



3. Мифы о «поколении Z» / Н. В. Богачева, Е. В. Сивак; Национальный исследовательский университет «Высшая школа экономики», Институт образования. — М.: НИУ ВШЭ, 2019. С. 33.

4. Дидактическая концепция цифрового профессионального образования и обучения / П. Н. Биленко, В. И. Блинов, М. В. Дулинов, Е. Ю. Есенина, А. М. Кондаков, И. С. Сергеев; под науч. ред. В. И. Блинова – М.: Изд. «Перо», 2019. С. 20, 22.

5. Там же. С. 38.

6. Нечаев В.Д., Дурнева Е.Е. «Цифровое поколение»: психолого-педагогическое исследование проблемы // Педагогика. 2016. №1. С. 36-45.

Е.С. Головина

НЕОБХОДИМЫЕ ОГРАНИЧЕНИЯ ДЛЯ СИСТЕМ АВТОМАТИЗИРОВАННОГО УПРАВЛЕНИЯ ОПАСНЫМ ПРОИЗВОДСТВОМ

(Самарский университет)

Введение

С развитием цифровизации и автоматизации всё большее количество процессов переходит в цифровую среду. Ученые и промышленники стараются переложить всё более сложные задачи на плечи машин, в том числе частично передавая «ключевые точки принятия решений». Широко распространена практика внедрения систем автоматизированного управления различных уровней, позволяющих выполнять очень большой спектр действий в пределах заданных диапазонов и в рамках поставленных задач. Таким образом, системы управления действуют в рамках процессов в одном заданном качественном состоянии.

В рамках исследования под «итоговой точкой принятия решения» понимается такая точка процесса, после которой изменяется качественное состояние управляемой системы или системы, на которую она воздействует (например, остановка производства, запуск ядерной ракеты, отключение пациента от системы жизнеобеспечения). В конечном итоге в ряде случаев от принятия решения в такой итоговой точке зависит жизнь человека, группы людей или всего человечества. В сферах повышенной ответственности и опасности – нефтепереработка, химическая промышленность, врачебные дисциплины и прочее – итоговая точка принятия решения так и не передана машине. Корни этого решения видятся не в простом опасении или невозможности создать программу управления таким образом, чтобы она учитывала все бесконечное разнообразие факторов, но в наличии значимых и трудно прогнозируемых рисков.

В статье вводится определение кибернетики четвертого рода и рассматривается ряд методологических и философских предпосылок,



обусловливающих необходимость установления ограничений для метасубъектов при развитии кибернетики четвертого рода, дается обзор вариантов ограничений, которые требуется учитывать при разработке современных киберфизических систем.

Рассматривается вопрос построения философско-этических принципов работы киберфизических систем. Учитывается, что решение человека напрямую связано с его этическими особенностями, в то время как перенос таких ограничений в киберфизическую систему связан с рядом сложностей.

Кибернетика четвертого порядка и ограничения к ней

В.Е. Лепский ввел понятие кибернетики третьего порядка, характеризующейся базовой парадигмой «субъект – метасубъект (саморазвивающаяся полисубъектная среда)» [1], «в которой объект и система управления сливаются в единое целое в саморазвивающейся рефлексивно-активной среде». Значительное количество разработок, касающихся современных систем управления производствами, могут быть отнесены к кибернетике третьего порядка. Наряду с этим продолжается обсуждение и попытки разработки сложных систем, «на плечи» которых планируется переложить ответственность за принятие решений более серьезного характера – «итоговых точек принятия решений». Продолжая рассуждения В.Е. Лепского касательно парадигм и философских оснований становления кибернетики третьего порядка, можно ввести понятие кибернетики четвертого порядка, когда взаимодействие идет между метасубъектами, а процесс саморазвития среды дополняется функцией целеполагания. В научном сообществе данный вопрос так или иначе обсуждается, в том числе в свете трансформации жизненного мира человека в «киберумвельт» - смешанный жизненный мир физически-цифровой или киберфизической действительности [2].

Под киберфизическими системами в общем смысле понимают такие системы, которые соединяют физические процессы производства или иные процессы, требующие практической реализации непрерывного управления в режиме реального времени, с программно-электронными системами [3, 4]. Киберфизическая система кибернетики третьего порядка имеет своей задачей удержать систему в заданном состоянии или перевести систему в требуемое состояние. При этом целевое состояние системы задается человеком. В концепции кибернетики четвертого рода система «принимает решение», в какое состояние ей необходимо перейти. Это подразумевает, что система формирует свою «мета-цель», и здесь возникает вопрос – насколько данное решение формируется человеком, а насколько – самой системой через призму построений с опосредованным влиянием человека?

Проблему принятия ключевых решений киберфизической системой следует рассматривать с позиции целеполагания: цель, стоящая перед человеком при развитии цифровизации – максимизация помощи человеку, «техника как исполнение человеческих целей посредством оформления задач» [2]. Передача ключевых решений машине подразумевает замену целей человека



целеполаганием машины: человек строит «между собой и Природой цепь из звеньев, в которой каждое последующее звено будет как усилитель Разума более мощным, чем предыдущее», оставляя первичное целеполагание за человеческим разумом [3]. Целеполагание машины может быть понимаемо лишь в том объеме, который ей дал в качестве установок человек, оставляя за скобками «личностное, скрытое знание, неразрывно связанное с субъектами, их порождающими» [4]. Однако при передаче функции принятия решений киберфизической системе, она превращается из объекта в субъект, а учитывая интегративные тенденции, принятые ей решения и совершаемые действия, могут затрагивать фактически каждого индивида [5].

«Эксклюзивные» точки принятия решения

Наличие функции целеполагания в киберфизической системе подразумевает, в том числе, что система может «выбрать» движение не по назначенной заранее программе действий, но и перейти к изменению этой программы в зависимости от факторов и внешних воздействий, получаемых системой извне. Алгоритмы действия системы или программы наиболее полно отражает граф или дерево решений, чья степень сложности вложенности может быть значительной. Однако на языке алгоритма возможность системы принять решение, выходящее за пределы допустимого, может быть отражена следующим образом: в схеме появляется действие, которое переводит систему в состояние поиска другого перечня действий, отличных от описанных в основной программе.

Условно можно разделить весь объем решений, принимаемых системой, на два блока: инклюзивные (включенные в группу чего-либо) и эксклюзивные (ведущие к исключению). Данные термины восходят к латыни и буквально обозначают «включать» и «исключать». Смыслы данных слов не просто антонимичны, каждое из них шло по собственному пути развития, у «эксклюзива» появилось значение единичности, неповторимости, поскольку оно обозначает нечто такое, что нельзя включить в общую массу, поэтому оно отсюда исключено. В этом плане слово «инклюзив» не является антонимом, оно не означает нечто среднее или распространенное, не противопоставляется единичному. Эксклюзивное решение требуется тогда, когда происходит аварийная, опасная или нестандартная ситуация. Фактически такое решение может закончить деятельность данной системы, то есть оно находится вне заданий системы (программы). По аналогии с авиационной терминологией, эксклюзивное решение можно сравнить с «точкой невозврата». Последствия включения в систему эксклюзивных решений могут быть необратимы.

Решения, последствия которых могут носить правовой или этический характер, так же следует считать эксклюзивными. Проблема ответственности автономных и интеллектуальных технологий в научных исследованиях получила название «мораль-машина» [2, 7]. При этом следует понимать, что в процессе разработки киберфизических систем программист может в явном виде не закладывать подобные возможности системы, однако отсутствие



ограничений на определенные параметры могут привести к подобным последствиям. Например, система жизнеобеспечения не запрограммирована на отключение пациента от аппарата, однако она может установить объем подаваемого воздуха в легкие пациента равным нулю, тем самым лишив его жизни. Аналогично, система управления технологическим процессом производства может иметь возможность нарушить процесс, тем самым спровоцировать аварию. Ряд аналогий можно продолжать бесконечно, но все эти случаи объединяет одно – последствия принятия решения системой носят правовой или этический характер. Во всех рассматриваемых случаях встает вопрос выставления претензии системе, что по современным меркам абсурдно, поскольку за решение системы ответственность может нести либо разработчик, либо пользователь — в любом случае, субъект технического сознания [8]. Здесь можно прийти к доведенной до абсурда ситуации, описанной в одноактной пьесе С. Лема «Существуете ли вы, мистер Джонс?», в которой в гипертрофированной форме указывается на правовой вакуум [9], существующий и по сей день, связанный с развитием цифровизации и роботизации.

Подобного рода ограничения и проверки должны выполняться для киберфизических систем, применяемых на опасных производствах: добыча, транспортировка и переработка нефти и газа, атомная промышленность, химическая промышленность и прочее. Конкретный алгоритм действий при проверке еще предстоит разработать, если предположить, что он в принципе возможен в обобщенном, сводном формате. Автор предполагает, что подобное требование по контролю автоматизированных систем должен взять на себя регулятор в виде государственных или подвластных правительству структур.

Принципиально иную проблематику дает проектирование автоматизированных систем, изначально связанных с необходимостью взаимодействовать и воздействовать на живых существ. В первую очередь это связано с системами, разрабатываемыми в военной отрасли, сфере медицины (в части автоматизированного проведения операций, разработки роботизированных имплантатов и пр.), то есть во всех сферах, связанных с применением «автопилотов» в той или иной степени. Изначальная цель создания подобного рода систем уже предполагает использование некоторых этических норм в качестве предпосылок, что ведет к ответственному принятию решения системой. Хотя сами по себе этические предпосылки, заложенные в основу таких систем, находятся вне зоны рассмотрения настоящей статьи, необходимо понимать, что в полном объеме, во всеобъемлющем понимании человеком этики, реализация этих предпосылок в системах не возможна. Проблема при реализации может крыться в каждой точке процесса принятия решения: начиная от собственно распознавания аварийной ситуации [10, 11, 12] и заканчивая решениями по применимым реакциям [13, 14]. При применении подобных киберфизических систем прогнозируется снижение влияния человеческого фактора в процессе их функционирования, однако нужно понимать, что человеческий фактор при этом



остается в самом ядре системы, заложенный его разработчиком с учетом его этических норм и представлений. Ярким и доступным примером демонстрации подобного явления может служить рассказ С. Лема «Ананке», описывающий факт передачи психического заболевания от разработчика машине и его катастрофические последствия.

По аналогии с ранее описанными предложениями такие системы также могут быть подвергнуты анализу и тестированию на предмет наличия эксклюзивных решений и решений, приводящих к последствиям, незапланированным изначальной этической установкой. Современные киберфизические системы, особенно основанные на технологиях многослойных нейросетей, не могут быть подвергнуты анализу в явном виде, поскольку система в таком случае играет роль «черного ящика», преобразующего сигналы на входе в некоторое решение на выходе. Для подобных систем следует разработать по аналогии с автомобильной промышленностью «краш-тест», в результате которого можно было бы определить, как будет поступать система в граничных условиях. Весьма желательно, чтобы эти краш-тесты проходили в лабораторных условиях, до передачи в промышленное или массовое использование.

Заключение

Повсеместная цифровизация промышленности вызывает ряд вопросов и опасений, связанных с передачей киберфизическим системам функций принятия решений, которые могут впоследствии привести к катастрофе или аварии. Человечество не должно останавливать прогресс, однако необходимо понимать и оценивать риски новых технологий.

В научно-фантастической литературе многократно были описаны примеры эксклюзивных (в терминологии статьи) решений, приводящих к трагедиям различных масштабов. Это вполне логично и ожидаемо, поскольку писатель-фантаст не ограничен необходимостью описания существующих реалий. Как показывает практика, многие подобные вымыслы находят воплощение в жизни.

Научное сообщество, не имея в значительной степени готовых работающих систем, не имеет предмета обсуждения. Большинство научных статей, затрагивающих аналогичные вопросы, только указывают на необходимость проработки проблем, связанных с расширением применения автоматизированных средств управления, не давая конкретных рекомендаций [5, 6, 7, 14]. Аналогичные противоречия и ограничения встают и перед автором настоящей статьи.

В данной статье описаны предпосылки к созданию проекта по разработке системы ограничений для метасубъектов кибернетики четвертого рода при проектировании киберфизических систем, которые имеют возможность принимать решения в итоговых точках (в терминологии статьи). Предлагается продолжение обсуждения научным сообществам ограничений для



метасубъектов кибернетики четвертого рода, а также вынесение данного вопроса на уровень государственного регулирования.

Литература

- 1) Лепский В.Е. Философско-методологические основания становления кибернетики третьего порядка / Философские науки, №10: 2018 — с.7-28.
2. Чеклецов В.В. Философские и социо-антропологические проблемы конвергентного развития киберфизических систем (блокчейн, Интернет вещей, искусственный интеллект) / Философские проблемы информационных технологий и киберпространства, №1: 2016.
 1. Куприяновский Д.Е. Киберфизические системы как основа цифровой экономики / В.П. Куприяновский, Д.Е. Намиот, С.А. Синягов / International Journal of Open Information Technologies, Т.4, №2: 2016 – с.18-24.
 1. Wolf W. Cyber-physical systems / Computer, №3, 2009 – с.88-89.
 - Дессауэр Ф. Спор о технике /перев. Нестеров А.Ю. Самара: Издательство Самарской гуманитарной академии, 2017 - с. 266.
 1. Лем С. Сумма технологии.
 1. Проблема виртуального бессмертия в электронной культуре. Лепский, В.Е. 2018 г., Институт философии РАН, стр. 170-184.
 2. Нестеров А.Ю. Вопрос о сущности техники в рамках семиотического подхода / Вестник Самарского государственного аэрокосмического университета. Т.14, №1: 2015 - с. 235-246.
 3. Андреева О.И. Технологии распознавания лиц в уголовном судопроизводстве: проблема оснований правового регулирования использования искусственного интеллекта / Андреева О.И., Иванов В.В., Нестеров А.Ю., Трубникова Т.В. / Вестник Томского государственного университета №449: 2019 - с.201-212.
 4. Слободская А.В. Этические проблемы робототехники / Евразийский Союз Ученых (ЕСУ), №7: 2014.
 5. Куксин И.Н. "Цифровая революция": правовые и этические проблемы / Куксин И.Н., Мархгейм М.В. 2019.
 6. Пономаренко Г.С. Об этических проблемах современной робототехники / Политехнический молодежный журнал: 2016 - с. 4.
 7. Камалетдинов Дамир. Сбивший американку насмерть самоуправляемый Uber «не знал», что люди могут оказаться на дороге без перехода / Tjournal.ru 9.11.2019 <https://tjournal.ru/tech/124873-sbivshiy-amerikanku-nasmert-samoupravlyaemyu-uber-ne-znal-chto-lyudi-mogut-okazatsya-na-doroge-bcz-perchoda>.
 8. Иран признал, что сбил украинский "боинг". Его приняли за крылатую ракету / bbc.com 11.01.2020 <https://www.bbc.com/russian/news-51073468>.
 9. Реннел Тони. 26 сентября 1983 года. День, когда едва не погиб наш мир. / Иносми.ру: 2007 <https://inosmi.ru/inrussia/20071229/238739.html>.



10. E. Awad. The Moral Machine experiment / E. Awad, S. Dsouza, R. Kim, J. Schulz, J. Henrich / Nature: 2018 (563) — с.59-64.
11. Грунвальд А. Беспилотный автомобиль в свете социальной оценки техники / А. Грунвальд, В.Н. Железняк, Е.В. Середкина / Технологос, №2 Пермь: 2019.

И.В. Демин

ВЛАДИМИР ВЕРНАДСКИЙ И ЛЕВ ЛОПАТИН: СПОР О НАУЧНОМ МИРОВОЗЗРЕНИИ

(Самарский университет)

В философских дискуссиях первой четверти XX вв. вопрос о «научном мировоззрении» ставился и решался в контексте обсуждения более общей и фундаментальной проблемы – соотношения метафизики и позитивных наук.

Выражение «научное мировоззрение» вошло в язык русской философии и публицистики примерно в то же время, когда стал широко использоваться термин «мировоззрение» (и его понятийные аналоги – «миросозерцание», «мировидение», «миросознание»), представляющий собой, как известно, кальку с немецкого *Weltanschauung*. Понятие «мировоззрение» впервые появляется у немецких романтиков в начале XVIII в. Широкое распространение этот термин получает в конце XIX – начале XX вв. благодаря работам Г. Риккерта, В. Дильтея, М. Шелера, К. Ясперса и некоторых других философов. В таких направлениях, как неокантианство, философия жизни, экзистенциализм, «мировоззрение» становится одной из центральных категорией и приоритетной темой философских изысканий⁴.

Вопросы о сущности и специфике научного познания, месте науки (естествознания) в жизни культуры и судьбе цивилизации, соотношении науки и религии, границах и опасностях «сциентизма» обсуждались русскими философами на протяжении всей второй половины XIX века, однако отправной точкой философских споров вокруг понятия «научное мировоззрение» стала статья В.И. Вернадского, впервые опубликованная в журнале «Вопросы философии и психологии» в 1902 г.⁵

Трактовка науки, философии и «научного мировоззрения» в статье Вернадского в значительной мере обусловлена явной или скрытой полемикой с *позитивизмом*. Эта трактовка может быть сформулирована в следующих тезисах.

⁴ Подробнее см.: Демин И. В. Мировоззрение гуманизма и гуманизм мировоззрения // Свобода совести и диктатура миросозерцания: история взаимоотношений и современные реалии. – Орёл: Картуш, 2016. – С. 88-101.

⁵ Вернадский В. И. О научном мировоззрении // Вопросы философии и психологии. – 1902. – № 65. – С. 1409-1465.