознания, заимствующие некоторые объяснительные модели из кибернетики и теории систем. Философ Дэниел Деннет и вовсе предпочел использовать концептуальный аппарат теории информации в абстрактно-метафорическом ключе. Упомянутые теории сознания объединяет как единая стратегия объединения философских построений с достижениями современных эмпирических и математических наук, так и попытка преодоления разрыва между строгим функциональным и субъективным описаниями сознания посредством использования таких тематически нейтральных категорий, как «информация», «сигнал», «вычисление» и т.д.

Несмотря на популярность информационных и вычислительных философских теорий, к 1980-м годам многие исследователи пришли к выводу, что вычислительная установка сталкивается с целым рядом сложных проблем, на многие из которых философы не способны найти последовательных решений и по сей день. Противники применения информационного словаря для описания природы сознания указывают, в первую очередь, как на общую неопределенность самих понятий «информация» и «вычисление», так и на отсутствие у ученых единого мнения по поводу необходимых и достаточных критериев вычислительной реализации. Аргумент тривиальности, допускающий возможность интерпретации любых физических систем как вычисляющих или осуществляющих обработку информации, ведет к неутешительным выводам: философы вынуждены либо признать положения панкомпьютериализма, либо согласиться с утверждением об относительности понятий «вычисление» и «информация» и зависимости их значений от установок условного интерпретатора.

Отдельная проблема связана с невозможностью информационно-вычислительного объяснения содержательной стороны ментальных переживаний. Данная проблема, источником которой отчасти является концептуальное смешение синтаксического и семантического аспектов

толкования информации, получила особую известность благодаря мысленному эксперименту Джона Серла «Китайская комната» и долгое время представлялась одной из самых насущных проблем для большинства информационных теорий сознания. Несмотря на то, что многие философы не только отмечали несостоятельность самого аргумента Серла, но и выдвигали различные теории ментальных содержаний, данная проблема, как и проблема тривиальности, остается пока нерешенной.

Исследование семантических свойств информационных компонентов ментального опыта также затронуло вопросы об онтологии феноменального сознания. Известные мысленные эксперименты Нэда Блока, Томаса Нагеля, Фрэнка Джексона и Дэвида Чалмерса показали общую несостоятельность большинства информационных теорий к объяснению феноменального сознания, неизбежно сталкивающихся с проблемой, связанной с описанием причин появления субъективных переживаний. Признание «трудной проблемы» побудило многих философов выйти за рамки строгого функционального объяснения и заняться поиском новых теорий сознания.

Тем не менее, данные критические аргументы и замечания не способствовали уменьшению популярности информационной интерпретации в аналитической философии, и многие современные исследователи стали разрабатывать концепции, призванные не только прояснить суть ключевых для данных информационных теорий концептуальных категорий, но и улучшить наше понимание возможной связи, существующей между информационными и феноменальными состояниями.

Один из наиболее последовательных информационно-вычислительных подходов к решению проблемы сознания представил Дэвид Чалмерс. Данный подход можно условно разделить на три составляющие:

- 1) «Минимальный компьютеризм».

В соответствии с данной методологической установкой, потенциально совместимой со многими известными сегодня вычислительными парадигмами, вычислительная реализация определяется структурным соответствием между

каузальной организацией некоторой физической системы и формальной структурой некоторого вычисления. Признавая возможность панкомпьютериализма, философ, тем не менее, отвергает возможность трактовки всякой физической системы как осуществляющей любое возможное вычисление. По замыслу Чалмерса, каузально-структурная трактовка избавляет вычислительную установку от проблем, связанных с возможностью антиреалистской интерпретации вычислительных состояний, а также от проблемы соответствия между информационными и вычислительными структурами.

Несмотря на то, что «минимальный компьютеризм», который сам философ трактует как частный случай нередуктивного функционализма, решает некоторые проблемы, с которыми обыкновенно сталкивается большинство информационных теорий сознания, данный подход еще не способен предоставить последовательное решение проблемы, связанной с объяснением природы феноменального опыта.

2) Двухаспектный принцип информации.

В рамках данного принципа принимается, что информация распространена повсеместно и может быть представлена как субъективно, так и объективно, что означает ее возможную феноменальную и физическую реализуемость. Трактую информацию в каузальных терминах, то есть как «различие, производящее различие», философ ставит данный принцип в один ряд с ключевыми для «минимального компьютеризма» принципами структурной когерентности и организационной инвариантности, доказывая таким образом наличие структурного изоморфизма между вычислительными и информационными состояниями. Будучи лишь своеобразной «прототеорией», двухаспектный принцип информации может быть потенциально развит в сторону различных метафизических теорий сознания: в тех или иных версиях данный принцип используют такие философы, как Грегг Розенберг, Сэм Колеман, Риза Мали, Макс Вэлманс и др.

3) Концептуальный структурный реализм.

Структурно-каузальная трактовка вычислительных и информационных пространств совместима с более широким представлением, затрагивающим вопросы, касающиеся не только онтологии сознания, но также эпистемологии и философии науки. Чалмерс отстаивает свою версию научного структурного реализма, отличную как от эпистемического, так и онтического структурного реализма: согласно данной концепции, существуют такие естественные свойства и отношения, что эти свойства и отношения находятся в определенных номических и каузальных связях друг с другом и с нашим опытом. Философ полагает, что принятие подобной теории в какой-то мере избавляет нас от выводов, ведущих к глобальному скептицизму в рамках философии науки; будучи совмещенным с принципами «минимального компьютеризма» и информационного двуаспектизма, концептуальный структурный реализм также способствует избавлению от своеобразного «мистерианства» в рамках философии сознания. С точки зрения философа, информационное объяснение способно обеспечить наши теории некоторыми знаниями о каузальном действии и номической организации фундаментальной реальности, познание подлинной природы которой, однако, невозможно. Термин «информация», таким образом, является удобным концептуальным конструктом, играющим в теории вспомогательную роль. Использование подобного термина для описания тех или иных структурных аспектов реальности вовсе не означает, что предельная природа данной реальности с необходимостью должна быть представлена как информационная.

Итак, в ходе настоящего исследования были достигнуты все поставленные цели и решены все сформулированные ранее исследовательские задачи: выявлены основные предпосылки появления информационных теорий сознания в рамках аналитической философской традиции, проанализированы наиболее известные версии данных теорий, представленные в работах отечественных и зарубежных философов, последовательно разобраны основные критические аргументы, направленные против информационной интерпретации природы

сознания, а также рассмотрены возможные пути развития информационных теорий, предложенные современными философами и учеными.

Несмотря на то, что одним из основных результатов данного исследования является вывод об отсутствии на сегодняшний день удовлетворительной информационной теории сознания, проект развития философских концепций, трактующих природу ментального опыта как информационную, представляется достаточно перспективным. Стратегию, развиваемую Чалмерсом и многими другими сторонниками предложенного им информационно-вычислительного подхода к объяснению сознания, не следует рассматривать как полноценную теорию, а скорее как возможную методологическую базу для дальнейших теоретических и практических исследований. Представляется, что наиболее актуальный вопрос, стоящий перед подобными исследованиями, касается не столько проблем, возникающих в ходе решения «трудной проблемы» сознания, сколько проблем, связанных с особенностями использования информационного концептуального инструментария в качестве основы для искомого тематически нейтрального объяснения нашего ментального опыта. Одна из первоочередных задач видится в поиске возможных точек пересечения рассмотренных в рамках настоящего исследования информационных концепций с другими философскими теориями и традициями (среди которых наиболее перспективными представляются различные виды процессуальных философий, а также некоторые версии так называемых «плоских онтологий», разрабатываемых в рамках современной континентальной традиции, потенциально применимых для изучения метафизики цифровых объектов), и, разумеется, с естественнонаучными дисциплинами. К условиям продуктивного взаимодействия философии с науками относится не только дальнейшее сопоставление результатов философского анализа и современных эмпирических исследований, но и критическое осмысление используемых в рамках данных наук концептуальных схем: очевидно, что до тех пор, пока не будет достигнут консенсус по вопросам, касающимся интерпретации таких основополагающих понятий, как «информация», «данные», «обработка», «вычислительная

реализация» и т.д., философские и естественнонаучные теории будут развиваться параллельно, никак не обогащая друг друга.

Как отмечают многие исследователи, развитие современных информационных теорий идет по двум альтернативным маршрутам: с одной стороны, интерпретации упомянутых выше понятий нередко подстраиваются под решения конкретных практических и технических задач; с другой стороны, данные термины все чаще приобретают крайне неопределенные значения и применяются для описания процессов, происходящих в различных сферах современного «информационного общества». По этой причине задача критического осмысления данных понятий сегодня стоит как никогда остро.

Несмотря на то, что применяемый некоторыми современными аналитическими философами структурно-каузальный подход к толкованию подобных понятий является далеко не единственным возможным вариантом интерпретации информационной терминологии, он представляется наглядной иллюстрацией того, в каком направлении развиваются информационные теории сознания после разоблачения компьютерной метафоры и выявления всех прочих проблем, связанных с полюбившейся многим исследователями грубой аналогией мозг-компьютер. Последовательная критика вычислительной установки способствовала отказу от безусловной веры в силу информационной обработки, действующей по принципу «черного ящика», магическим образом превращающего функциональную активность нейронов в феноменальные переживания. Подобный отказ побудил ученых и философов подойти критически к осмыслению объяснительного потенциала информационного концептуального аппарата и очистить когнитивные науки от необоснованных метафор и аналогий. Проект структурно-каузальной интерпретации информационных и вычислительных пространств является результатом подобного предельного очищения: трактуя вычислительные и информационные состояния достаточно широко, он предоставляет исследователям возможность начать строить новую теорию практически с чистого листа. Это, в частности,

объясняет, почему предложенная Чалмерсом стратегия описания сознания в информационных терминах еще очень далека от статуса полноценной теории.

На сегодняшний день достаточно сложно сказать, в какую сторону будут развиваться информационные теории сознания: весьма вероятно, что в дальнейшем будет представлено более строгое определение понятия «информация», а также будет сформулирован более четкий критерий вычислительной реализации, который, в свою очередь, будет дополнен последовательным концептуальным анализом, призванным ограничить спектр потенциально совместимых с данным подходом теорий сознания. Будучи соотнесенным с достижениями в области математических, компьютерных и биологических наук, информационный подход сулит обновление методологической программы столь популярных вычислительных когнитивных теорий и, возможно, решение проблемы сознания.

Библиография

На русском языке:

1. Августин. О Троице. Пер. А.А. Тащиана. Краснодар: Издательство ГЛАГОЛ. 2004.
2. Алексеев А.Ю., Ефимов А.Р., Финн В.К. Будущее искусственного интеллекта: тьюринговая или посттьюринговая методология? // Искусственные общества. М.: Издательство Федеральное государственное бюджетное учреждение науки Центральный экономико-математический институт РАН. 2019. Т. 14. № 4.
3. Барышников П.Н. Компьютерные метафоры в теориях сознания // Философские проблемы информационных технологий и киберпространства. 2011. № 2. С. 24-36.
4. Барышников П.Н. Феноменальное и вычислимое в структурах сознания // Вестник Российского университета дружбы народов. Серия: Философия. 2017. Т. 21. № 2. С. 229-239.
5. Беркли Дж. Алкифрон, или Мелкий философ. Работы разных лет. Перевод с английского, составление и комментарии А. А. Васильева. СПб.: Алетейя. 2000.
6. Бодякин В.И., Дубровский Д.И., Колин К.К., Лекторский В.А., Мелик-Гайказян И.В., Пружинин Б.И., Урсул А.Д. Информационный подход в междисциплинарной перспективе. Материалы круглого стола // Вопросы философии. 2010. №2. С. 84-113.
7. Братко А. А, Кочергин А.Н. Информация и психика. М.: Наука. 1977.
8. Васильев В.В. Трудная проблема сознания. М.: Прогресс-традиция. 2009.
9. Волков Д.Б. Бостонский зомби: Д. Деннет и его теория сознания. М.: Книжный дом «ЛИБРОКОМ». 2012.
10. Гаспарян Д. Э. Методология кибернетики второго порядка в применении к эпистемологическим проблемам философии сознания // Вестник Томского государственного университета. Философия. Социология. Политология. 2020. № 54. С. 5-18.

11. Гоббс Т. Сочинения в 2 т. Т.1. М. Мысль. 1989.
12. Дубровский Д. И. Информация, сознание, мозг. М. 1980.
13. Дубровский Д. И. Психические явления и мозг: философский анализ проблемы в связи с актуальными задачами нейрофизиологии, психологии и кибернетики. М. 1971.
14. Дубровский Д. И. Расшифровка кодов (Методологические аспекты проблемы) // Вопросы философии. 1979. № 12.
15. Дубровский Д. И. Сознание, мозг, искусственный интеллект. М. 2007.
16. Дубровский Д.И. Проблема «сознание и мозг»: информационный подход // Знание. Понимание. Умение. 2013. №4. С. 92-98.
17. Дубровский Д.И. Психические явления и мозг: философский анализ проблемы в связи с некоторыми актуальными задачами нейрофизиологии, психологии и кибернетики. М. Наука. 1971.
18. Дубровский Д.И. Сознание, мозг, искусственный интеллект: сб. статей. М. 2007.
19. Иванов Д.В. Природа феноменального сознания. М.: Книжный дом «ЛИБРОКОМ». 2013.
20. Ключева Н.Ю. Компьютерное моделирование интеллектуальных функций // Философия сознания: Аналитическая традиция. Третьи Грязновские чтения. Материалы Международной научной конференции (6-7 ноября 2009 г.). М.: Современные тетради. 2009.
21. Колин К.К. Философия информации и фундаментальные проблемы современной информатики // Alma mater (Вестник высшей школы). 2010. №1. С. 29-35.
22. Кузнецов А.В. Аргумент против конститутивного панпсихизма // Вестник Томского государственного университета. Философия. Социология. Политология. 2016. № 4. С. 60-68.
23. Люгер Д.Ф. Искусственный интеллект: стратегии и методы решения сложных проблем. М.: Издательский дом «Вильямс». 2003.

24. Марр Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов. М.: «Радио и связь». 1987.
25. Михайлов И.Ф. Концепции вычислений в современных науках о человеческом познании // Философские проблемы информационных технологий и киберпространства. 2018. № 1(14). С. 4–22.
26. Нагуманова С.Ф. Материализм и сознание: анализ дискуссии о природе сознания в современной аналитической философии. Казань. 2011.
27. Патнэм Х. Философия сознания. Пер. Макеевой Л.Б., Назаровой О.А., Никифорова А.Л. М.: Дом интеллектуальной книги. 1999.
28. Патнэм Х. Химера когнитивной науки // Тезисы докладов и выступлений IV Российского Философского конгресса (Москва, 24-28 мая 2005). Пер. В.В. Васильева. [Электронный ресурс]. URL: https://www.academia.edu/4091972/Споры_о_сознании_в_аналитической_философии_предисловие_к_публикации (дата последнего обращения: 06.04.2021).
29. Пинкер С. Как работает мозг. Пер. О.Ю. Семиной. М.: Кучково поле. 2017.
30. Плеханов И.А. Д.И. Дубровский и Э.В. Ильенков: неоконченный спор о природе психики и идеальном // Вестник Нижегородского университета им. Н.И. Лобачевского. 2007. №1(6). С. 153-162.
31. Рапопорт Г.Н., Герц А.Г. Искусственный и биологический интеллекты. Общность, общность структуры, эволюция и моделирование процессов познания. М.: КомКнига. 2005.
32. Редько В.Г. Эволюция, нейронные сети, интеллект: модели и концепции эволюционной кибернетики. М.: Книжный Дом "ЛИБРОКОМ". 2013.
33. Серл Д. Открывая сознание заново. М. 2002.
34. Сущин М.А. Байесовский разум: новая перспектива в когнитивной науке // Вопросы философии. 2017. №3. С. 77-88.
35. Урсул А.Д. Проблема информации в современной науке. М.: Наука. 1975.
36. Чалмерс Д. Сознательный ум. В поисках фундаментальной теории. Пер. В.В. Васильева. М.: УРСС: Книжный дом «ЛИБРИКОМ». 2013.

37.Юлина Н.С. Аналитическая философия в XX веке // Новейшие тенденции в аналитической философии. М.: ИФ РАН. 1985. С. 5-40.

38.Юлина Н.С. Головоломки проблемы сознания. Концепция Дэниела Деннета. М. 2004.

На английском языке:

1. Adams F. The Informational Turn in Philosophy // *Minds and Machines*. 2003. №13. P. 471–501.

2. Adams F., Aizawa K. Causal Theories of Mental Content. [Электронный ресурс] // *The Stanford Encyclopedia of Philosophy* (ed. Edward N. Zalta). 2017. URL: <https://plato.stanford.edu/archives/sum2017/entries/content-causal/> (дата последнего обращения: 12.12.19).

3. Adriaans P., Pet van Emde B. Computation, Information, and the Arrow of Time // *Computability in Context: Computation and Logic in the Real World* (ed. by B. Cooper, A. Sorbi). London: Imperial College Press. 2011. P. 1–17.

4. Adriaans P. Information // *The Stanford Encyclopedia of Philosophy* (N. Zalta ed.), 2019. [Электронный ресурс]. URL: <https://plato.stanford.edu/archives/spr2019/entries/information/> (дата последнего обращения: 09.05.2020).

5. Alistar M.C.I. Computational Thought from Descartes to Lovelace // *The Routledge Handbook of the Computational Mind* (ed. M. Spevak, M. Colombo). New York. 2019. P. 9-23.

6. Armstrong D.M. *A Materialist Theory of the Mind*. Routledge & Kegan Paul. 1968.

7. Atmaspacher H. 20th Century Variants of Dual-Aspect Thinking // *Mind&Matter*. 2014. №12(2). P. 245-288.

8. Baars B. *In the Theater of Consciousness: The Workspace of the Mind*. NY: Oxford University Press. 1997.

9. Bar-Hillel Y., Carnap R. An Outline of a Theory of Semantic Information // Technical Report. 1952. №247.
10. Benovsky J. Dual-Aspect Monism // Philosophical Investigations. 2016. № 39(4). P. 335-352.
11. Blamauer M. Panpsychism without Subjectivity? A Brief Commentary on Sam Coleman's «Mental Chemistry» and «The Real Combination Problem» // Disputatio. 2013. № 5(37). P. 299-309.
12. Block N. Searle's Arguments Against Cognitive Science // Views Into the Chinese Room: New Essays on Searle and Artificial Intelligence (eds. Preston, Bishop). Oxford University Press. 2003. P. 70-79.
13. Block N. Troubles with Functionalism // Minnesota Studies in the Philosophy of Science. 1978. №9. P. 261-325.
14. Block N., Fodor J. What Psychological States Are Not // The Philosophical Review. 1972. №81. P. 159–181.
15. Boden M.A. Information, Computation and Cognitive Science // Handbook of the Philosophy of Science. Vol.8: Philosophy of Information (ed. by P. Adrians, J. van Benthem). Amsterdam: Elsevier. 2008. P. 741-763.
16. Boden M.A. Mind as Machine: A History of Cognitive Science. Oxford University Press. 2006.
17. Boole G. An Investigation of the Laws of Thought. London: Walton and Maberly. 1854.
18. Brey P., Soreker J. Philosophy of Computing and Information Technology // Handbook for Philosophy of Science. (ed. A. Meijers, gen. ed. D. Gabbay, P. Thagard, J. Woods). 2009. №14.
19. Capurro R., Birger H. The Concept of Information // Annual Review of Information Science and Technology (ed. by B. Cronin). 2003. №37. P. 343–411.
20. Cerullo M. Integrated Information theory: A promising but ultimately incomplete theory of consciousness // Journal of Consciousness Studies. 2011. №18. P. 45-58.
21. Chalmers D. J. Constructing the World. Oxford University Press. 2012.

22. Chalmers D.J. A Computational Foundation for the Study of Cognition // *The Journal of Cognitive Science*. 2011. №12. P. 323–357.
23. Chalmers D.J. Absent qualia, fading qualia, dancing qualia // *Conscious Experience*. (ed. by T. Metzinger). Imprint Academic.1995. P. 309-328.
24. Chalmers D.J. Syntactic transformations on distributed representations // *Connection Science*. 1990. №2. P. 53-62.
25. Chalmers D.J. *The Consciousness Mind: In Search of a Fundamental Theory*. NY: Oxford University Press. 1996.
26. Chalmers D.J. Computing the Thinkable // *Behavioral and Brain Sciences*. 1990. №13. P. 658-659.
27. Chalmers D.J. *Consciousness and Cognition*. 1990. [Электронный ресурс] // Персональный сайт Д. Чалмерса. URL: <http://consc.net/papers/c-and-c.html> (дата последнего обращения: 11.04.2020).
28. Chalmers D.J. Does a Rock Implement Every Finite State Automaton? // *Synthese*. 1996. №108. P. 309–333.
29. Chalmers D. J. Facing up to the Problem of Consciousness // *Journal of Consciousness Studies*. 1995. №2(3). P. 200-219.
30. Chalmers D.J. Mind, Pattern, Information. 1988 [Электронный ресурс] // Персональный сайт Д. Чалмерса. URL: <http://consc.net/papers/oxford2.pdf> (дата последнего обращения: 10.04.2020).
31. Chalmers D.J. Minds, Machines, and Mathematics // *Psyche*. 1996. №2. P. 11–20.
32. Chalmers D.J. On Implementing a Computation // *Minds and Machines*. 1995. №4. P. 391–402.
33. Chalmers D.J. Structuralism as a Response to Skepticism // *Journal of Philosophy*. 2018. №115(12). P. 625-660.
34. Chalmers D.J. The Combination Problem for Panpsychism. 2017. [Электронный ресурс] // Персональный сайт Д. Чалмерса. URL: <http://consc.net/papers/combination.pdf> (дата последнего обращения: 12.04.2020).
35. Chalmers D.J. *The Matrix As Metaphysics* // *Philosophers Explore the Matrix* (ed. by C. Grau). Oxford University Press. 2005. P. 132-169.

36. Chalmers D.J. The Varieties of Computation: A Reply // *The Journal of Cognitive Science*. 2012. №13. P. 213–248.
37. Chalmers D.J. The Virtual and the Real. // *Disputatio*. 2017. № 9(46). P. 309-352.
38. Chalmers D.J. Why Fodor and Pylyshyn Were Wrong: The Simplest Refutation // *Philosophical Psychology*. 1993. №63. P. 305–319.
39. Chalmers. D.J. Panpsychism and Panprotopsychism. 2013. [Электронный ресурс] // Персональный сайт Д. Чалмерса. URL: <http://consc.net/papers/panpsychism.pdf> (дата последнего обращения: 12.04.2020).
40. Chomsky N. Review of Verbal Behavior // *Language*. 1959. №35. P. 26–58.
41. Churchland P.M., Sejnowski T.J. *The Computational Brain*. Cambridge, MA: MIT Press. 1992.
42. Churchland P.S. *Neurophilosophy*, Cambridge: MIT Press. 1986.
43. Churchland P.S., Koch C., Sejnowski T. What Is Computational Neuroscience? // *Computational Neuroscience*, E. Schwartz (ed.), Cambridge: MIT Press. 1990.
44. Cole D. The Chinese Room Argument // *The Stanford Encyclopedia of Philosophy* (ed. N. Zalta). 2020. [Электронный ресурс] URL: <https://plato.stanford.edu/archives/win2020/entries/chinese-room> (дата последнего обращения: 16.12.2020).
45. Coleman S. The Real Combination Problem: Panpsychism, Micro-Subjects, and Emergence // *Erkenntnis*. 2013. №(1). P. 1-26.
46. Copeland J. What is Computation? // *Synthese*. 1996. №108. P. 335–359.
47. Daiseanu F. Informational Model of Consciousness: From Philosophic Concepts to an Information Science of Consciousness // *Philosophy Study*. 2019. №9(4). P. 181-196.
48. Dehaene S. *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. New York. 2014.
49. Dennet D.C. The logical geography of computational approaches: A view from the east pole // *The Representation of Knowledge and Belief* (eds. M. Brand, R. M. Harnish). University of Arizona Press. 1986. P. 59-78.

50. Dennett D.C. *Brainstorms: Philosophical Essays on Mind and Psychology*. MIT Press. 1981.
51. Dennett D.C. *Consciousness explained*. Boston, MA. 1991.
52. Dennett D.C. *Content and consciousness*. N.Y. 1969.
53. Dennett D.C. *From Bacteria to Bach and Back: The Evolution of Minds*. NY: W.W. Norton Company. 2017.
54. Dennett D.C. *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. MIT Press. 2005.
55. Deutsch D. *The Fabric of Reality*. New York: Allan Lane. 1997.
56. Dretske F. *Dretske and His Critics*. Cambridge: Blackwell. 1991.
57. Dretske F. *Explaining Behavior: Reasons in a World of Causes*. Cambridge, Massachusetts The MIT Press. 1988.
58. Dretske F. *Knowledge and the Flow of Information*. MA: Massachusetts Institute of Technology. 1981.
59. Dretske F. *Naturalizing the Mind*. Cambridge, MA: The MIT Press. 1995.
60. Dreyfus H. *What Computers Can't Do*. Harper&Row. 1972.
61. Edelman S. *Computing the Mind. How the Mind Really Works*. Cornell University Press. 2007.
62. Floridi L. A Defence of Informational Structural Realism. // *Synthese*. 2008. №161(2). P. 219-253.
63. Floridi L. Against digital ontology // *Synthese*. 2009. №168 (1). P. 151-178.
64. Floridi L. Informational Realism // [Электронный ресурс]: ResearchGate.net. URL: https://www.researchgate.net/publication/262350693_Informational_Realism (дата последнего обращения: 07.05.2020).
65. Floridi L. *The Philosophy of Information*. Oxford University Press. 2011.
66. Floridi L. Trends in the Philosophy of Information // *Handbook of philosophy of information* (ed. by P. Adriaans and V. Benthem). Amsterdam; London, North-Holland. 2008. P. 113-131.
67. Fodor J. *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press. 1990.

- 68.Fodor J. Information and Representation // Information, Language, and Cognition (ed. P. Hanson). Vancouver: University of British Columbia Press. 1990. P. 175–190.
- 69.Fodor J. Psychosemantics: The Problem of Meaning in the Philosophy of Mind. Cambridge, MA: MIT/Bradford. 1987.
- 70.Fodor J. Representations. Philosophical Essays on the Foundations of Cognitive Science. Harvester Press. 1981. P. 225-257.
- 71.Fodor J. The Mind Doesn't Work This Way. The Scope and Limits of Computational Psychology. MIT Press. 2000.
- 72.Fodor J.A. The Language of Thought. Cambridge, MA: Harvard University Press. 1975.
- 73.Fodor J.A., Pylyshyn Z. Connectionism and cognitive architecture: A critical analysis // Cognition. 1988. №28. P. 3-71.
- 74.Frankish K. Illusionism as a Theory of Consciousness //Journal of Consciousness Studies. 2016. №23 (11-12). P. 11-39.
- 75.Fredkin E. Introduction to Digital Philosophy // International Journal of Theoretical Physics. 2003. №42(2). P. 189-247.
- 76.Fresco N. Objective Computation Versus Subjective Computation // Erkenntnis. 2015. № 80 (5). P. 1031-1053.
- 77.Gabora L. Amplifying Phenomenal Information: Toward a Fundamental Theory of Consciousness // Journal of Consciousness Studies. 2002. №9(8).P. 3-29.
- 78.Gamez D. Are Information or Data Patterns Correlated with Consciousness? // Topoi. 2016. №35(1). P. 225-239.
- 79.Gleick J. The Information: A History, A Theory, A Flood. NY: Pantheon Books. 2011
- 80.Godfrey-Smith P. Triviality Argument against Functionalism // Philosophical Studies. 2009. № 145. P. 273-295.
- 81.Graziano M. S.A. Rethinking Consciousness: A Scientific Theory of Subjective Experience. W.W. Norton Company. 2019.
- 82.Grice H.P. Meaning // The Philosophical Review. 1957. №66. P. 377–388.

- 83.Hardcastle V.G. A Critique of Information Processing Theories of Consciousness // Minds and Machines. 1995. №5(1). P. 89-107.
- 84.Harms W.F. The use of information Theory in Epistemology // Philosophy of Science. 1998. №65. P. 472-501.
- 85.Haugeland J. Artificial Intelligence: The Very Idea. Cambridge, MA: MIT Press. 1985.
- 86.Haugeland J. Syntax, Semantics, Physics // Views Into the Chinese Room: New Essays on Searle and Artificial Intelligence (eds. Preston and Bishop). 2002. P. 379–392.
- 87.Hinton G.E., McClelland J.L., Rumelhart D.E. Distributed representations // Parallel Distributed Processing (Rumelhart D.E., McClelland J.L. & PDP Research Group eds.). Cambridge, MA: MIT Press. 1986. P. 77-109.
- 88.Horst S.W. Symbols, Computation and Intentionality. A Critique of the Computational Theory of Mind. University of California Press, Berkeley and Los Angeles. 1996.
- 89.Hutto D., Myin E. Radicalizing Enactivism: Basic Minds without Content. Cambridge, MA: MIT Press, 2013.
- 90.Jackendoff R. Consciousness and the Computational Mind. Cambridge, Massachusetts: MIT Press. 1987.
- 91.Kay L.E. Who Wrote the Book of Life? A History of the Genetic Code. Stanford University Press. 2000.
- 92.Kripke S. Naming and Necessity. Oxford. 1980.
- 93.Ladyman J. What Does it Mean to Say that a Physical System Implements a Computation? // Theoretical Computer Science. 2009. №410. P. 376–383.
- 94.Le Bihan B. Aspects in Dual-Aspect Monism and Panpsychism: A Rejoinder to Benovsky // Philosophical Investigations. 2019. №42(2). P.186-201.
- 95.Levin J. Functionalism and the Argument from Conceivability // Canadian Journal of Philosophy, 1985. №11. P. 85–104.
- 96.Lewis, D. Lucas Against Mechanism // Philosophy. 1969. №44. P. 231–243.

97. Lloyd S. *Programming the Universe: A Quantum Computer Scientist Takes On the Cosmos*. Knopf, 2006.
98. Loewer B. *From Information to Intentionality // Synthese*. 1987. №70. P. 287-317.
99. Lucas J.R. *Minds, Machines, and Gödel // Philosophy*. 1961. №36. P. 112–137.
100. Lycan W. *Representational Theories of Consciousness // The Stanford Encyclopedia of Philosophy* (ed. by E.N. Zalta). 2019. [Электронный ресурс] // plato.stanford.edu.
URL: <https://plato.stanford.edu/archives/fall2019/entries/consciousness-representational/> (дата последнего обращения: 18.10.2021).
101. MacKay D. M. *Information, Mechanism, and Meaning*. Cambridge: MIT Press. 1970.
102. Maleeh R. *Pragmatic Information as a Unifying Biological Concept // Information*. 2014. №5. P. 451-478.
103. Marr D. *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman and Company. 1982.
104. McCulloch W., Pitts W. *A Logical Calculus of Ideas Immanent in Nervous Activity*. *Bulletin of Mathematical Biophysics*. 1943. №5. P.115-133.
105. Milkowski M. *Explaining the Computational Mind*. Cambridge: MIT Press. 2013.
106. Millikan R. *What Has Natural Information to Do with Intentional Representation? // Naturalism, Evolution and Mind* (ed. D.M. Walsh). Cambridge: Cambridge University Press. 2001. P. 105-125;
107. Mørch H. H. *Panpsychism and Causation: A New Argument and a Solution to the Combination Problem*. Ph.D. thesis, University of Oslo, 2014. [Электронный ресурс] // philpapers.org. URL: <https://philpapers.org/archive/HASPAC-2.pdf> (дата последнего обращения: 26.04.2020).
108. Nagel T. *The View from Nowhere*. NY: Oxford University Press. 1986.
109. Navon D. *The Function of Consciousness or of Information? // Behavioral and Brain Sciences*, №14(4). 1991.

110. Neander K. A Mark of the Mental: In Defence of Informational Teleosemantics. Cambridge, MIT Press. 2017.
111. Newell, A., Simon H. The Logic Theory Machine: A Complex Information Processing System // IRE Transactions on Information Theory. 1956. №3. P. 61–79.
112. Oizumi M., Albantakis L., Tononi G. From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0. // PLoS Computational Biology. 2014. №10. P. 23-57.
113. Penrose R. The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics. Oxford: Oxford University Press. 1989.
114. Piccinini G. Computing Mechanisms // Philosophy of Science. 2007. №74. P. 501–526.
115. Piccinini G. Physical Computation: A Mechanistic Account. Oxford University Press. 2015.
116. Piccinini G. The Resilience of Computationalism // Philosophy of Science. 2010. №77. P. 852–861.
117. Piccinini G., Scarantino A. Computation vs. Information processing: Why their Difference Matters to Cognitive Science // Studies in History and Philosophy of Science. 2010. №41. P. 237–246.
118. Pinker S. How the Mind Works. New York: Norton. 1997.
119. Pinker S. So How Does the Mind Work? // Mind & Language. 2005. №1 (20). P. 1–24.
120. Pockett S. Problems with theories that equate consciousness with information or information processing // Frontiers in Systems Neuroscience. 2014. №8(255).
121. Pratt V. Thinking Machines. The Evolution of Artificial Intelligence. Oxford: Basil Blackwell. 1987.
122. Putnam H. A Half Century of Philosophy, Viewed From Within // Daedalus. 1997. №1. P. 175-208.
123. Putnam H. The Best of All Possible Brains? // The New York Times. 1994. №20. P.7-17.

124. Putnam. H. Mind, Language, and Reality. Cambridge: Cambridge University Press. 1975;
125. Pylyshyn Z. Computation and Cognition. Toward a Foundation for Cognitive Science. Cambridge, Massachusetts: Bradford Books/MIT Press. 1984.
126. Rescorla M. A Theory of Computational Implementation // Synthese. 2014. №191. P. 1277–1307.
127. Rescorla M. The Computational Theory of Mind // The Stanford Encyclopedia of Philosophy (N. Zalta ed.). 2019. [Электронный ресурс]. URL: <https://plato.stanford.edu/entries/computational-mind/#AltConCom> (дата последнего обращения: 04.12.2020).
128. Rosenberg G.H. A Place for Consciousness: Probing the Deep Structure of the Natural World. Oxford University Press. 2004.
129. Rosenberg G.H. Causality and the combination problem // Consciousness in the Physical World: Perspectives on Russellian Monism (T. Alter, Y. Nagasawa, eds.). Oxford University Press. 2014. P. 224-245.
130. Rosenbluelh A., Wiener N., Bigelow J. Behavior, Purpose and Teleology // Philosophy of Science. Baltimore. 1943. № 10. P. 18–24.
131. Sayre K.M. Consciousness: A Philosophic Study of Minds and Machines. Random House. 1969.
132. Sayre K.M. Cybernetics and the Philosophy of Mind. Routledge and Kegan Paul. 1976.
133. Sayre K.M. Intentionality and information processing: An alternative model for cognitive science // Behavioral and Brain Sciences. 1986. №9(1). P.121-138.
134. Schweizer P. Triviality Arguments Reconsidered. Minds & Machines. 2019. №29. P. 287–308.
135. Seager W. A New Idea Of Reality: Pauli on the Unity of Mind and Matter // Mind and Matter. 2011. № 9(1). P. 37-52.
136. Seager W. Consciousness, Information and Panpsychism // Journal of Consciousness Studies. 1996. №2(3). P. 272-88.

137. Searle J. Is the Brain a Digital Computer? // Proceedings and Addresses of the American Philosophical Association. 1990. №64. P. 21–37.
138. Searle J. Minds, Brains, and Programs // Behavioral and Brain Sciences. 1980. №3. P. 417–457
139. Searle J. The Chinese Room // The MIT Encyclopedia of the Cognitive Sciences (eds. R.A. Wilson, F. Keil), Cambridge, MA: MIT Press. 1999.
140. Searle J. The Rediscovery of the Mind. Cambridge, MA: MIT Press. 1992.
141. Searle J.R. Why I Am Not a Property Dualist // Journal of Consciousness Studies. 2002. №9. P. 57-64.
142. Searle J.R., Can Information Theory Explain Consciousness? [Электронный ресурс] // Hhimwich. 2013. URL: <https://www.hhimwich.com/files/can-information-theory-explain-consciousness.docx> (последнее обращение 28.03.2019).
143. Shallice T. Information-processing models of consciousness: Possibilities and problems // Consciousness in Contemporary Science (eds. J. Anthony, M. Bisiach). Oxford University Press. 1988.
144. Shannon C. E., Weaver W. The Mathematical Theory of Communication. University of Illinois Press. 1963.
145. Skrbina D. Dualism, Dual-Aspectism and the Mind // Contemporary Dualism: A Defence (ed. A. Lavazza, H. Robinson). NY: Routledge. 2014. P. 220-245.
146. Sloman A. The Computer Revolution in Philosophy: Philosophy, science and models of mind. Harvester press and Humanities press. 1978.
147. Smolensky P. On the Proper Treatment of Connectionism // Behavioral and Brain Sciences. 1988. №11. P. 1–74.
148. Sprevak M. Triviality Arguments About Computational Implementation // The Routledge Handbook of The Computational Mind (ed. by M. Sprevak, M. Colombo). NY: Taylor&francis Group. 2019. P. 175-192.
149. Stoljar D. Chalmers v. Chalmers // Noûs. 2020. №54(2). P. 469-487.
150. Stubenberg L. Neutral Monism // The Stanford Encyclopedia Tononi G. Consciousness as Integrated Information: A provisional manifesto // Biological Bulletin. 2008. №7. P.216-242.

151. Turing A. M. Computing Machinery and Intelligence // Mind. 1959. №49. P. 433-460.
152. Usher M. A Statistical Referential Theory of Content: Using Information Theory to Account for Misrepresentation // Mind and Language. 2001. №16. P. 311-334.
153. Varela F. J., Thompson E., Rosch E. The Embodied Mind: Cognitive Science and Human Experience. Cambridge, MA: The MIT Press. 1991.
154. Varela F. J. Present time consciousness. Journal of Consciousness Studies. 1999. №6(2-3). P. 111-140.
155. Velmans M. Is human information processing conscious? // Behavioral and Brain Sciences. 1991. №14(4). P. 651-669.
156. Velmans M. Reflexive monism // Journal of Consciousness Studies. 2008. №15(2). P. 5-50.
157. Von Neumann J. The Computer And The Brain. New Haven: Yale University Press. 1958.
158. Wiener N. Cybernetics or Control and Communication in the Animal and the Machine. John Wiley & Sons Inc., New York. 1948.
159. Williams M.R. A History of Computing Technology. CA: Computer Society Press. 1997.
160. Zuse K. Calculating Space. Cambridge, MA: MIT Press. 1970.