



ТЕХНОЛОГИЧЕСКАЯ СИНГУЛЯРНОСТЬ В КОНТЕКСТЕ ТЕОРИИ МЕТАСИСТЕМНЫХ ПЕРЕХОДОВ

Потапов А. С.¹

¹ Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург, Россия

Аннотация

Идея технологической сингулярности проистекает из экстраполяции эмпирических кривых ускорения технического развития. Нередко, однако, из этих ненадежных экстраполяций делаются далеко идущие следствия, которые подвергаются активной критике, и эта критика некорректно переносится на саму концепцию технологической сингулярности. В статье делается попытка отделить данную концепцию от дополнительных предположений и убеждений, в частности, касательно возможности создания искусственного интеллекта, а также выделить ряд заблуждений. Для этих целей привлекается теория метасистемных переходов, и технологическая сингулярность определяется как точка, в которой ранее наблюдаемые закономерности уменьшения времени между метасистемными переходами и уменьшения времени удвоения производительности кибернетических систем внутри одного уровня перестают выполняться. Указывается, что до достижения этой точки, вероятно, успеет произойти ряд метасистемных переходов, и возникшие системы будут показывать экспоненциальный рост с периодом удвоения менее полугода и превысят сложность существующих систем в течение нескольких десятилетий.

Ключевые слова: технологическая сингулярность; метасистемные переходы; искусственный интеллект.

Цитирование: Потапов А. С. Технологическая сингулярность в контексте теории метасистемных переходов // Компьютерные инструменты в образовании. 2017. № 6. С. 12–24.

1. ВВЕДЕНИЕ

Технический прогресс имеет вполне ощутимо ускоряющийся характер. Существует много экспоненциальных трендов, включающих, помимо широко известного закона Мура, также, к примеру, экспоненциальное увеличение числа подключенных к сети интернет компьютеров или объемов данных, получаемых посредством технологий сканирования мозга. Более того, более поздние технологические тренды имеют больший показатель экспоненты. Таким образом, технический прогресс в целом оказывается гиперэкспоненциальным, асимптотически устремляющимся к бесконечности за конечное время, составляющее от текущего момента несколько десятилетий (см., например, [1]). Точка, в которой некоторая функция стремится к бесконечности или не определена, в

математике называется особенностью, или сингулярностью. По аналогии гипотетическую точку, при приближении к которой технический прогресс оказывается неограниченным, называют технологической сингулярностью (Сингулярностью).

Концепция Сингулярности получила широкое распространение и многими своими сторонниками воспринимается некритически. Как результат, на ее основе делаются далеко идущие (см., например, [2]), но плохо обоснованные выводы и предсказания. Критика таких предсказаний, однако, не должна переноситься на саму базовую концепцию Сингулярности и обесценивать ее.

Относительно концепции Сингулярности существуют разные заблуждения. Одно из наиболее часто встречающихся заблуждений состоит в отождествлении Сингулярности как события с созданием искусственного (сверх)интеллекта. Даже статья “Technological singularity” на англоязычной Wikipedia [3] начинается с определения Сингулярности как гипотезы о том, что изобретение искусственного сверхинтеллекта запустит процесс стремительного технологического роста, который кардинальным образом изменит человеческую цивилизацию (хотя далее и упоминает о разных путях к Сингулярности, рассмотренных В. Винджем [4], впервые популяризовавшим данную концепцию).

Как следствие, одним из направлений критики идеи Сингулярности является критика мыслящих машин (см., например, [5]), возможность создания которых сама по себе вызывает споры, особенно если под этими машинами неявно подразумеваются существующие компьютеры. В то же время базовая концепция Сингулярности совершенно не требует создания «Сильного ИИ» на базе цифровых компьютеров.

Вина за подобные заблуждения во многом лежит на самих сторонниках идеи Сингулярности, поскольку они нередко пытаются одновременно защищать целый комплекс своих убеждений и желаний, полагая, что они усиливают доказательность друг друга. В то же время оппоненты ищут наиболее слабые свидетельства, полагая, что их критика опровергает весь комплекс идей целиком. При этом они также зачастую руководствуются своими субъективными предпочтениями (которые утрированно можно сформулировать как «Сингулярность невозможна, так как мы не хотим исчезновения людей»).

В данной статье проводится попытка разделить различные факторы, включая обоснованные утверждения и факты от личных убеждений и желаний, и установить, что мы можем сказать о Сингулярности с использованием минимальных дополнительных предположений и спекулятивных прогнозов. В частности, понятие Сингулярности описывается в терминах метасистемных переходов как объективного феномена.

2. БАЗОВЫЕ ПОЛОЖЕНИЯ

2.1. Теория метасистемных переходов

Хотя теория метасистемных переходов редко упоминается в связи с концепцией Сингулярности (исключением является работа [6], в которой данная теория применяется к концепции «Глобального Мозга»), ее вполне можно считать одним из возможных научных оснований данной концепции. Эта теория, предложенная В. Ф. Турчиным [7], рассматривает эволюцию кибернетических систем (то есть систем управления), как биологических (преимущественно нервной системы как естественной кибернетической системы), так и культурных. Эволюция этих систем состоит из последовательности метасистемных переходов, каждый из которых заключается в возникновении системы управления следующего уровня, которая осуществляет выбор между различными состояниями

ми или экземплярами уже существующих более низкоуровневых систем управления.

К примеру, кибернетическая система сначала обладает положением в пространстве. Управление этим положением — это движение, которое само исходно неуправляемо. Простейшее управление движением (реакция на элементарные стимулы) — это раздражимость. Управление раздражимостью в условиях, когда у кибернетической системы есть много сенсоров и эффекторов, приводит к простому рефлексу (координированной, но негибкой реакции множества эффекторов на определенный сенсорный сигнал). Управление множеством простых рефлексов — это сложный (условный) рефлекс, или ассоциация (в авторской терминологии). Управление ассоциациями — это мышление. Управление мышлением (в кибернетической перспективе) — культура (см. табл. 1).

Таблица 1. Этапы эволюции по В. Ф. Турчину [7]

Химическая эра	Химические формы жизни
	Движение
	Раздражимость (простой рефлекс)
Кибернетическая эра	Нервная сеть (сложный рефлекс)
	Ассоциирование (условный рефлекс)
Эра разума	Мышление
	Культура, социальная интеграция

Кроме того, В. Ф. Турчин рассматривал метасистемные переходы внутри культуры (в частности, математики), но и прочие этапы эволюции могут быть описаны более детально. К примеру, в ходе генетической эволюции возникали разные уровни управления генами. Также и в техническом развитии (в особенности, вычислительной техники) могут быть выделены различные метасистемные переходы.

Существует ряд расширений и детализаций этой теории (см., например, [8]), однако в рамках нашего рассмотрения их детали не существенны.

2.2. Временные закономерности

В рамках теории метасистемных переходов не проводился анализ временных и количественных характеристик эволюции кибернетических систем. В то же время как раз графики возрастания производительности систем, построенные по некоторым «ключевым событиям» в биологической эволюции и/или в истории развития человеческой цивилизации для некоторых количественных показателей (емкости памяти, скорости вычислений, объемов производства и т. д.), обычно рассматриваются в качестве эмпирического основания концепции Сингулярности. Хотя разные авторы выбирают разные «ключевые события», кривые возрастания сложности и уменьшения интервалов между «сменами парадигм» друг с другом хорошо согласуются, как показано Р. Курцвейлом на примере 15 различных списков ключевых событий [1].

Вполне естественно объединить теорию метасистемных переходов, указывающую на общие механизмы глобальной, или универсальной эволюции (объединяющей разные уровни организации материи [9]), с эмпирическим анализом временных и количественных характеристик эволюционирующих кибернетических систем. Тогда можно заключить следующее:

1. Производительность или сложность систем определенного уровня управления возрастает с течением времени экспоненциально.
2. Время между последующими метасистемными переходами сокращается геометрически, а темпы роста на следующих уровнях управления с каждым метасистемным переходом возрастают геометрически.

Таким образом, Сингулярность — это точка, в которой данные закономерности перестают выполняться.

Эти закономерности, хотя и не являются столь же точными, как, скажем, физические законы, но хорошо обоснованы эмпирически. Весьма трудно отрицать, что, скажем, рост числа нейронов в нервных системах в ходе эволюции или рост числа транзисторов в процессорах компьютеров экспоненциальный, причем время удвоения последних существенно меньше, чем первых. Сомнению может быть подвержена, разве что, значимость этих наблюдений и выводимые из них следствия.

3. ПРЕДСКАЗАНИЯ

Касательно Сингулярности обычно делаются предсказаний (см., например, [1, 4]), которые можно разделить на два типа — экстраполяция кривых возрастания сложности и качественное описание возможных сценариев развития технологий.

3.1. Экстраполяция кривых роста

Вполне естественно, что экстраполяция эмпирических данных не дает достоверного знания, так как сводится к проблеме индуктивного вывода. В этой связи никакие предсказания нельзя абсолютизировать. Тем не менее, можно пытаться построить более вероятную экстраполяцию. При отсутствии дополнительной информации простейшее (например, в смысле алгоритмической сложности по А. Н. Колмогорову) продолжение некоторого временного ряда является наиболее вероятным. В рассматриваемом случае простейшая экстраполяция ведет к цепочке метасистемных переходов, между которыми интервалы времени будут все сокращаться, а показатель экспоненты роста для каждой последующей системы — увеличиваться, что в итоге приводит к асимптотическому приближению кривой возрастания сложности к бесконечности за конечное время. «Истинная» Сингулярность — воображаемая точка, конкретное положение которой определено не вполне точно, но по многим прогнозам технических трендов (см., например, [1]) оказывается в пределах нескольких десятилетий от текущего момента.

Одним из сопутствующих аргументов для этих предсказаний являются предсказания возможного момента появления условий для создания искусственного интеллекта (например, [10–13]). Здесь мы, однако, не будем опираться на эти предсказания в целях избежать обсуждения спорного вопроса о самой возможности создания мыслящих машин.

Любопытнее то, что, как показано в работе [14], различные подходы к предсказанию долговременных трендов, включающие помимо чисто технических трендов также и экономические и социальные циклы, независимо прогнозируют технологический всплеск в 2040-х годах, который может соответствовать Сингулярности, если определенные технологии окажутся доступными к этому моменту.

Естественно, в соответствии с существующей научной картиной мира, «истинная» сингулярность (технологическая или нет) физически невозможна. Не исключено изменение этой картины в будущем, однако на настоящий момент вряд ли ее стоит пере-

смагивать только потому, что простейшая экстраполяция одной эмпирической кривой подразумевает бесконечность. Скорее, следует полагать эту экстраполяцию не вполне адекватной.

Вероятно, стоит ожидать замедления скорости роста, что, однако, не разрушает всю концепцию Сингулярности: черная дыра не перестает быть объектом с уникальными свойствами, даже если материя в ней не сжимается в точку с бесконечной плотностью. Таким образом, действительный интерес представляет вопрос, насколько далеко и быстро продолжит возрастать сложность кибернетических систем.

Следующая простейшая экстраполяция — это S-образная кривая, подразумевающая смену ускорения роста замедлением с перегибом в некоторой точке и дальнейшим приближением к некоторому конечному порогу. Некоторые авторы полагают, что такие S-образные кривые являются типичным паттерном в кривых возрастания сложности индивидуальных технологий (как отмечается, например, в [14]), который вполне может повторяться фрактально в последовательностях метасистемных переходов на различных временных масштабах.

Конкретная форма не имеет принципиального значения. Для людей в текущий момент времени даже не столь существенно выйдет ли кривая развития на насыщение или будет участвовать в S-образном паттерне еще большего масштаба. Для нас важно, когда и как быстро начнется замедление. Пока нет оснований полагать, что замедление будет неожиданным и резким.

S-образная кривая — достаточно консервативная экстраполяция, и мы не видим признаков того, что замедление уже началось. Напротив, во многих областях очевидны ускоряющиеся темпы развития. Вероятно, мы еще не прошли точку перегиба, а после этой точки мы будем наблюдать хотя и замедляющийся, но все еще очень быстрый рост (если не рассматривать катастрофические сценарии), который приведет к возникновению кибернетических систем на переднем фронте глобальной эволюции, существенно превосходящих по сложности существующие сейчас системы, что для людей может быть мало отличимо от «истинной» Сингулярности (что ниже мы будем характеризовать как «фактическая» Сингулярность).

3.2. Возможные сценарии

Ни один конкретный сценарий в рамках концепции Сингулярности не является достаточно обоснованным предсказанием и не может быть использован для критики общей концепции Сингулярности. Означает ли это, что данная концепция не может давать предсказаний, которые могут быть подвержены проверке, то есть не удовлетворяет критерию фальсифицируемости Поппера, а следовательно, ненаучна? Не вполне. Мы не можем точно сказать, в какой конкретно форме произойдет следующий метасистемный переход, но мы можем предсказать, что **некоторый метасистемный переход произойдет в указанном временном интервале, и возникшая метасистема будет развиваться экспоненциально с коротким временем удвоения (например менее полугодом) и превзойдет производительность существующих кибернетических систем в течение нескольких десятилетий.**

Кроме того, мы все же можем выдвинуть некоторые предположения о том, какой сценарий является относительно более вероятным. Все сценарии основаны на так называемых сингулярных технологиях, то есть технологиях, которые ускоряют свое собственное развитие (что обычно связывается с некоторой формой сверхинтеллекта). Обычно в ка-

честве иллюстрации приводится искусственный интеллект, который становится быстрее за счет изобретения более быстрых компьютеров, что позволяет ему изобретать эти компьютеры еще быстрее. Но мыслящие машины здесь не обязательны. К примеру, генная инженерия может привести к появлению более умных людей, которые будут быстрее выполнять генетические исследования, приводящие к появлению еще более умных людей еще быстрее и т. д. В целом подобные сценарии можно охарактеризовать как «взрыв интеллекта».

Широкие классы сингулярных технологий включают био-, нано-, инфо- и, возможно, некоторые другие технологии и их комбинации (например, можно представить нанороботов, населяющих человеческий мозг и расширяющих его возможности). Эти технологии имеют различные времена удвоения. К примеру, для появления генетически модифицированных людей нужны годы для роста и обучения, тогда как другие формы суперинтеллекта могут развиваться быстрее, опережая путь генетических модификаций (особенно если говорить о социальных факторах, связанных с генной инженерией).

Естественно, подобный анализ будет далек от строгого и не может привлечь внимание неизвестные будущие технологии, а также не рассматривает взаимодействия между разными технологиями. Он также не учитывает социальные, геополитические и экономические факторы, которые могут быть необходимы для предсказания будущего развития (см., например, [14]). Тем не менее, он может нам позволить предположить, какие технологии с меньшим временем удвоения могут привести к следующему метасистемному переходу. Такой анализ, однако, не является предметом настоящей статьи, и мы лишь хотим подчеркнуть, что предположения о возможных реализациях последующих метауровней не должны использоваться для критики концепции Сингулярности как таковой.

Дополнительным источником для предсказаний является сама теория метасистемных переходов, так как она предлагает для них общие механизмы. К примеру, можно высказать предположение, что следующей метасистемой будет «контроль культур». Кто-то может даже сказать, что этот переход уже осуществляется в том смысле, что люди взаимодействуют друг с другом и искусственными агентами через интернет, образуя метасистему, либо же он будет достигнут в форме «глобального мозга» (например [6, 15]). Хотя подобные сценарии выглядят логическим следствием теории метасистемных переходов, эта теория не является достаточно детальной, чтобы описать и предсказать «аппаратное обеспечение» метасистем. Например, она ничего не говорит о том, почему ДНК как «аппаратное обеспечение» кибернетических систем сменилась (или, скорее, была расширена) нейронами. Аналогично, теория не говорит нам о том, на каком аппаратном обеспечении должен воплощаться уровень культуры и последующие метауровни.

Исходно культура «исполнялась» посредством нервных систем людей, дополненных такими артефактами как письменность. К человеку эти «культурные сети» относятся, скорее, как к генам относятся генные сети, а не нейроны. Так и в «глобальном мозгу» роль человека может соответствовать не нейрону, а лишь гену, экспрессирующемуся внутри нейрона. Сам же «нейрон» может быть реализован на другой аппаратной основе. Вполне возможно, что роль нейрона будущего «глобального мозга» может остаться за компьютерами, которые уже сейчас участвуют в управлении нашей культурой через системы рекомендаций, фильтрации новостей и т. д. Это нисколько не противоречит структуре метасистемных переходов, как и сценарий, в котором люди объединяются в «глобальный мозг», напрямую соединяя свои нервные системы. В этой связи конкретная роль людей в будущих кибернетических системах не может быть предсказана теорией метасистемных переходов, из которой можно лишь заключить, что **следующая метаси-**

стема будет основана на человеческой культуре в том или ином смысле.

4. ЗАБЛУЖДЕНИЯ

4.1. «Истинная» Сингулярность

Идея «истинной» Сингулярности (принимаемой не просто как модельное упрощение, а как ожидаемая реальность) имеет религиозный оттенок, так как предполагает возникновение бесконечно могущественной богоподобной сущности. Как мы отмечали, единственным указанием на подобный сценарий является одна (пусть и априори наиболее вероятная) из экстраполяций единственной эмпирической кривой, в то время как против него свидетельствует весь корпус научных знаний. Естественно, из исследователей вряд ли кто-то действительно придерживается этой идеи в ее крайней форме. Однако существует ее более мягкая версия, в соответствии с которой образующаяся сущность остается конечной, но расширяется, скажем, до всей Вселенной. При этом ограничение на скорость света преодолевается благодаря разработке «онтотехнологий», которые позволяют менять физические законы Вселенной так же, как люди могут менять свой геном. Вселенная, приобретшая разум, начинает общаться с другими разумными вселенными внутри Мультиверса. Хотя подобные идеи имеют некоторые основания в физических теориях (касательно возможного места нашей Вселенной в Мультиверсе [16, 17]), они, конечно, крайне далеки от доказательности.

Хотя подобные идеи и заняты, их нельзя считать в качестве сколько-нибудь вероятных предсказаний. И, напротив, их низкая правдоподобность не должна рассматриваться в качестве контраргумента против Сингулярности как таковой. Заблуждением здесь является то, что под Сингулярностью понимается «истинная» Сингулярность.

Однако также необходимо отметить, что хотя многие определения Сингулярности в явном виде не опираются на асимптотический технический прогресс, а бесконечные темпы роста выглядят физически бессмысленными, но и экспоненциальный рост не достаточен для достижения Сингулярности, как обсуждается в [18]. Представление о Сингулярности как простой экстраполяции закона Мура также является заблуждением.

4.2. Люди

Что можно сказать о «фактической» Сингулярности, которая не столь фантастична, как «истинная» Сингулярность, но и гораздо большее, чем экспоненциальный прогресс? Можно было бы сказать, что это точка «в истории нашей расы, за которой человеческие дела, как мы их знаем, не смогут продолжаться», как это было сформулировано Стэнном Уламом со ссылкой на его разговор с Джоном фон Нейманом более 60 лет назад. Однако это определение далеко от строгого. С одной стороны, многие человеческие дела сейчас весьма отличны от того, что было 200 лет назад. С другой стороны, некоторые человеческие дела (например наука) могут продолжаться и без людей, особенно людей в существующем виде. Будут ли генетически модифицированные или аугментированные люди все еще людьми? Является ли человек, использующий ручку и бумагу или компьютер, все еще просто человеком? Это риторические вопросы.

Мы не можем определить «фактическую» Сингулярность относительно людей, которые постоянно меняются как часть большей метасистемы. Возможно, Сингулярность уже свершается в соответствии с определением 60-летней давности. Будем ли мы так

считать или нет, не изменит реальности. Конечно, для самих людей судьба их жизней имеет значение, но это то, что в контексте Сингулярности предсказать трудно. Мы можем лишь сказать, что глобальная эволюция будет продолжаться, и метасистемные переходы будут иметь место и будут вести к возникновению кибернетических систем гораздо большей сложности, чем отдельный человек, лишенный дополнительных инструментов.

4.3. У нас есть выбор

Глобальная эволюция продолжается миллиарды лет. Ее законы метасистемных переходов не являются строгими, но являются объективными, не зависящими от человеческой воли и желаний. Конечно, люди гораздо разумнее, чем ДНК или отдельные нейроны. Складывается впечатление, что мы можем выбирать, будет иметь место Сингулярность или нет. Но можем ли мы выбрать прекращение глобальной эволюции? Вряд ли в полной мере. Разные люди имеют разные мнения на этот счет. Страны и корпорации имеют собственный интерес. Сложно представить, что развитие всех сингулярных технологий (а вернее, даже всех технологий) будет запрещено во всех странах и не будет действительно абсолютно никем осуществляться. Конечно, люди могут совершать действия, ускоряющие или замедляющие прогресс, возможно, даже в сравнительно больших пределах. Однако все сценарии, при которых Сингулярность не достигается в XXI веке, видимо, связаны с глобальными катастрофами.

Люди также полагают одни сценарии более вероятными по сравнению с другими сценариями на основе своих субъективных предпочтений. Например, трансгуманисты предпочитают говорить о «загрузке сознания», рассматривая общий искусственный интеллект как не особо релевантный или даже нежелательный, просто потому что «цифровое бессмертие» выглядит привлекательнее, чем сюжет «Терминатора». Мы здесь не обсуждаем возможность «загрузки сознания» или «мыслящих машин», однако мы можем сравнить эти два сценария, поскольку они принадлежат одной парадигме (реализация мышления с использованием компьютеров). Моделирование естественных нейронов требует огромных накладных расходов. Таким образом, вычислительные ресурсы для ИИ уровня человека будут доступны существенно раньше, чем ресурсы для эмуляции человеческого разума в масштабе реального времени. Для уменьшения накладных расходов требуется детальное понимание того, как можно абстрагироваться ото всех биохимических и физических деталей (например от трехмерного фолдинга белков, моделирование которого является крайне ресурсоемкой задачей, но который требуется для эмуляции результатов экспрессии генов, в частности, участвующих в консолидации следов долговременной памяти). Следовательно, у нас уже будет достаточно детальная модель (человеческого) интеллекта для этого, то есть и ИИ может быть уже создан.

Общий ИИ как сингулярная технология будет обладать меньшим временем удвоения, поскольку знание о принципах его построения откроет доступ к самооптимизации или к расширению дополнительными модулями, сенсорными модальностями и т. д. Таким образом, хотим мы этого или нет, но если эти технологии возможны, то общий ИИ возникнет раньше и будет развиваться быстрее, чем «загрузка сознания» или полная эмуляция мозга.

Правительства, корпорации, научные сообщества и т. д. могут влиять на скорость разработок различных технологий через финансовую поддержку, законодательные ограничения и т. д., но они не повлияют на имманентные объективные свойства этих

технологий, которые играют основную роль в том, какой путь будет выбран глобальной эволюцией. Предсказание правдоподобности различных сценариев должно делаться на детальном сравнении объективных свойств различных технологий и их возможном взаимном влиянии, но не на основе наших желаний. Тем не менее, это вовсе не означает, что реализация будущего технического прогресса predetermined, и никакого влияния на нее оказано быть не может.

4.4. Искусственный сверхинтеллект

Как было отмечено, Сингулярность часто ассоциируется с созданием искусственного сверхинтеллекта (и даже обосновывается им, см., например [19, 20]), что, однако, также является и источником критики самой концепции Сингулярности [5].

Так, хрестоматийный пример компьютерного ИИ, который разрабатывает все более быстрые компьютеры и сам становится все быстрее и т. д., часто подвергается критике, хотя является лишь иллюстрацией (а не предполагаемой реализацией) к идее «взрыва интеллекта». Как отмечалось, в этот воображаемый сценарий может быть подставлена любая сингулярная технология или набор технологий. Например, можно представить, что люди используют компьютеры, чтобы проводить генетические исследования и совершенствовать сами компьютеры, что приводит как к более умным людям, так и более быстрым компьютерам, ускоряя оба направления исследований с положительной обратной связью.

В этой связи можно высказать два утверждения.

- Сама концепция Сингулярности может быть описана как последовательность ускоряющихся метасистемных переходов без опоры на идею сверхчеловеческого «сильного» ИИ и может быть обоснована независимо от нее.
- Напротив, концепция Сингулярности (в совокупности с рассмотрением скорости удвоения различных сингулярных технологий) может служить дополнительным обоснованием возможности возникновения сверхчеловеческого общего ИИ в течение нескольких десятилетий.

Может показаться, что второе утверждение говорит в точности то же самое, что и упомянутая статья [3]. Однако это не так, поскольку причинно-следственные связи отличны. Если кто-то говорит, что создание ИИ приведет к Сингулярности, то если мы поставим под сомнение возможность создания такого ИИ, то мы в еще большей степени будем сомневаться касательно реализуемости Сингулярности.

Напротив, если сначала будут предъявлены независимые свидетельства в пользу концепции Сингулярности, то, даже если мы будем склоняться к негативному ответу на дискуссионный вопрос относительно возможности мыслящих машин, это не окажет влияния на достоверность указанной концепции. Далее мы можем предоставить аргументы, что общий ИИ более вероятен как реализуемая сингулярная технология. Это самостоятельное и более слабое утверждение, которое не оказывает влияния на первое утверждение.

Действительно, мы видели, что концепция Сингулярности может быть введена независимо от идеи искусственного сверхинтеллекта (хотя и не вполне независимо от идеи сверхинтеллекта как такового). Необходимость сверхинтеллекта быть «искусственным» — это дополнительная независимая посылка. Кроме того, если от сверхинтеллекта не требуется обладание самосознанием или целостностью, то такой сверхинтеллект не является чем-то фантастическим, а вполне может считаться уже существующим и постоянно прогрессирующим (причем у нас нет оснований полагать, что

этот прогресс внезапно остановится). К примеру, человеко-машинные системы сейчас способны решать такие задачи, решение которых ранее было бы сложно представить без сверхинтеллекта.

Однако и сугубо искусственный суперинтеллект возможен. Нет каких-либо фундаментальных ограничений, препятствующих этому, особенно если мы не будем ограничиваться существующими компьютерами, но и учтем возможные будущие компьютеры, основанные на других (возможно, сейчас неизвестных) физических принципах. Мы вовсе не утверждаем, что такие компьютеры действительно нужны для реализации мыслящих машин, но лишь указываем на тот факт, что даже такие критики «сильного» ИИ как Дж. Сёрл и Р. Пенроуз адресовали свою критику только цифровым компьютерам, а не всем возможным вычислительным устройствам в принципе. Можно также добавить (см., например, [19]), что от искусственного сверхинтеллекта вовсе не обязательно требовать быть «сильным», то есть обладающим всеми субъективными человеческими качествами, но достаточно лишь быть общим, а к общему ИИ не применима основная часть критики «сильного» ИИ.

Любопытно отметить, что те особенности человеческого интеллекта, воспроизведение которых в компьютере чаще всего ставят под сомнение критики «сильного» ИИ (например, чувство юмора, эмоции и т. д.), в наименьшей степени затронуты прогрессом. Экстраполяция этой тенденции делает более вероятным общий, а не «сильный» ИИ, который может не обладать субъективными переживаниями в том же виде, что и человек. Можно сокрушаться односторонностью прогресса и тем, что в ходе этого прогресса что-то сугубо человеческое может оказаться утерянным, но не отрицать факт (гипер)экспоненциального развития технологий. Тем не менее, хотя мы здесь и хотим избежать споров о возможности «сильного» ИИ, но все же необходимо отметить, что многие исследователи не считают создание «сильного» ИИ невозможным и при этом разрабатывают теории искусственной креативности, любопытства, юмора и т. д. (см., например, [21]).

Таким образом, утверждать, что искусственный сверхинтеллект строго не возможен ни в какой форме, и отвергать концепцию Сингулярности на этой основе, представляется ненаучным. Однако не следует также утверждать, что конкретная форма ИИ (например, на основе цифровых компьютеров) будет неизбежным шагом на пути к Сингулярности. Хотя личным мнением автора является то, что общий ИИ может быть создан на базе цифровых компьютеров и является наиболее вероятным этапом технического прогресса в связи с наименьшим системным временем, это действительно мнение, которое может оказаться ошибочным, в связи с чем оно здесь не отстаивается, чтобы его дискуссионность не была перенесена на концепцию Сингулярности.

5. ЗАКЛЮЧЕНИЕ

Технологической сингулярности даются различные определения. Какие-то из них в явном виде включают возникновение искусственного сверхинтеллекта, тогда как другие говорят о технологических изменениях, недоступных человеческому пониманию или прогнозированию. Подобные расхождения являются источником полемики. Да, и сами понятия «искусственного» и «сверхинтеллекта» могут пониматься существенно различным образом. В этой связи не следует спорить об определениях, но нужно быть точным в своих утверждениях.

В этой статье была выполнена попытка выделить утверждения о Сингулярности, ко-

торые могут защищаться независимо от предсказаний о конкретном характере технического прогресса, в особенности, в части создания искусственного интеллекта.

Хотя некоторые точки зрения касательно ИИ приводились, попытки их обосновать не делалось, и основной целью статьи было выделить те утверждения, которые можно сделать о Сингулярности. В частности, можно утверждать, что вскоре произойдет некоторый метасистемный переход с уровня человеческой культуры, и возникшая метасистема будет развиваться экспоненциально со временем удвоения менее полугода и превзойдет производительность существующих кибернетических систем в течение нескольких десятилетий. Весьма вероятно, что будут иметь место последующие метасистемные переходы, хотя скорость их возникновения может в какой-то момент начать уменьшаться.

Будет ли это Сингулярностью? Ответ на этот вопрос зависит от определений, а также от конкретного сценария, который будет иметь место, и который проблематично предсказать с достаточной степенью уверенности (с учетом количества футурологов, делающих различные предсказания, кто-то из них вполне может оказаться прав, но это не означает, что в настоящий момент правильное предсказание достаточно обосновано). Мы определили Сингулярность как точку, в которой перестают выполняться закономерности развития кибернетических систем в ходе последовательных метасистемных переходов. Достижение этой точки неизбежно в силу характера указанных закономерностей, однако оно не обязательно будет для людей выделяться на фоне других событий.

Таким образом, за концепцией Сингулярности стоит вполне реальный феномен ускользящей глобальной эволюции, и данная концепция не должна отбрасываться лишь потому, что для этого феномена предложены лишь простые предиктивные модели. Вся критика должна быть отнесена на использование этих моделей за область их применимости.

Список литературы

1. *Kurzweil R.* The Singularity is Near. N.Y.: Viking Books, 2005.
2. *Garis H.* How Will the Artilect War Start? In *The End of the Beginning. Life, Society and Economy on the Brink of the Singularity*, *Goertzel B., Goertzel T.* (eds.), 2015.
3. Wikipedia: Technological Singularity. URL: https://en.wikipedia.org/wiki/Technological_singularity (дата обращения 20.12.17).
4. *Vinge V.* The Coming Technological Singularity: How to Survive in the Post-Human Era // *Vision-21: Interdisciplinary Science and Engineering in the Era of Cyberspace*. San Diego: San Diego State University. 1993. P. 11–22.
5. *Braga A., Logan R. K.* The Emperor of Strong AI is Naked: Limits to Artificial Intelligence // *Information (Switzerland)*. 2017. № 8(4). P. 156.
6. *Heylighen F.* Return to Eden? Promises and Perils on the Road to Global Superintelligence // *The End of the Beginning. Life, Society and Economy on the Brink of the Singularity / Goertzel B., Goertzel T.* (eds.), 2015.
7. *Турчин В. Ф.* Феномен науки: кибернетический подход к эволюции / В.Ф. Турчин. 2-е изд. М.: ЭТС, 2000. [Turchin, V.F. *The Phenomenon of Science: A Cybernetic Approach to Human Evolution*. 1st ed. N.Y.: Columbia University Press, 1977].
8. *Heylighen F.* (Meta)systems as Constraints on Variation: a classification and natural history of metasystem transitions // *World Futures: the Journal of General Evolution*. 1995. Vol. 45. P. 59–85.
9. Wikipedia: Universal Evolution. URL: https://en.wikipedia.org/wiki/Universal_evolution (дата обращения 20.12.17).
10. *Moravec H.* When will computer hardware match the human brain? // *Journal of Transhumanism*. 1998. Vol. 1.

11. *Bostrom N.* How Long Before Superintelligence? // Linguistic and Philosophical Investigations. 2006. Vol. 5, № 1. P. 11–30.
12. *Baum S., Goertzel B., Goertzel T.* How long until human-level AI? Results from an expert assessment. Technological Forecasting and Social Change. 2011, vol. 78, iss. 1, pp. 185-195.
13. *Grace K. et al.* When Will AI Exceed Human Performance? Evidence from AI Experts. arXiv:1705.08807. 2017.
14. *Goertzel T., Goertzel B.* Predicting the Age of Post-Human Intelligences. In *The End of the Beginning. Life, Society and Economy on the Brink of the Singularity*, Goertzel, B., Goertzel, T. (Eds.), 2015; ISBN: 0692457666.
15. *Vidal C.* Distributing Cognition: From Local Brains to the Global Brain. In *The End of the Beginning. Life, Society and Economy on the Brink of the Singularity*, Goertzel, B., Goertzel, T. (Eds.), 2015; ISBN: 0692457666.
16. *Smolin L.* The status of cosmological natural selection. arXiv:hep-th/0612185. 2006.
17. *Susskind L.* *The Cosmic Landscape: String Theory and the Illusion of Intelligent Design*, 1st ed.; Little, Brown. 2005. ISBN: 978-0316155793.
18. *Grey A.* The Singularity and the Methuselahry: Similarities and Differences. In *The End of the Beginning. Life, Society and Economy on the Brink of the Singularity*, Goertzel, B., Goertzel, T. (Eds.), 2015; ISBN: 0692457666.
19. *Muehlhauser L., Salamon A.* Intelligence Explosion: Evidence and Import. In *Singularity Hypotheses: A Scientific and Philosophical Assessment*; Eden, A.H., Moor, J.H., Soraker, J.H., Steinhart, E. (Eds.); Springer, 2012; pp. 15–42, ISBN: 978-3642325601.
20. *Loosemore R., Goertzel B.* Why an Intelligence Explosion is Probable. In *Singularity Hypotheses: A Scientific and Philosophical Assessment*; Eden, A.H., Moor, J.H., Soraker, J.H., Steinhart, E. (Eds.); Springer, 2012; pp. 83–98, ISBN: 978-3642325601.
21. *Schmidhuber J.* Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, 2010, vol. 2, iss. 3, pp. 230–247.

Поступила в редакцию 21.11.2017, окончательный вариант — 20.12.2017.

Computer tools in education, 2017

№ 6: 12–24

<http://ipo.spb.ru/journal>

TECHNOLOGICAL SINGULARITY IN THE CONTEXT OF THE THEORY OF METASYSTEM TRANSITIONS

Potapov A. S.¹

¹ITMO University, Saint Petersburg, Russia

Abstract

The idea of a technological singularity stems from extrapolations of the empirical curves of the acceleration of technological development. Quite often, however, from these unreliable extrapolations far-reaching consequences are made, which are subject to active criticism, and this criticism is incorrectly transferred to the very concept of technological singularity. The article attempts to separate this concept from additional assumptions and beliefs, in particular, regarding the possibility of creating artificial intelligence, as well as to highlight a number of misconceptions. For these purposes,

the theory of metasystemic transitions is involved, and the technological singularity is defined as the point at which previously observed patterns of decreasing time between metasystem transitions and reducing the time of doubling the performance of cybernetic systems within one level cease to be fulfilled. It is pointed out that before reaching this point, a number of metasystemic transitions are likely to occur, and the resulting systems will show exponential growth with a doubling period of less than six months and will exceed the complexity of existing systems for several decades.

Keywords: *technological singularity; metasystem transitions; artificial Intelligence.*

Citation: A. S. Potapov, "Technological Singularity in the Context of the Theory of Metasystem Transitions," *Computer tools in education*, no. 6, pp. 12–24, 2017 (in Russian).

Received 21.11.2017, the final version — 20.12.2017.

Alexey S. Potapov, Ph.D., Dr.Sc., professor at Computer photonics and videomatics dept., ITMO University; 197101, Russia, St. Petersburg, Kronverksky ave. 49, potapov@aideus.com



Наши авторы, 2017.
Our authors, 2017.

**Потапов Алексей Сергеевич,
доктор технических наук, доцент,
профессор кафедры КФВИ
Университета ИТМО, 197101, Россия,
Санкт-Петербург, Кронверкский пр., 49,
potapov@aideus.com**