

ОТЗЫВ
официального оппонента на диссертацию
Ефимова Альберта Рувимовича
«Философско-методологические основы посттьюринговой
интеллектуальной робототехники»,
представленную на соискание ученой степени кандидата философских
наук по специальности 09.00.08. – философия науки и техники

Диссертация соответствует специальности и отрасли науки, по которой она представлена.

Актуальность темы диссертации определяется масштабной интенсификацией исследований и практических разработок по проблематике искусственного интеллекта в целом и интеллектуальной робототехники в частности. Это видно, в частности, по взрывному росту количества публикаций, патентов и инвестиций в тематику. Указанный процесс сопровождается интенсификацией алармистских настроений на тему «бунта машин», «Четвертой мировой войны искусственного интеллекта против человека», подогреваемых такими произведениями массовой культуры, как сериал ВВС «Черное зеркало» и т.д. Наряду с этим формируется крайне серьезная повестка цифрового регулирования, нормативной и правовой базы, наконец, новой цифровой морали и этики. Все это показывает абсолютную необходимость ускорения процесса формирования именно философского фундамента данной области.

В названии своей работы автор дал отсылку к гениальному математику и философу Алану Тьюрингу, который завершил формализацию рациональной, алгоритмической деятельности человека в 1936 г., и поставил проблему соотношения искусственного и естественного интеллекта в 1950-ых гг. Наряду с этим, художественные пророчества Карела Чапека (с его роботами R.U.R.) 1920-ых гг. и «Законы робототехники» Айзека Азимова 1940-ых гг. задают систему координат, в которой разворачивается исследование.

За период, прошедший с момента создания самого термина «искусственный интеллект» Дж. Маккарти и его соратниками по Дартмутскому семинару 1956 г., жизнь человечества изменилась коренным образом. При этом, и в теоретической проработке, и в практической реализации идеологии ИИ существенную роль сыграли отечественные исследователи: А. И. Берг, Д.А. Поспелов, М.М. Бонгард, В. И. Варшавский, М.Л. Цетлин и многие другие.

В этих условиях диссертация А.Р. Ефимова охватывает ряд принципиально новых аспектов и подходов заявленной проблематики. Самым важным из них, пожалуй, является

сочетание всестороннего рассмотрения когнитивного аспекта ИИ и его телесного, физического аспекта, сконцентрированного, подчеркну, в философской концепции интеллектуального робота, развиваемой и практически реализуемой автором.

Диссертация состоит из трех глав, дополненных введением, заключением и списком литературы.

Введение обрисовывает контекст исследования и содержит краткое изложение его основных результатов.

Первая глава диссертации посвящена обзору и анализу современных исследований в интеллектуальной робототехнике и уточнению понятий, включая ключевое понятие «умного», интеллектуального робота. В данной области некорректно говорить об исчерпывающем обзоре, однако автору удалось сделать действительно представительную выборку, составившую 199 источников; ко многим ключевым текстам читателя отсылают также подстраничные примечания. Отметим, что А.Р. Ефимов не ограничивается только публикациями последних лет, он устанавливает их преемственность с классическими работами. Критическая глубина предложенного в диссертации обзора области далеко превосходит обычные литературные обзоры кандидатских и докторских диссертаций. Этот обзор может рассматриваться как, в определенной степени, синтезирующий идеи предшественников и бросающий новый свет на данную проблематику.

Во второй главе автор анализирует исходный подход Тьюринга и выявляет ряд его принципиальных особенностей и ограничений. Особую роль играет глубокий анализ, проведенный соискателем в одной из наиболее знаменитых, противоречивых и символических областей философии ИИ – «Тесте Тьюринга». Собранный автором уникальный материал потребовал систематизации, осмысления и подъема на мета-уровень. Автор использовал введенное А.Ю. Алексеевым различие частного и комплексного теста Тьюринга. Результатом этой работы А.Р. Ефимова стало построение им «посттьюринговой методологии». На ее основе он вводит системы диалектических оппозиций «*виртуальное-физическое* и *вербальное-невербальное*». В этой «системе координат» автор располагает рассмотренные им 18 видов теста Тьюринга. Ключевым для «посттьюринговой методологии» Ефимова становится понятие техно-умвелта. Оно восходит к прорывным работам Якоба Иоганна Иксуля – выдающегося эстонско–русско–немецкого исследователя, заложившего основу для современных представлений биосемиотики и этологии, в какой-то степени параллельных учению о ноосфере В.И. Вернадского и дополняющих его. Замечательным достижением диссертанта является перенос концепции

Иксуля в техносферу и обогащение ее развиваемыми Ефимовым представлениями о технологии интеллектуальных роботов. Результатом является осмысление проблемы Общего искусственного интеллекта, дающее надежду на принципиальные продвижения в этой проблеме. Посттюринговая методология и понятие техно-умвельта позволяет автору найти подход к снятию эпистемологического ограничения робота, устранению «стены Тьюринга», разделяющей наблюдателя-человека и робота в одноименном тесте, и, что конечно, важнее – в окружающем нас сегодня цифровом мире.

В третьей главе диссертации предложенная автором методология и система понятий получают дальнейшее развитие и приложения в практических вопросах робототехники. Построенная А.Р. Ефимовым система четырех квадрантов техно-умвельта дает ему единый подход к построению прикладных систем интеллектуальных роботов различного назначения и движению к Общему искусственному интеллекту. Как уже отмечалось в Главе 1, интеллектуальная робототехника играет ту же концептуальную роль для Общего искусственного интеллекта, что человеческое тело играет для естественного интеллекта.

Особое внимание, в том числе – для будущих работ, заслуживают результаты Ефимова, относящиеся к сообществам интеллектуальных роботов и к системам, гармонично сочетающим совместно действующих человека и робота.

Достоверность результатов диссертации подтверждается их логической согласованностью и разносторонней апробацией, в том числе – многочисленными своевременными публикациями в авторитетных изданиях, полно отражающих содержание диссертации.

От философской диссертации не принято ожидать непосредственных практических применений. Обычно речь идет, в основном, об использовании в учебных курсах. И действительно, материал, представленный в диссертации, должен войти в курсы по философским вопросам искусственного интеллекта. В частности, планируется, что он будет представлен в курсе «Основы цифровых технологий и искусственного интеллекта» в рамках Программы магистратуры школы «Мозг, когнитивные исследования и искусственный интеллект МГУ им. М.В. Ломоносова. Также стоит отметить, что идеи, изложенные в диссертации, включены в коллективную монографию Сбербанка «Сильный искусственный интеллект: на подступах к сверхразуму».

При этом можно выразить уверенность, что практическая ценность рассматриваемой диссертации не ограничится учебными курсами. Это обусловлено отчасти спецификой самой отрасли, где тесно переплетены вопросы человеческого сознания и цифровых

технологий, отчасти новаторским характером самой работы. А.Р. Ефимов смело вторгается своими философскими построениями в технологические решения возглавляемых им разработок в области искусственного интеллекта и робототехники. Полученные результаты подробно представлены в третьей главе диссертации. Есть и еще одно направление, представляющееся весьма перспективным. Это – интерпретация концепции умвельта в результативном образовании на цифровой платформе. С одной стороны – это умвельт людей – участников образовательного процесса, с другой – это техно-умвельт платформы, содержащей внутри себя умвельт школы, целостно взаимодействующей с сообществом учащихся, для каждого учащегося учитывающий его учебные цели, режим дня и время затрачиваемое на выполнение заданий, питание и т.д.

Диссертация не свободна от недостатков, которые, при этом, не влияют на ее целостную оценку, приведенную ниже. Укажем основные из них.

Наименее обоснованными и существенными бывают претензии к диссертационным исследованиям, высказанные, исходя из того, чего в диссертации нет. Тем не менее, не могу не указать на то, что при общем скудном представлении в диссертации отечественных исследований (например, упомянуты А.И. Берг и А.Н. Колмогоров, но нет ссылок на их работы), что может иметь свои объяснения, ничего не говорится о работах Дмитрия Александровича Поспелова, которого часто называют «отцом отечественного искусственного интеллекта». Именно Д.А. Поспелов ввел в научный обиход понятие «гиромат» (термин был взят им у С. Лема, цитируемого диссертантом). Это понятие имеет существенные общие элементы с построениями А.Р. Ефимова.

Ключевое для диссертации понятие техно-умвельт вводится в ней косвенным образом, это затрудняет его восприятие читателем.

Одной из центральных фигур области, безусловно значительной в мировом масштабе, является Алан Тьюринг. В тексте упоминается, что он возглавлял Ratio Club – важный британский кибернетический кружок конца сороковых – пятидесятих годов прошлого века, в которых входили многие блестящие умы Великобритании и, до некоторой степени, США. Однако источников информации о Тьюринге именно, как о лидере, не указано. Утверждение о его лидерстве остается необоснованным. Есть и еще ряд мелких неточностей, относящихся к роли Тьюринга и его результатам.

На Рис. 1. «Пространство тестов Тьюринга» (с. 92 диссертации), играющем принципиальную роль в исследовании, графически представлены варианты «Теста». Однако состав тестов на рисунке отличается от их состава в предшествующем тексте.

При обсуждении Э.ЛЕНЫ, или где-нибудь еще в тексте, напрашивается ссылка на Элизу (*ELIZA*) Дж. Вейценбаума. Такой ссылки в работе нет. Также можно отметить, что убедительно выявленная автором контекстная неотличимость Э.ЛЕНЫ от человека не соответствует полностью идеологии даже частичного теста Тьюринга, и представляет интерес, прежде всего с точки зрения выхода за пределы этого теста и преодоления «стены».

В тексте имеются опечатки и другие речевые ошибки, например, отклонения от принятого русского написания иноязычных фамилий, что «в пределах нормы» для кандидатских работ.

Мы уже упоминали большую библиографическую работу, сделанную автором. Эта работа была бы еще более полезной, если бы ссылки на русские переводы изданий сопровождалась указанием на оригинальные тексты. Это важно, например, с точки зрения хронологии и преемственности исследований.

Подытоживая сказанное, можно констатировать, что диссертация А.Р. Ефимова представляет собой законченное исследование, обладающее высокой степенью актуальности и научной новизны, существенно продвигающее методологию работ в области искусственного интеллекта, практически значимое.

Рецензируемая работа имеет высокую научную и практическую значимость и отвечает требованиям, установленным Министерством науки и высшего образования Российской Федерации. Проведенное исследование говорит о глубоком понимании автором темы и его высокой квалификации. По своему содержанию, структуре, завершенности, обоснованию теоретических положений и практического использования результатов диссертация полностью соответствует требованиям п.п. 9 и 10 Положения о присуждении ученых степеней, утвержденного Постановлением Правительства Российской Федерации от 23 сентября 2013 г. № 842, а ее автор – Ефимов Альберт Рувимович – заслуживает присуждения ученой степени кандидата философских наук по специальности 09.00.08 – философия науки и техники.

Доктор физико-математических наук,
профессор,
академик РАН и РАО
Зав. кафедрой математической логики
и теории алгоритмов МГУ им. М. В. Ломоносова

Семенов Алексей Львович

27 апреля 2021 г.

Погнис

уговорено
(А.И. Шафаревич)