

Автономные и самовоспроизводящиеся роботы. Искусственный интеллект и разум.

Оглавление.

РОБОТЫ ДЛЯ ОСВОЕНИЯ КОСМОСА. ОБЩИЕ ПРОБЛЕМЫ.	4
<i>Философские аспекты «разумности» машин.</i>	4
<i>Искусственная панспермия.</i>	6
РЕАЛИИ НАСТОЯЩЕГО.	9
<i>Надежность и работоспособность роботов.</i>	10
<i>О теории.</i>	11
<i>Всё только начинается.</i>	15
<i>О самовоспроизведении автоматов.</i>	17
<i>Развитие теории фон Неймана.</i>	20
ГДЕ НАЧИНАЕТСЯ ДВИЖЕНИЕ?	23
<i>Двигаем возбуждение.</i>	25
<i>Счетная логика.</i>	27
<i>Применение счетной логики.</i>	30
ОБ ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ ... И РАЗУМЕ.	32
<i>Разум и интеллект. Что о чем?</i>	33
<i>Классификации и теории интеллектов.</i>	35
<i>Искусственный разум.</i>	37
<i>Появление машинного интеллекта.</i>	39
О ТЕХНИЧЕСКОМ «СУБЪЕКТЕ Я».	40
<i>Составляющие «технического субъекта Я».</i>	41
<i>Основы формирования технического субъекта Я.</i>	43
<i>Оценка качеств машинного интеллекта и разума.</i>	44
<i>Сверхразум и суперинтеллект.</i>	46
ВРЕМЯ СОБИРАТЬ КАМНИ. ...	47
<i>Подводим итоги.</i>	49
ЛИТЕРАТУРА.	50

Не так давно в интернете появилась статья о возможности освоения нашей Солнечной системы с применением роботов [1]:

Эксперты Космического центра Кеннеди НАСА (США) описали сценарий, согласно которому, по их мнению, Солнечная система, в частности, Луна, может быть освоена автономной и расширяющейся популяцией роботов.

Каким образом планируется это начать сделать? А вот так [1]:

Моделирование, проведенное учеными, показало, что для создания успешной автономной колонии роботов достаточно в течение 20 лет доставить на Луну минимум 12 тонн полезной нагрузки. Она будет представлять собой управляемую с Земли колонию роботов. Ресурсы Луны и астероидов позволят колонии в дальнейшем перейти к автономному существованию.

Открывающиеся перспективы этого проекта захватывают:

По мнению авторов исследования, популяция роботов ввиду доступа к практически неограниченным объемам полезных ископаемых и энергии будут развиваться в геометрической прогрессии.

Эксперты прогнозируют, что спустя несколько десятилетий, без привлечения дополнительных инвестиций, объем промышленного производства на Луне превзойдет индустриальный потенциал США в миллионы раз.

Правда, в другом источнике, тут же следует предостережение [2]:

Такая стратегия может быть очень опасной, считает научный обозреватель телеканала «Культура» Александр Галкин.

... «Если же роботы будут совсем примитивными — все может сработать, однако если они будут представлять из себя самообучающуюся колонию — проблем не избежать. Вообще человечеству необходимо будет всегда контролировать такой улей, и в случае появления первых зачатков разума — уничтожать отдельные единицы», — заключил Галкин.

О, как...

Зачатки разума – уничтожать. Иначе проблем не избежать. Научный обозреватель боится появления зачатков разума у роботов уже на этапе освоения Луны.

Не очень это круто для компьютерных роботов?

Ну, ладно, отставим шутки в сторону и попробуем понять, о чем конкретно мы говорим.

В НАСА решили в течение 20 лет создать на Луне управляемую с Земли колонию роботов. И на их основе организовать добычу и переработку ценного сырья на Луне [1]. «Ресурсы Луны и астероидов позволят колонии в дальнейшем перейти к автономному существованию»...

И если тут всё получится, то и начать программу подготовки полета автоматической станции к ближайшей к нам звезде - альфа Центавра [1]. Но некоторые специалисты уже боятся начала «восстания машин» [13], если у роботов появятся зачатки разума.

Вот так, ни больше, ни меньше.

Хоть мелочь, а приятно. Управление с Земли ..., это они хорошо придумали. Насколько я помню, водители нашего «Лунохода 1» в свое время тяжело вздыхали [25]:

... мы были телеоператорами, управляя луноходом по телевидению. Только сидели в отдалении от самого лунохода в 400 тыс. км. Изображение в виде черно-белого слайда приходило с Луны на пульт водителя через каждые 20 секунд. Была естественная задержка 4,1 секунды за счет скорости распространения радиоволн от Земли до Луны и обратно. Плюс аппаратурная, техническая из-за самого сигнала по малокадровому телевидению. Представляете: пока мы получаем картинку, луноход проходит 6 метров. Потому мы должны были не только видеть, но и предвидеть последующий ход.

Конечно, сегодня кое-какие трудности уже остались в прошлом. Например, передача изображения каждые 20 секунд, это уже совсем неактуально. Сегодня изображение идет непрерывно и с хорошим качеством.

А вот «естественная задержка 4,1 секунды за счет скорости распространения радиоволн от Земли до Луны и обратно» никуда не делась. Может быть она немного изменилась с ростом качества средств связи, но не более. Но даже и 3 секунды от появления фактора риска до первой реакции на его устранение, это очень большое время.

За это время может случиться очень много.

Если на этом полностью роботизированном производстве случится что-то серьезное, например, авария с крупными поломками, то тут уже на таких скоростях управления ничего не сделаешь. И... - бригада на выезд ..., т.е. в полет на Луну по аварийному графику...

Но и то, не факт, что успеют.

А у нас пока все роботы – управляемые. На каждого «умного» сложного робота – свой оператор. Можно себе представить, какая армия операторов с Земли будет управлять этой колонией роботов на Луне.

У водителей «Лунохода-1» это выглядело так [25]:

Пять человек получилось в каждом экипаже. Потому что сделали два расчета, которые менялись через каждые два часа работы. Перерыв, смена.

Ну, так или не совсем так это будет, но, примерно, ход мыслей понятен.

Если колония, например, будет состоять из 20 роботов, то на Земле её будет обслуживать какой-то штат операторов и другой персонал. Если колония роботов будет увеличиваться, то ... операторов на них не напасешься. Расширение количества рабочих мест на Земле будет на порядок больше роста количества новых (автономных!) роботов.

Но пока другой альтернативы для управления колонией роботов на Луне прямо с Земли у нас нет. Даже если робот полностью автономный и самостоятельный, ... как мы пока это понимаем. Непонятно и, как колония роботов с таким ресурсом самостоятельности в управлении может перейти на «автономное существование»?

Но это не всё ...

На научных чтениях памяти К.Э.Циолковского в Государственном музее истории космонавтики им. К.Э. Циолковского, г. Калуга сделал доклад А.В.Колесников [3].

Вот главная мысль этого доклада:

Однако доставить живых людей на реактивных приборах к иным, даже ближайшим звёздным мирам вряд ли удастся. Иное дело, наноразмерные объекты — наследственный материал и наномшины. Это даёт основание задуматься о концепции колонизации иных миров путём транспортировки через межзвёздное пространство не «готовых», живых организмов, а лишь наследственного материала с последующей «распечаткой» биологических объектов уже непосредственно на месте назначения. То есть, космические путешественники-колонисты будущего должны будут преодолеть межзвёздные расстояния в нерожденном состоянии. При всей кажущейся фантастичности, это, вероятно, едва ли не единственная реальная возможность воплощения мечты Циолковского о расселении человечества по Галактике.

Особая роль в этой фантастической пока схеме должна принадлежать машинам, на которые ляжет основная функция первичного исследования и освоения места прибытия, а также последующей инкубации и выращивания земных организмов в условиях иной планеты. Но возможно ли создать такие машины? Некоторые соображения, основанные на мыслительных моделях, а также анализе современных технических достижений, особенно в области нанотехнологий, дают определенные основания для оптимизма.

Весьма интересно, но совершенно непонятно. Что здесь представляют из себя наномшины или нанометрические объекты? Надо бы уточнить...

Читаем [3]:

А ведь в принципе **живые организмы, по существу есть те же наномшины**, только, очень совершенные. Это становится возможным, благодаря молекулярной самоорганизации, а также **принципу построения живых организмов из самоподобных блоков по рекурсивным алгоритмам**. То есть, упрощенно говоря, построение организма осуществляется **путём многократного повторения одной и той же процедуры, запускающей саму себя на разных масштабных уровнях и с несколькими различными значениями управляющих параметров**.

После недолгих раздумий о самоподобии мы приходим к пониманию, что, может быть, это что-то на основе фракталов, иначе, фрактальные структуры. Примерно так: будем строить по уже отработанной ранее биологической технологии, из клеток. Одинаковых и ... разных. По нашему образу и подобию. Как говорит А.В.Колесников: на основе «многократного повторения одной и той же процедуры, запускающей саму себя на разных масштабных уровнях и с несколькими различными значениями управляющих параметров».

Ну пусть, как-то так ...

Если отбросить все несущественные детали, то автор начинает разговор о рукотворной панспермии [4]. Он предлагает на Земле создать какие-то «зародыши жизни» и отправить их для освоения космических просторов. Других уточняющих материалов по данному вопросу у этого автора больше не обнаружилось. То, что представлено здесь, явно урезано по возможности понимания.

И всё же ...

Что у нас есть в реальности для проведения сегодня освоения Луны и дальнего космоса с помощью колоний роботов? Есть ли реальные предпосылки для появления у этих роботов зачатков разума?

Роботы для освоения космоса. Общие проблемы.

Придется начинать с дальних подступов.

Сначала разберем то, что уже прочитали. Здесь мы имеем два разных пути применения роботов в освоении космоса. Сначала речь идет о колонии роботов на Луне, которых привезут с Земли. С помощью роботов предлагается наладить добычу полезных ископаемых, первичную переработку и получение готового продукта, пригодного для отправки на Землю космическим транспортом. Управление роботами предполагается проводить с Земли.

А вот далее идет основной момент для нас. В дальнейшем эта колония роботов как-то должна перейти на *автономное существование*. Каким же это образом?

Первое, что приходит на ум – видимо, эти роботы имеют большой запас надежности. Конструкция и вложенная программа самостоятельных действий должны обеспечить, в том числе, и полную автономность работы на длительный период. Так?

Как мы уже поняли, всё это должен обеспечить искусственный интеллект. На основе компьютерных управляющих программ. Сегодня смысл управления для всех подобных роботов примерно одинаков. Оператор принимает информацию от робота и управляет его основными действиями. А «интеллект» контролирует безопасность действий и защиту робота в сложных внешних условиях. Чем выше устойчивость работы робота, возможности его защиты в сложных условиях реальности и шире спектр выполняемых функций, тем, как считается, выше его ИИ. Но, в любом случае, основные функции управления находятся в руках оператора.

Понятно, что «интеллект» в этих системах чисто программный. Среда, в которой он функционирует – логическое пространство компьютера, управляющего его действиями.

Дополняют всё команды оператора.

Это стало настолько привычным, что для оператора уже идет постепенное слияние действительной реальности с виртуальной, игровой или моделируемой реальностью.

И погружения в эту реальность все более затягивают.

Правдивость отображения «реальности» все более возрастает. Имитаторы или симуляторы уже становятся основой как компьютерных игр, так и тренажеров для получения навыков водителей, летчиков, подводников, танкистов...

Мы уже почти не различаем разные виды компьютерного моделирования. Они стали повседневностью нашей жизни. Но в результате их применения моделирование реальности постепенно замещает саму реальность во всех аспектах деятельности человека. Боюсь, что скоро и реальные войны будут вестись на уровне компьютерных моделей. Человек здесь постепенно теряет понимание реальности происходящего.

Конечно, на этом направлении развития роботов и ИИ, человек в полной мере ощущает себя Творцом. Причем, не только самих роботов и различных автоматических устройств, но и создаваемой им виртуальной реальности, которая часто оказывается для него уже привычнее действительности.

Сегодня человек почти уверен в том, что так будет всегда. Он будет моделировать реальность и управлять, а роботы будут подчиняться человеку и превращать модель в действительность. На любом уровне «умности» роботов.

Человек планирует «поручить» и супермозгом, если он когда-нибудь будет создан.

И уж конечно, планирует широко использовать роботов всех мастей для завоевания космоса. Ближнего и дальнего.

Но пока он «рулит» только компьютерным моделированием...

Философские аспекты «разумности» машин.

Это самый разработанный комплекс проблем. Об этом уже была работа [8]. Сейчас мы немного вспомним основные направления понимания.

Первым здесь, конечно же, должен стоять вопрос, который рассматривался многими фантастами и философами: Отношение к личности робота, если он обретет разум.

А если не обретет? Ну, тогда это просто – машина, что с неё взять.

Вот, примерно с этих позиций и начнем рассмотрение...

Собственно, всё сводится к пониманию проблем взаимоотношений с искусственным Интеллектом ... или Разумом. Тут понимание сегодня достаточно размыто. Наверное, точнее будет говорить о взаимоотношениях человека и машин, способных мыслить на уровне человека. Примерно с этих позиций всё и началось.

Основных направлений здесь немного:

- **Катастрофические гипотезы «восстания машин»** [13] или **Противостояние**. Машинны вдруг осознали себя личностями, и ... стали перекраивать мир под себя, попутно уничтожая человечество. Таких апокалиптических сценариев уже десятки.
- **Дружеские «союзы людей и роботов»** как **Союз интеллектов**. Машинны осознали себя личностями, и захотели стать полноправными партнерами людей, создать единый конгломерат людей и машин. Сценариев различного развития сюжетной линии здесь тоже хватает.

Из этих философских обобщений возникло несколько формальных теорий развития ИИ до разума. Вполне практичных, надо сказать. Правда, как это у нас часто бывает, и здесь моделирование было ограничено интеллектуальной стороной развития машин.

При этом было исследовано несколько вариантов самостоятельного роста разумности машины в зависимости от моделируемых внешних условий:

- **«Квазибиологическое развитие»**.

Машина самостоятельно постигает внешний мир. Она строит себе модель внешнего мира и встраивает себя в этот мир [8]. Постижение внешнего мира идет по проторенному пути биологических организмов. Методом проб и ошибок.

Хороший метод, но он не предполагает других ограничений, кроме физических. Разрешено всё, что не запрещено. Для машины в этой системе, человека, как равноправного субъекта – нет. Понятно, что это прямой путь к «восстанию машин».

Есть другой путь.

- **«Путь самоосознания»**.

Машина исследует себя. Свои способности и возможности без прямого соотнесения с внешним миром. Задачи машине выдает человек, машина их решает. В результате рационального соотнесения результатов решений разных задач постепенно вырабатывается «мышление», усиливающее уже имеющийся ИИ до человеческого.

Тоже неплохой путь саморазвития. Развитие идет на основе увеличения информации о «гипотетическом» для машины мире, в котором она создает и разрушает те или иные модели достижения результата в поставленной задаче. Здесь мы прямо ощущаем путь построения искусственного «чистого разума». Это Машинный Разум, чьи решения не ограничены рамками реальности. Разум «вообще».

Примерно такой вид разума человек видит в глобальных системах ИИ, основанных на математических нейронных сетях в глобальном масштабе физического Интернета. Например, проект электронного Всемирного Мозга, начавшийся еще во времена публикации «Суммы технологии» Станислава Лема [14].

И наконец

- **«Путь озарений»**.

Робот вдруг «прозревает». Он вдруг начинает понимать чувственные состояния людей и воспринимать их команды в этом понимании, а не только как формальные приказы к исполнению. Машинный Интеллект превращается в Разум.

Это наиболее часто встречающийся переход от ИИ к машинному разуму, встречающийся в фантастических романах и различных философских исследованиях. Почти стандартный пример «очеловечивания» машины. Точно так же с давних времен идет очеловечивание всех сказочных в обычной жизни «неживых» персонажей. Например, вспомним сказку «Гуси-лебеди», где печка и яблоня разговаривали и помогали героям сказки ...

Наверное, можно найти еще несколько подходов к развитию ИИ в различных исследованиях.

Но ..., как-то не вдохновляют нас уже все эти вариации.

Сегодня мы много знаем и понимаем из того, что полвека назад ученые и представить себе не могли. Но тем не менее, поток статей на тему философского осознания проблемы машинного разума не высох. Работы пишутся, публикуются, читаются, порождая споры и длинные размышления. Вот пример.

Ну, и наконец...

Практика взаимодействия человека с машиной постепенно начинает осознаваться как эволюционный процесс. Машины усложняются, возможности их растут.

И возможно, что где-то впереди то время, когда человек уже будет не в состоянии контролировать все действия «умных машин» вокруг себя. *Он вынужден будет, или полностью положиться на их возможности для продолжения своего существования, или вынужден будет ... умереть от невозможности что-то изменить в этой модели машинного мира...*

Некоторые ученые сегодня понимают наше настоящее, как конечный этап пути развития машин, до точки *технологической сингулярности*:

Технологическая сингулярность — гипотетический момент, по прошествии которого, по мнению сторонников данной концепции, технический прогресс станет настолько быстрым и сложным, что окажется недоступным пониманию^{[1][2][3]}, предположительно следующий после создания искусственного интеллекта и самовоспроизводящихся машин, интеграции человека с вычислительными машинами, либо значительного скачкообразного увеличения возможностей человеческого мозга за счёт биотехнологий.

Вернон Виндж считает, что технологическая сингулярность может наступить уже около 2030 года^{[4][5]}, в то же время Рэймонд Курцвейл даёт 2045 год^[6]. На Саммите Сингулярности в 2012 году Стюарт Армстронг собрал оценки экспертов, медианное значение этой выборки составило 2040 год^[7].

Оказывается, не так уж долго осталось нам ждать момента проверки правильности выбора технократического пути для человеческой цивилизации. Недавно футуролог и идеолог сингулярности Рэй Курцвейл снова подтвердил свой прогноз:

По его прогнозам, сингулярность наступит в 2029 году. Большинство экспертов считает, что уровень интеллекта машин и человека сравняется не раньше 2045 года — если это произойдет в принципе. Но Курцвейл настроен оптимистично — он уверен, что за слиянием человека с искусственным интеллектом будущее. Люди перестанут мыслить линейно, и это приведет к небывалому прежде прогрессу.

Вот так. То ли до «Судного дня», то ли до великого будущего ... остались считанные десятилетия.

Искусственная панспермия.

А если машины себя личностями и не осознают? Если это просто автоматы различного назначения и размеров? В этом случае все направления их глобального развития моделируются в области освоения космоса. В основном, дальнего.

Если уж мечтать, то ни в чем себе не отказывать...

И разработанные философами различные сценарии освоения космоса машинами часто оказываются вполне апокалиптичны. Удивительно, но, глобальность катастрофы и катаклизмы зависят от размеров роботов, осваивающих космос.

Начнем мы с более или менее мирных процессов...

Для распространения роботов во вселенной был разработан сценарий «зонда фон Неймана»:

Зонд фон Неймана - характерный пример гипотетической концепции, основанной на работе американского математика и физика венгерского происхождения Джона фон Неймана (нем. *John von Neumann*).

Фон Нейман тщательно исследовал идею самовоспроизводящихся машин, которых он назвал «универсальными сборщиками», и которые часто упоминаются как «машины фон Неймана». Хотя фон Нейман никогда не рассматривал свою работу в приложении к идее космического корабля, позднее теоретики проделали

это. Идея самовоспроизводящегося корабля применялась (в теории) для решения нескольких определённых задач, и детальное развитие этого понятия в применении к исследованию космоса известно как **зонд фон Неймана**.

И как всегда бывает в таких случаях, мысль ученых сразу пошла дальше.
До «искусственной панспермии» на уровне отдельных клеток или эмбрионов.
Тут есть своя гипотеза:

Самовоспроизводящий корабль-сеятель.

...Такой корабль может хранить генетические образцы жизненных форм своей родины, возможно даже, своих создателей. При обнаружении обитаемой или пригодной для терраформации экзопланеты, он пытается воспроизвести эти жизненные формы — из сохранённых эмбрионов (см. Эмбриональная колонизация космоса) или сохранённой информации с использованием молекулярной нанотехнологии для «сборки» зигот с различной генетической информацией из местного сырья. Детальное описание корабля-сеятеля (с ограниченными возможностями самовоспроизведения) можно найти в рассказе «Дальний прицел» Вернора Винджа.

Такие клетки или эмбрионы надо спроектировать, создать, проверить жизненный цикл получаемых искусственных организмов на много поколений и лишь потом развозить по вселенной, как биоматериал для массового воспроизводства.

Конечно, сегодня успехи биотехнологий впечатляющие. Возникли новые науки и получили новое дыхание уже созданные ранее: Биоинформатика. Бионика. Биоремедиация. Искусственный отбор. Клонирование. Гибридизация (биология). Генетическая инженерия.

При таком бурном развитии новых наук неплохо бы помнить и о биологической безопасности этих исследований. Как показали первые результаты применения ГМО, количество проблем от этого только увеличилось. Правда, геновая инженерия делает лишь первые шаги. У специалистов по биотехнологиям и у генетиков здесь еще целина непаханая.

Правда, пока никакие биотехнологии не дали необходимых материалов для массового распространения на других планетах. Этот сценарий сегодня осуществим только в объеме клеток организмов, уже живущих на Земле.

Тогда возникает вопрос устойчивости их воспроизведения. Если на других планетах нет вирусов и бактерий, которые есть на Земле, то и защита от них, заложенная в ДНК, постепенно перестанет работать. И их быстрое вымирание гарантировано.

Или мы и вирусы с бактериями будем забрасывать в дальний космос?

А может, что-то еще более мелкое?

У нас есть следующий, уже вполне апокалиптический сценарий - «серая слизь»:

Серая слизь (англ. *grey goo*) — гипотетический сценарий конца света, связанный с успехами молекулярных нанотехнологий и предсказывающий, что неуправляемые самореплицирующиеся нанороботы поглотят всё доступное им вещество Земли^{[1][2]}, выполняя свою программу саморазмножения (или вещество биосферы — биомассу, данный сценарий известен под названием «экофагия»^[3]).

Впервые самореплицирующиеся машины были описаны математиком Джоном фон Нейманом, и поэтому иногда называются машинами фон Неймана. Впервые термин «серая слизь» был применён в 1986 году пионером нанотехнологий Эриком Дрекслером в книге «Машины создания»,^[4] где при описании подобного сценария Дрекслер предостерег, что «Мы не можем позволить себе определённого рода аварии с реплицирующимися ассемблерами»^[3] В 2004 году Дрекслер сказал, что из-за разразившейся шумихи жалеет о том, что придумал этот термин.^[5]

Как правило, термин используется в популярной прессе или научной фантастике. В худших постулируемых сценариях, требующих больших, способных к космическим полётам машин, материя вне Земли также обращается в серую слизь. Под этим термином понимается большая масса самовоспроизводящихся наномашин, которые не обладают структурой в большом масштабе, которая может оказаться, а может и не оказаться подобной слизи. Бедствие случается по причине преднамеренного включения Машины судного дня или от случайной мутации в самореплицирующихся наномашинах, используемых в других целях, но созданных для работы в естественной среде.

Эрик Дрекслер предупреждает, с такими нанороботами шутки плохи.

Эта опасность страшнее чумной пандемии будет. И ядерной войны. Одна утечка, и ... конец всему. Вообще - всему. Тихо и незаметно.

Как мы видим, научная и философская мысль здесь выделила два основных пути «искусственной панспермии» от человечества Земли. В одном случае, техническим агентом передачи информации и частицей «жизни» выступает наноробот, а в другом случае – биологическая молекулярная машина. Эрик Дрекслер вполне справедливо не стал различать биологические «ассемблеры» и механистические нанороботы.

Сегодня эти направления объединились на уровне *наноассемблеров*:

Наноассемблер (*нано* — 10^{-9} и *англ.* assembler — *сборщик*) — это разрабатываемое устройство наноразмеров, способное собирать из отдельных атомов или молекул сколь угодно сложные конструкции по вводимому в них плану. Термин был введён Эриком Дрекслером и в настоящее время широко используется в фантастике в стиле нанопанк. Первые работы в этом направлении были сделаны ещё в 1986 году, когда компания IBM с помощью туннельного сканирующего микроскопа (размеры которого очень далеки от нанометра) выложила на металлической пластине свой логотип отдельными атомами ксенона.^[11]

Наноассемблер является частным случаем не созданной на данный момент нанофабрики — более крупного устройства, предназначенного для сборки объектов из отдельных атомов. По мнению Дрекслера, наноассемблер можно будет запрограммировать как репликатор: устройство, способное производить свои собственные копии. Более простым, чем наноассемблер, устройством может быть фабрикатор, способный создавать наноконструкции под внешним управлением.

Начиная с 2007 года, Британский совет по инженерным и физическим исследованиям финансирует разработки молекулярных ассемблеров, подобных рибосомам. Очевидно, что молекулярные ассемблеры в этом ограниченном понимании точно возможны. Проект технологической дорожной карты, возглавляемый институтом им. Баттелла и расположенный в нескольких национальных лабораториях США, исследовал область технологий производства с атомарной точностью, включая и ближайшие, и отдалённые перспективы разработки программируемого молекулярного производства. Этот отчёт был выпущен в декабре 2007 года.^[12]

Ну, мы поняли...

Наноассемблер способен собирать из отдельных атомов сколь угодно сложные объекты, конструкции, системы. Он и самого себя сможет реплицировать... когда-нибудь.

А пока..., с помощью огромного электронного микроскопа и уникальных технологий мы немножко поработали с отдельными атомами вещества, и ... всё. Далее идет уже чистое моделирование возможных вариантов развития. Ничего еще нет, кроме красивой модели применения этого направления. Но, ведь до сборки атомов мы всё же добрались, это факт.

И всё же, с этими успехами говорить о завоевании космоса на уровне любых нано и микророботов несколько рановато...

Если все предыдущие варианты панспермии в виде больших самовоспроизводящихся роботов, молекулярных машин, наноассемблеров, клеток или эмбрионов, предусматривали их конкурентное совмещение с имеющимися формами жизни дальних планет или безразличие к ним, то есть проекты, настораживающие своей начальной установкой для распространения...

Вот например Берсеркер:

Более пугающая разновидность самовоспроизводящего космического корабля называется **Берсеркер**. В отличие от безопасной концепции зонда, Берсеркеры запрограммированы обнаруживать и уничтожать все обнаруженные жизненные формы и населённые экзопланеты.

Этот термин происходит из серии романов Фреда Саберхагена, которые описывают непрекращающуюся войну между человечеством и такими машинами (см. Берсеркер). Саберхаген показывает (посредством одного из персонажей), что корабли-берсеркеры в его романах не являются собственно машинами фон Неймана, но сложный комплекс машин-берсеркеров, включая автоматические верфи, составляет машину фон Неймана. Это снова приводит к понятию *экологической системы машин фон Неймана*, или даже *роя машин фон Неймана*.

Это одно из направлений проведения искусственной панспермии. В этом случае - сначала очистим планету от прошлой жизни, потом начнем выращивать свою. Как первый этап для *терраформирования*:

Терраформирование (*лат.* terra — земля и *forma* — вид) — изменение климатических условий планеты, спутника или же иного космического тела для приведения атмосферы, температуры и экологических условий в состояние, пригодное для обитания земных животных и растений. Сегодня эта задача представляет в основном теоретический интерес, но в будущем может получить развитие и на практике.

Термин «терраформирование» был придуман Джеком Уильямсоном в научно-фантастической повести, опубликованной в 1942 году в журнале Astounding Science Fiction^[1], хотя идея преобразования планет под земные условия обитания присутствовала уже в более ранних произведениях других писателей-фантастов.

Такой вариант искусственной панспермии, по крайней мере, уже не скрывает конечную цель идеи освоения космоса - *формирование условий для жизни человека на всех возможных планетах и спутниках ... любыми средствами.*

Как мы видим, показан и вполне четкий способ осуществления миссии - *продолжение политики конкистадоров при колонизации новых обитаемых земель.*

И наконец, ещё немного о кораблях-сеятелях:

Такие корабли могут быть кораблями терраформации, подготавливая миры для последующей колонизации другими кораблями, или — если они запрограммированы для воссоздания, выращивания и обучения представителей создавших их видов — воспроизведения самих колонизаторов.

Надо заметить, эта модель терраформации и колонизации не обязана быть «автоматизированной». Пилотируемые межзвёздные колониальные корабли могут следовать похожей модели — и рассматриваться как комбинированный корабль (зонд фон Неймана/сеятель), в котором репликация выполняется с помощью живых обитателей. Некоторые сторонники космической среды обитания утверждают, что цивилизации, использующей такой подход, планеты совершенно не нужны.

Вот он, заключительный аккорд...

Модель панспермии плавно перетекает в модель «космического человека», которому планеты вообще не нужны. Его дом — звездолет, а родина — космос. Планеты он, видимо, рассматривает только как новые кладовые полезных ископаемых и биозапасов.

Для себя любимого...

Реалии настоящего.

Даже если не очень критично относиться к предлагаемым вариантам и конечным целям освоения космоса с помощью роботов, то нам все же надо понять: Зачем нужна космическая панспермия человечеству, которое еще и от собственной планеты оторваться толком не может?

Возможно, ученые что-то ещё могут придумать за ближайшие полвека, и экспансия космоса роботами всё же начнется. Тогда - вполне очевидным способом...

Полетят «корабли-сеятели».

Они начнут высаживать десанты из крупных роботов на подходящие планеты. А уж эти крупные роботы, может быть, «человеческой» размерности, начнут создавать процессы производства нанороботов, как компонентов преобразования планеты для принятия человека.

Да, наверное, это и будет один из вариантов той самой «серой слизи».

Как цианобактерии в свое время создали на Земле кислородную атмосферу, так массы нанороботов должны создать нечто подобное на других планетах. И умереть. Оставив после себя планету, пригодную для высадки на ней человека, без скафандра и защитных масок.

Или они будут продолжать трудиться, создав новую Жизнь, на другой основе, иначе полученные условия на планете долго сохранить не удастся. И эта новая «нано» Жизнь должна быть дружественна человеку, иначе зачем всё это?

Цель понятна. А сколько времени потребуется для приведения одной планеты в должное состояние? Сотни, тысячи, миллионы лет?

Нашей осмысленной человеческой цивилизации на Земле, вместе с периодом её заселения человеческими племенами, около 300 тыс. лет, а истории взаимоотношений народов не более 20 тыс. лет. Что будет с нашей цивилизацией через 1 тыс. лет?

С учетом нарастания скорости прогрессивного развития это представить невозможно.

Сегодня моральное старение техники уже часто наступает прежде, чем эта техника дошла до потребителя. Крупные корпорации вынуждены задерживать выпуск новых образцов техники

только потому, что ещё не успели окупить продукцию, лишь недавно запущенную в производство. Так происходит, например, с выпуском блоков памяти на мемристорах [12].

Человек уже не успевает за прогрессом техники. Можем ли мы представить себе, как мы будем осваивать дальний космос через тысячу лет? На чем и как будем передвигаться?

А тем временем по космосу все еще будет расползаться армада посланных когда-то давно кораблей-сеятелей с роботами на борту и, возможно, с запасами «серой слизи», опасной больше, чем мы можем себе это сегодня даже предположить...

Если же пока отложить в сторону околосказочные сюжеты космических панспермий и вселенских катаклизмов от непродуманных действий человека, и заняться реальными проблемами тех же роботов для космических исследований, работ на астероидах и планетах солнечной системы, то окажется, что для осуществления задуманного здесь есть огромные проблемы.

Вполне реальные проблемы уже сегодняшнего дня.

Надежность и самовосстановление роботов.

Человек находится на перепутье сложной проблемы.

Что выбрать? Создавать специализированных роботов для разных условий работы, или стремиться делать универсальных, пригодных к любым условиям?

Любой специалист по роботам скажет сразу – вопрос некорректен.

Нужны разнообразные роботы, от роботов манипуляторов, специальных летающих роботов и глубоководных роботов, до аналогов животных и человекоподобных роботов. Сегодня на горизонте появились микро и нанороботы со своим кругом задач. Слишком широк круг планируемых работ, чтобы все мог выполнять один универсальный робот.

А с другой стороны? Спектр задач увеличивается, условия самостоятельности действий робота в разных средах постоянно расширяются. Это приводит к усилению универсальности когда-то очень специализированных роботов. И чем дальше, тем больше. Роботы начинают походить на какие-то биологические прототипы. Так диктует необходимость. Природа давно отработала разнообразные кинематические схемы движения, защитные средства для внутренних устройств автомата, датчики контроля поверхности автомата и системы контроля внешней среды...

Мимо этого информационного массива давно готовых, проверенных технических решений пройти невозможно. И роботы начинают закрываться защитными экранами от удара, гибкими оболочками от попадания жидкости и пыли во внутренние системы, покрываться сенсорными датчиками, регулирующими скорости перемещения и усилия давления при движении, контроля созданного давления на переносимые грузы... и пр., и пр., и пр.

Это и есть движение к универсальности роботов.

Конечно, какое-то разнообразие роботов будет оставаться всегда, но оно всё время будет стремиться сжаться до нескольких основных вариантов универсальности, до системы из пары десятков типоразмеров универсальных роботов, от микро и нанороботов, до макроразмерных роботов, в том числе и межзвездного транспорта. Правда, будет это, ой, как не скоро.

Но эту, пусть пока и очень далекую цель надо держать в голове уже сегодня.

Теперь немного об автономности и поддержании работоспособности.

Этот вопрос при рассмотрении освоения космоса роботами наука пока старательно обходит стороной. Ну, как-нибудь, может быть, само получится ... когда-нибудь.

Далее обычно эта тема не развивается. Почему?

Нет пока никаких реальных мыслей в этом направлении.

Технические способы увеличения надежности работы тех или иных сложных механизмов мы, в общем, знаем. Если мы говорим о модульных сложных системах управления, то прежде всего на ум приходит дублирование модулей, а то и комплектов целиком. Дублирование позволяет при возникновении неисправности в основном канале системы управления быстро

перейти с неисправного канала управления на резервный, и таким образом сохранить работоспособность машинного комплекса. Дублирование или резервирование до последнего времени широко используется в машинных комплексах управления и вычисления, требующих повышенной надежности для сохранения своей работоспособности.

Ну, а если дублирование на уровне целых каналов управления или отдельных блоков уже нецелесообразно? Сегодня вся вычислительная машина, за исключением нескольких её частей или блоков, уже имеет высокие показатели надежности.

Кроме того, мы подошли к осознанию и такого факта, что каждая выполняемая машиной задача должна иметь свой индивидуальный набор вычислительных модулей. И тогда мы вспоминаем о реконфигурируемых вычислительных системах (РВС) [37, 38]. Основы РВС были сформулированы академиком А.В.Каляевым в середине 80-х годов прошлого века. Вычисления [37] проходят примерно как на рис.1.

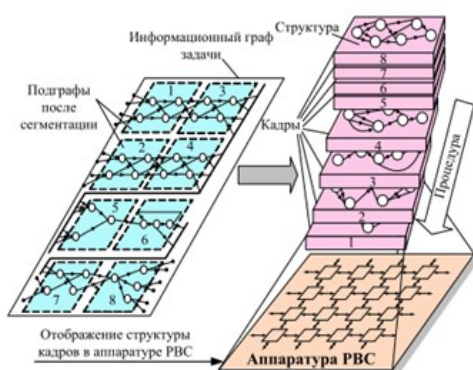


Рис. 1. Процесс решения задачи в РВС.

Сегодня такие реконфигурируемые модули однородного нейроморфного пространства создаются на базе ПЛИС. Например здесь. Или на базе мемристоров [12], как здесь. Вместе с повышением надежности вычислений уже можно говорить и о гибкой схеме их проведения, если это позволяет управляющая программа.

Несколько иной путь повышения надежности машины дают автоматы Л.Ф.Мараховского. Это выбор использования компонентов схемы в зависимости от их исправности. Здесь же [42] показаны и схемы многоуровневой оперативной памяти цифрового автомата.

Ну хорошо, электронные системы и другие основные системы машины как-то можно поддерживать в рабочем состоянии за счет дублирования и реконфигурации блоков в составе машины. Но, космос поставил перед нами такие проблемы, которые раньше просто не приходили нам в голову. Например, понимание *проблемы реальной самостоятельности, самовосстановления и самовоспроизводимости роботов без участия человека*.

Решена эта проблема? Пока – нет.

Для нас сегодня это две разные проблемы. Проблема надежности работы и проблема самовосстановления. Первую, как уже стало понятно, мы решаем прямым повышением надежности всех компонентов устройства, а далее резервированием и реконфигурацией. Проблема самовосстановления, а тем более самовоспроизведения ... пока вообще не рассматривается. Прежде всего потому, что человек даже не рассматривает возможности работы машины без человеческого пригляда, без ремонта и восстановления в случае крупной поломки.

И только космос свел эти проблемы в одну - *обеспечение длительного существования робота на очень большой период времени. На десятки, сотни лет и более*.

В этом случае ни резервирование ни реконфигурация уже не помогают. Здесь всё решает возможность самовосстановления и полного самовоспроизведения робота.

Конечно, самовосстановление входит в самовоспроизведение как составная часть.

Потому, далее мы будем говорить только о самовоспроизведении.

Как о наиболее сложной и общей проблеме...

О теории...

Как мы уже знаем, проблема самовоспроизведения робота соприкасается с проблемой самовосстановления. В живой природе самовосстановление имеет название регенерации.

А самовоспроизведение, это понятно - размножение.

Знаем мы что либо подобное в технике вокруг нас? Ну, хоть пока только в теории?

Да, из теории машин и автоматов известно понятие - самовоспроизводящиеся машины. И вот здесь, например, читаем [3]:

Проблема самовоспроизводящихся автоматов ставилась в свое время еще Джоном фон Нейманом. При упоминании о самовоспроизводящихся машинах не следует непременно представлять себе робот-манипулятор, который каким-то образом вытачивает и собирает свои собственные детали, собирая из них свои копии.

Вот к этому понятию мы сейчас и присмотримся.

Чуть ранее мы рассматривали сценарии искусственной панспермии. В космосе от планеты к планете путешествуют какие-то корабли-сеятели и высаживают на поверхность планет десант из наноассемблеров, нанороботов, молекулярных машин или клетки, зиготы и т.д., например, для начала процесса искусственного терраформирования. Это когда меняется газовый состав атмосферы, климатические условия и пр. и пр., для возможности начала жизни здесь человека. В идеальном случае, без скафандра и кислородной маски.

Заметим, все высаживаемые на другие планеты автоматические машины должны иметь возможность саморепликации, самокопирования, самовоспроизведения....

Термины разные а основная суть одна – размножение.

Есть сегодня у нас такая техника? Нет.

Ну, бог с ней, с техникой ...

Нет у нас пока и хорошей теории самовоспроизводящихся роботов. Нет нормальной механистической теории, позволяющей разработать и сделать любой самовоспроизводящийся робот любой сложности и с любым комплексом выполняемых задач.

Если отбросить философские, в основном, гуманистические, и эмоциональные разработки этических сторон этого процесса по отношению к роботам, отбросить большой объем технократической фантастики, рисующей картины как светлого будущего, так и апокалипсиса, связанных с появлением самовоспроизводящихся роботов, что останется в «сухом остатке»?

Только работа Джона фон Неймана "Теория самовоспроизводящихся автоматов" [5], на которую и ссылаются все пишущие об этой проблеме.

О чем же размышлял когда-то Джон фон Нейман?

Приведем слова составителя книги [5] А. Бёркса:

Этот вопрос объединяет понятия из различных областей: биохимии, техники, логики. Он очень важен для нового (возникшего примерно в 1948 г.) направления — теории автоматов, или кибернетики. Главными создателями этого направления были Джон фон Нейман и Норберт Винер. Хотя они и находились под взаимным влиянием, но все же их подходы и интересы были различными. Фон Нейман назвал свой вариант «теорией автоматов», а Винер — «кибернетикой».

Теория автоматов — это наука об основных принципах, общих для искусственных автоматов (цифровые вычислительные машины, аналоговые вычислительные машины, управляющие системы) и естественных автоматов (нервная система человека, самовоспроизводящиеся клетки, организмы в эволюционном аспекте).

... В планы фон Неймана входило создать систематическую теорию, математическую и логическую по форме, которая упорядочила бы понятия и принципы, касающиеся структуры и организации естественных и искусственных систем, роли языка и информации в таких системах, программирования и управления такими системами. Теория автоматов лежит на стыке разных дисциплин, объединяет различные подходы (с точки зрения логики, теории связи, физиологии), но в конце концов ей предстоит стать отдельной самостоятельной дисциплиной.

... Винер понимал важность цифровых вычислительных машин для кибернетики, а фон Нейман хотел расширить сферу теории автоматов, чтобы включить в нее непрерывные механизмы. Он предполагал построить непрерывную модель самовоспроизведения, основанную на нелинейных дифференциальных уравнениях в частных производных, описывающих диффузионные процессы в жидкости.

С рассказом о перспективах, в общем – всё. Далее идут уже более практические обоснования появления теории. Они очень интересны сегодня [5]:

В самом деле, между проблемой понимания самовоспроизведения естественной системы на генетическом и биохимическом уровнях и логической, автоматико-теоретической проблемой самовоспроизведения, сформулированной фон Нейманом, существует глубокое различие. Я сейчас постараюсь показать, что это различие не столь велико, как кажется. Для этого я сделаю краткий обзор основных результатов **части I («Теория и**

организация сложных автоматов») и части II («Теория автоматов: конструирование, воспроизведение, однородность») настоящей книги.

Оценим сказанное.

Как мы видим, вся теория должна была состоять из двух частей:

- часть I «Теория и организация сложных автоматов»;
- часть II «Теория автоматов: конструирование, воспроизведение, однородность».

Начнем смотреть материал...

Лучше сразу краткое содержание I части [5]:

Умозрительные и численные методы математики. Роль последних в прикладной математике и математической физике. Роль в чистой математике. Положение в анализе. Численная обработка для эвристических целей.

Различные формы численного подхода: аналоговая и цифровая.

Аналоговая процедура: использование физического эксперимента для вычислительных целей; аналоговые вычислительные машины.

Цифровая процедура: вычисления вручную; простые машины; полностью автоматизированные вычисления.

Современное состояние вычислительной техники. Роль аналоговых и цифровых вычислительных машин в настоящее время. Вопросы скорости, программирования и точности. Понятие элементарной операции в вычислительной машине. Ее роль в аналоговых вычислительных машинах и в цифровых машинах. Несколько слов о компонентах аналоговой вычислительной машины, цифровой машины.

Релейный орган. Основные типы: электромеханическое реле; электронная лампа; другие возможные типы релейных элементов.

Измерение длины или сложности вычисления. Логические и арифметические операции. Линейные и нелинейные арифметические операции. Роль количества операций умножения. Устойчивость статистических характеристик различных частей математики. Особая роль анализа.

Различные характерные уровни длины или сложности. Объем характерных задач, решаемых на автоматической цифровой вычислительной машине.

Требования к точности.

Требования к памяти: измерение емкости памяти; важнейшие характеристики памяти: время выборки и емкость; причины иерархической организации памяти; реальные требования к памяти автоматической цифровой машины.

Вход — выход: основные из имеющихся способов.

Понятие равновесия: равновесие скоростей различных компонентов, равновесие емкостей памяти на различных ступенях иерархической лестницы и скоростей. Баланс между скоростью и точностью. Баланс между скоростью, емкостью памяти и объемом программирования.

Термодинамические аспекты понятия равновесия. Термодинамические аспекты, связанные с емкостью памяти. Необходимость количественной теории вместо применяемых сейчас эмпирических методов. Предварительные замечания о надежности и ошибках.

Напомню, разговор идет о теории сложных автоматов...

Здесь же мы видим скорее вопросы по конструированию «искусственного мозга» на основе вычислительной машины, чем «умного» автомата «в целом».

Ну, хорошо, может быть мы что-то не поняли..., заглянем во вторую часть [5]:

Основные вопросы. Наше исследование автоматов будет проходить по двум важным и связанным друг с другом направлениям: логическому и конструктивному. Оставаясь в рамках указанного выше ограничения, мы рассмотрим пять главных вопросов, которые по крайней мере в начальной фазе развития изучаемого предмета являются, пожалуй, центральными.

(А) **Логическая универсальность.** При каких условиях некоторый класс автоматов логически универсален, т. е. может выполнять все выполнимые операции с помощью лишь конечного (но произвольно большого) числа средств? При наличии какого вспомогательного оборудования — переменного, но по существу стандартного — одиночный автомат будет логически универсальным?

(В) **Конструируемость.** Можно ли сконструировать, т. е. скомпоновать и построить из (надлежащим образом определенного) «сырья», один автомат с помощью другого? Или, формулируя вопрос с другого конца и расширяя его: какой класс автоматов можно сконструировать с помощью надлежащим образом заданного автомата? Здесь допускается переменное, но по существу стандартное оборудование (как в (А)).

(С) **Конструктивная универсальность.** Конкретизируем второй вопрос из серии (В): может ли какой-нибудь надлежащим образом заданный автомат быть конструктивно универсальным, т. е. способным сконструировать в смысле (В) (располагая подходящим, но по существу стандартным оборудованием) любой другой автомат?

(D) **Самовоспроизведение.** Сузим вопрос (С): может ли какой-нибудь автомат сконструировать в точности такой же автомат? Можно ли его заставить выполнять также и другие задачи, например сконструировать еще и другие, заранее указанные автоматы?

(Е) **Эволюция.** Объединим вопросы (С) и (D): может ли конструирование автоматов автоматами прогрессировать от более простых типов к более сложным? Может ли эта эволюция двигаться от менее эффективных (при некотором разумном определении эффективности) автоматов к более эффективным?

Вот они, правильные вопросы по теме!

Теперь осталось понять, как они решаются [5]...

Характер ответов. Ответ на вопрос (А) известен ^х). На вопросы (В) — (D) будут получены утвердительные ответы ²). На вопрос (Е) не позволяет дать подобный ответ неоднозначность в его формулировке, в особенности в определении понятия эффективности. Далее мы рассмотрим вопросы (А) — (Е), более жестко определив, что представляет собой автомат. Это мы сделаем с помощью понятия, которому больше всего подходит название *кристаллическая упорядоченность*. Этот дополнительный результат столь же важен и поучителен, как и возможность ответить на вопросы (А) — (D) {и до некоторой степени на вопрос (Е), см. выше}.

Оставшуюся часть гл. 1 мы посвятим предварительному эвристическому обсуждению вопросов (А) — (Е). В гл. 2 разработаем специальную модель, в рамках которой сможем рассмотреть вопросы (А) — (D) со всеми подробностями и достаточно строго. В гл. 3 дадим анализ другой, более естественной, но технически менее удобной модели. Глава 4 посвящена дополнительным эвристическим соображениям, которые целесообразнее привести после детально разработанных конструкций гл. 2 и 3.

Модели самовоспроизведения, рассматриваемые фон Нейманом. В предыдущем абзаце представлен план, намеченный фон Нейманом. К сожалению, ему не удалось осуществить свои намерения полностью. Чтобы понять сам план и ссылки на него, надо иметь некоторое представление о рассмотренных им различных моделях самовоспроизведения. Эти модели мы сейчас кратко опишем. По необходимости многое из этого описания основано на личных беседах с теми, с кем фон Нейман обсуждал свои модели.

Всего фон Нейман рассмотрел 5 моделей самовоспроизведения: кинематическую, клеточную, модель типа «возбуждение — порог — усталость», непрерывную и вероятностную модели.

Кинематическая модель связана с геометрическими и кинематическими проблемами движения, контакта, расположения, прикрепления и разделения. Вопросы силы и энергии в ней игнорируются. Основные элементы кинематической модели: логические (переключательные) и запоминающие (задержки) элементы (хранят и обрабатывают информацию); стержни (обеспечивают структурную жесткость); чувствительные элементы (чувствуют объекты в окружающем пространстве); кинематические (мышцеподобные) элементы (двигают близлежащие объекты); соединяющие и разделяющие элементы (осуществляют соединения и разъединения элементов). Кинематическая модель самовоспроизведения описана в лекции 5 части I. Там отмечалось, что фон Нейман начал работать над ней уже в 1948 г.

Вторая модель самовоспроизведения, принадлежащая фон Нейману, — **клеточная модель**. Ее создание было стимулировано Уламом, который в процессе обсуждения кинематической модели высказал предположение, что логический и математический анализ легче будет производить в рамках клеточной структуры, чем этой кинематической модели ^х). В клеточной модели самовоспроизведение осуществляется в неограниченном пространстве, разбитом на клетки, причем все они содержат один и тот же автомат. Фон Нейман называл это пространство «кристаллической упорядоченностью», «кристаллической средой», «гранулированной структурой» и «клеточной структурой» ¹). Мы будем употреблять термин **клеточная структура**.

Для самовоспроизведения можно использовать много различных типов клеточных структур. Фон Нейман выбрал бесконечный массив квадратных ячеек. В каждую ячейку он поместил один и тот же конечный автомат с 29 состояниями. Каждая ячейка непосредственно сообщается с 4 примыкающими соседями, причем информация приходит с задержкой по крайней мере в 1 единицу времени. Фон Нейман разработал эту модель в рукописи, которая составляет часть II настоящей книги. Относительно этой рукописи г-жа Клара фон Нейман писала мне: «Я абсолютно уверена в том, что он начал работать над ней в конце сентября 1952 г. и кончил где-то в конце 1953 г.» Насколько мне известно, после 1953 г. фон Нейман почти ничего или совсем ничего не делал с этой рукописью.

Рукопись, оставленная фон Нейманом, содержала две законченные главы и длинную, но незавершенную третью главу. Глава 1 рукописи совпадает с гл. 1 настоящей книги. В главе 2 рукописи формулируются правила переходов, которым подчиняется клеточный автомат с 29 состояниями; это гл. 2 настоящей книги. В незаконченной главе 3 делаются основные шаги на пути создания клеточного самовоспроизводящегося автомата; их описание составляет 3 и 4 гл. Фон Нейман так и не закончил конструирование своего клеточного самовоспроизводящегося автомата; в гл. 5 я показываю, как это сделать.

Пока достаточно о том, что мы хотели узнать...

С самого начала и Дж. фон Нейман, а потом и составитель книги [5] А. Бёркс, пошли проверенным путем. Краткое описание проблемы, плавный поворот на математику и вычислительные проблемы того времени и ..., всё. Общая теория автоматов» плавно перешла в прикладную теорию клеточных автоматов. Остальное - вскользь...

Да, Дж. фон Нейман немного рассматривает «кинематические автоматы» с точки зрения возможности их самовоспроизводства, но в подробности воспроизводства не вдаётся.

И получается, что заявленная им «теория сложных автоматов» так и не сформирована.

Он ограничился математическими приближениями. И вопрос совсем не в том, что не смог или не успел. Он и не мог тогда ничего сделать, кроме разработки математического аппарата для взаимодействия клеточных автоматов в виде картинок на экране монитора.

Необходимые знания стали появляться значительно позже.

Всё только начинается...

Если вспомнить вопросы, поставленные Дж. фон Нейманом в работе [5], то естественно хочется понять, что сделано со времен написания работы. И предложить что-то свое.

Попробуем разобраться...

Смотрим основные вопросы «теории автоматов» Дж.фон Неймана [5]:

- (A) *Логическая универсальность.*
- (B) *Конструируемость.*
- (C) *Конструктивная универсальность.*
- (D) *Самовоспроизведение.*
- (E) *Эволюция.*

Конечно, многие проблемы, затрагиваемые при рассмотрении этих вопросов со времен написания работы Дж. фон Нейманом уже нашли какое-то техническое решение.

И все же ...

Давайте чуть более внимательно пройдемся по пунктам:

(A) **Логическая универсальность.** Так, как её понимал фон Нейман – не решена. Её пока обеспечивает универсальный компьютер, используемый в качестве «искусственного нейрона» нейронной сети. Все другие «искусственные нейроны» [11] пока не обеспечивают полноту логической универсальности в условиях самостоятельной работы. И всё бы хорошо, но ... компьютер требует внешнего программирования и контроля работы.

С полной самостоятельностью у него плохо...

(B) **Конструируемость.** Задача - *сконструировать, т. е. сконструировать и построить из (надлежащим образом определенного) «сырья», один автомат с помощью другого...* - пока не решена. Никак. Если задача проектирования какого-то автомата на основе заданных параметров еще как-то может быть решена, то задача его самостоятельного производства только автоматом пока даже не ставилась ...

Здесь возникает очень сложная задача, о которой фон Нейман задумывался, но решения не было и нет, ни тогда, ни сейчас.

Как и из какого «сырья» можно построить нового робота?

К космической теме самовоспроизводящихся роботов этот вопрос имеет прямое отношение. В идеале необходимо взять на неизвестной планете какой-то «воздух», грунт под ногами или горную породу, произвести с ними необходимые физические и химические операции и получить металлы и материалы, необходимые для производства новых роботов. Фантасты в своих романах примерно так эту картинку и рисовали...

А как такое производство роботов должно выглядеть в реальности?

Вот, пара вариаций:

1. *На планету высаживается команда специализированных роботов.* Они начинают работать в составе общей технологической цепочки. Одни находят нужные породы и газы для производства необходимых материалов и металлов, и начинают производить эти виды «сырья» в нужном количестве. Вторая группа, из этого «сырья», создает все комплектующие для сборки новых роботов всех классов. Третья группа занимается сборкой и отладкой новых роботов. Четвертая группа проводит профессиональное обучение и специализацию новых роботов, и включает их в производственный процесс. И где-то еще работает высший центр управления, контролирующий весь процесс до его завершения и выполнения поставленной задачи. Далее идет переселение новой команды созданных роботов на новое место или новую планету...
2. *На планету высаживается универсальный робот, имеющий какое-то количество специализированных микро и нанороботов.* Роботы узкоспециализированы и не имеют большой производительности, но их начальное количество и разнообразие вполне обеспечивают постоянное их самовоспроизводство под контролем универсального робота. В какой-то момент времени при выполнении программы, универсальный робот должен дать команду на начало сборки нового универсального робота из полученных материалов и металлов. Новые универсальные роботы со своим запасом микро и нанороботов переселяются на другие места или планеты... И цикл повторяется.

Различие в вариантах небольшое, но весьма существенное.

Если есть только большие универсальные или специализированные роботы, то они образуют легко локализуемую колонию, а вот множество нанороботов, да еще не обладающих высокой производительностью выполняемых работ, ... локализовать и уничтожить очень трудно. Такая «серая слизь» может наделать много хлопот одним и дать большие преимущества в освоении новых планет другим. При этом:

Только микро или нанороботам, под контролем большого универсального робота, будет «по размеру» создать новую логическую структуру «мозгов» для нового робота.

Это связано с трудностью создания общей координации размеров и перемещений при этих работах. В этих условиях возможны только местные привязки и измерения только «в попугаях», т.е. в относительных единицах измерения, основанных на отношениях в размерах самих микророботов. Возможно, это будут нанометры и микроны, а возможно и нет.

«Мозги» робота для нас являются наиболее показательной частью самовоспроизводимых машин, хоть возможно, и не самой сложной технологически. Тем не менее, именно эта часть определяет пока технологию самовоспроизведения машин. Если мы сумеем построить новые «мозги» силами колонии микророботов под управлением робота «нашей» размерности в условиях его полной автономности, то мы уже сможем говорить о каком-то решении проблемы самовоспроизведения для больших и сложных роботов.

Из чего лучше создавать новые «мозги» робота?

Или из имеющихся металлов и материалов или ... из самих воспроизводимых в массовом количестве нанороботов. Тогда каждый наноробот будет «нейроном» в новой нейронной сети. Её соединения могут тоже формироваться из нанороботов. Цепочки нанороботов могут быть металлическими или иными проводниками сигнала. Кстати, использование нанороботов в качестве частей системы и проводников информации имеет некоторые преимущества в сборке «мозга» большого робота [6, 8].

(С) **Конструктивная универсальность.** К вопросу - *может ли какой-нибудь надлежащим образом заданный автомат быть конструктивно универсальным, т. е. способным сконструировать в смысле (В) (располагая подходящим, но по существу стандартным оборудованием) любой другой автомат* – наука ещё не приблизилась. Сегодня существует множество систем автоматизированного проектирования, САПР, но они пока далеки от требований полной самостоятельности проведения этих работ.

(D) **Самовоспроизведение.** У Неймана здесь две задачи: самовоспроизведение автоматом самого себя, и производство им автоматов других классов. Обе задачи наукой пока не решены. Прежде всего из-за отсутствия реального спроса на решение такой задачи. Но вместе с этим, даже если такая необходимость появится, то ... нет базы для разработки таких технологий. Сегодняшняя теория строительства роботов [15] подходит для решения этой задачи очень ограниченно. Тут, как выясняется, всё надо начинать «с чистого листа».

Многие технические «изюминки» для сборки робота роботом в технологиях конечно уже есть, но эти «крупницы золота» надо извлекать из «тонн пустой породы», отрабатывать и включать в новую общую технологию, как стандартные элементы. Так, постепенно, что-нибудь соберем и отработаем.

(E) **Эволюция.** Вопросом самостоятельной эволюции роботов наука только начинает заниматься. И пока результаты не впечатляют. Причина та же – нет реальной необходимости в понимании этого процесса. Задача начнет решаться, когда в её решении появится жесткая необходимость. Правда, здесь есть одна технологическая трудность.

Для начала эволюции должно появиться целевое «мышление» роботов. Как стимул для решения поставленной задачи саморазвития. А вот с этим у наших роботов пока очень сложно.

Ну, оно и понятно. Любое упоминание о «мышлении» робота тут же ассоциируются с «восстанием машин» и законами роботехники А.Азимова.

Тем не менее, эта задача уже поставлена. И решать её придется.

О самовоспроизведении автоматов.

Как оказалось, оно не очень зависит от интеллекта автомата.

Автомат может не иметь интеллекта, но иметь возможности самовоспроизведения.

С другой стороны, автомат, имеющий высокий уровень сложности и интеллекта, почти всегда практически лишен возможности точного самовоспроизведения. Это касается не только наших «железных» автоматов, но и биологической жизни.

Если на уровне клеток и вирусов еще соблюдается более или менее точное копирование «по образцу», то уже на уровне простейших многоклеточных организмов этот принцип нарушается, и переходит в «примерное копирование» по информации ДНК.

Почему примерное? Потому, что на уровне клеток геном воспроизводит всё почти точно, но последующее множественное клеточное разделение по типам и видам уже не может предельно точно воспроизводить родительский организм. Программа развития сложных организмов включает очень много составляющих. На результат здесь влияет всё. От особенностей генома, ферментов, гормонов и т.д., до влияния внешней среды.

Рассматривая весь период эволюционного развития жизни на Земле, невольно приходишь к выводу, что всё эволюционное развитие постоянно крутилось только вокруг механизмов самовоспроизведения [21-24]. Это реализация клеткой сформированного ею же вирусного принципа самосохранения и самовоспроизводства для поддержания своего существования. Да и сам эволюционный процесс начался только после того, как были выработаны хоть примерные, но, варианты воспроизведения «себя».

Самовоспроизведение работает за счет создания *внутренней среды*, которую любой земной организм продолжает сохранять, как основу собственного благополучия и возможности продолжения рода.

Самоподдержание и воспроизведение всех биологических функциональных и матричных автоматов [21-24] начинается с наличия активной внутренней среды, создающей возможность для регенерации изнутри всех частей внутренней системы автомата.

Вот где начинаются молекулярные машины, которые создали биологическую жизнь. Точнее, они создали её *внутреннюю среду*. Замкнутый объем, в котором происходит поддержание всех жизненных процессов. Насколько сами молекулярные машины являются

«биологическими», это еще большой вопрос, а вот то, что они имеют наноразмерность, это совершенно точно. И в техническом их понимании, это и есть нанотехнологии.

Чем сложнее самовоспроизводимый автомат, тем сложнее у него должна быть *внутренняя среда*, обеспечивающая ему надежное существование, а так же, частичное или полное воспроизведение. Так, по крайней мере, эта задача решена в биологическом мире.

Человек же пока идет другой дорогой.

До последнего времени автоматы создавались только с помощью человека и лишь с частичным участием роботов или автоматов в процессе их воспроизведения на автоматических линиях конвейеров. И конечно же, без использования какой либо *внутренней среды* для поддержания внутреннего самовосстановления.

Сегодня все автоматические конвейерные сборочные линии требуют постоянного контроля и прямого участия человека. Пока на Земле нет ни одного производства автоматов полного цикла без участия человека. И не планируется.

По нескольким причинам...

- Полностью автоматизированное производство без участия человека сложно для нас психологически. Мы привыкли считать себя самыми умными и способными решить все возникающие вопросы, даже самые сложные и трудные. Роль машины у нас изначально вторична. Независимо от наших рассуждений.
- Земля, имеющая окислительную атмосферу, исходно не самое лучшее место для организации полностью автоматического производства. Многие процессы такого производства должны проводиться в специальных условиях, которых на Земле нет. Вакуум, атмосфера из нейтральных газов, легко испаряющаяся жидкость, масляная ванна и т.д.
- Вероятность катастрофического сочетания событий при организации и работе на Земле полностью автоматизированного самовоспроизводства роботов, может быть и мала, но, все же реальна. Это заставляет с большой осторожностью относиться самому факту существования такого производства на Земле.

По этой причине самовоспроизводство роботов по полному циклу, например, на Луне, имеет важнейшее значение. Хотя мы, для начала этого процесса там, пока имеем только «отверточные технологии», не более. Но с другой стороны, именно Луна и является самым лучшим полигоном для отработки всего комплекса технологий самовоспроизводства роботов. Это условия открытого космоса, наличие радиации, большой перепад «дневных и ночных» температур, присутствие некоторой гравитации....

В этом смысле, для отработки технологий условия там почти идеальные.

Что у нас есть для начала самовоспроизведения машин?

Если быть честным – крайне мало.

Оказалось, что имеющиеся сегодня технологии для создания самовоспроизводящихся роботов не очень-то подходят. По сути, пока нет таких технологий. Их еще только предстоит разрабатывать и отрабатывать. И на это уйдут годы...

Но, некоторые реальные технологии уже появляются.

Это например, 3D принтеры. Такие устройства уже позволяют создать почти серийное производство самых разнообразных деталей, в том числе и для постройки роботов. Из самых разнообразных материалов.

Здесь выделим проект *RepRap*:

Проект RepRap (от англ. *Replicating Rapid Prototyper* — самовоспроизводящийся механизм для быстрого изготовления прототипов) — инициатива, направленная на создание самокопирующегося устройства, которое может быть использовано для быстрого прототипирования и производства.

Устройство RepRap представляет собой 3D-принтер, способный создавать объемные изделия на основе моделей, сгенерированных компьютером. Одной из целей проекта является «самокопирование», определяемое авторами как способность аппарата воспроизводить компоненты необходимые для создания другой версии себя. Аппарат представляет собой разработку с общедоступными наработками.

Конечно, в космосе он работать не может, но как факт наличия такого направления, очень примечателен.

Да, пока из всего арсенала копирования «по образцу» у нас есть только 3D-принтер и его технологии работы. Где-то на подходе такие же принтеры для порошковых технологий и напыления, в основном - газотермического... В общем, и всё.

Может быть, что-то из этого и подойдет для изготовления «мозгов» новых роботов в условиях дальнего космоса, не знаю. Но учитывая современные нанотехнологии, применяемые сейчас – вряд ли. Невозможно делать что-то очень сложное и бесконечно долго вакуумным напылением по сверхтонким трафаретам, а тем более «грубой» наплавкой в условиях открытого космоса или в агрессивной атмосфере неизвестного состава. У нас все технологии производства роботов сделаны для людей. Они уникальны и практически неповторимы в других условиях.

И пока, чем мельче робот, тем больше комплекс оборудования для его изготовления.

Если для роботов больших размеров проблемы самокопирования еще как-то решаются, хотя бы «отверточными технологиями», то в наноразмерности никаких отработанных вариантов нет.

Для наноразмерных роботов на основе *молекулярных машин* нет и самого главного – памяти, способной удержать в сохранности технологию процесса репликации и какие-то результаты проведенного самовоспроизведения. Кроме РНК и ДНК, других носителей на этом уровне не предлагалось.

А уж о движущихся микророботах вообще говорить что либо очень сложно. Здесь проблемы во всем. От движителя и двигателя, до энергетики, обеспечивающей это движение. Кроме природного варианта на АТФ-синтазе и белковых двигателях, как например, в жгутиках пока ничего не предложено. Ну и, ... нет пока для них никаких «мозгов».

И все же, начинать придется с микророботов.

Потому, что, скорее всего, в технологиях самовоспроизведения всех роботов основную роль должны играть микророботы, управляемые каким-то внешним автоматическим координационным центром. Более всего нано и микророботы подходят для создания *внутренней среды* робота «человеческой» размерности и больше. Их наличие внутри большого робота создает активную внутреннюю среду и позволяет решить проблему поддержания постоянной работоспособности и самовосстановления робота изнутри. Размеры микророботов позволяют проводить работы в наноразмерности.

Ну да, это всё действительно очень похоже, на «рождение», в биологическом мире.

И тем не менее, это наиболее эффективный и безопасный путь создания как больших роботов в условиях полной автономности, так и малых микророботов при сохранении внутренней среды и механизмов её передачи и поддержания.

Ничего подобного у нас пока нет и не предвидится.

Сегодня уже понятно, проблему машинного самовоспроизводства ученые оценили и начали решать. Хотя бы на том техническом уровне, который сегодня есть.

Видимо скоро будет проведена общая стандартизация узлов для всех типов роботов, составлена их классификация, организация САПР для перевода роботов на машинное самовоспроизводство.

Возможно, сначала на Земле.

Потом этот процесс можно будет запускать на Луне из местных материалов, в режиме контролируемого производства. И только тогда может начаться подготовка к самостоятельной работе самовоспроизводящихся роботов в условиях «большого» космоса.

Пока к этому светлому будущему сделан только первый шаг.

Развитие теории фон Неймана...

Сегодня «Теория и организация сложных автоматов», в том виде, как когда-то её определил Дж. фон Нейман, получила мощное развитие именно там, где и предполагал автор. В вычислительной технике.

Все поставленные Дж. фон Нейманом проблемы получили какое-то разрешение.

Но главное развитие идет вот здесь:

Термодинамические аспекты понятия равновесия.

1. Термодинамические аспекты, связанные с емкостью памяти.
2. Необходимость количественной теории вместо применяемых сейчас эмпирических методов.
3. Предварительные замечания о надежности и ошибках.

На вопросы 2 и 3 ответила теория информации К.Шеннона, написанная на основе работ В.А. Котельникова, Г.Найквиста и других ученых, уже хорошо известных Дж. фон Нейману на момент написания этой работы.

Главную сложность и тогда и сейчас представляет вопрос 1.

Путь решения проблемы в общем был Нейману уже известен, но реальные решения стали находиться только сейчас. Тогда вся научная мировая элита уже была увлечена первыми работами И.Пригожина по термодинамике диссипативных систем, написанными им в 1948г. Это там снова появился термин «энтропия», получивший новую жизнь во многих науках, включая и теорию информации. Диссипативные системы (открытые системы) вошли составной частью в общую теорию систем. Тогда же, в работах У.Эшби в 1947г. появилось и понятие самоорганизации материи. А вот замыкающий эту пирамиду новый термин «синергетика» появился уже позже, в работах Г.Хакена. Об этом моя работа [9].

Правда, всё понимание проблем шло в общем русле развития науки того времени. Понимание информации было сопряжено с проблемами энергетических взаимодействий. Теория информации К.Шеннона была написана именно в этом ключе. Рассматривались абстрактная система источников информации, каналы связи, как беспроводные, так и проводные, объемы передаваемой по ним информации и возможность ошибки при передаче.

Позже на этой же основе стали оценивать память вычислительных машин. И энергозатраты при получении и выдаче информации.

Вот тогда и появился принцип обратимости [17] Ландауэра¹:

Принцип Ландауэра — принцип, сформулированный в 1961 году Рольфом Ландауэром (IBM)^[1] и гласящий, что в любой вычислительной системе, независимо от её физической реализации, при потере 1 бита информации выделяется теплота...

Этот принцип наконец-то вывел науку на минимально возможную энергоёмкость информационной системы. На понимание того, что любые изменения в машине связаны с потреблением энергии. Как в процессоре и во всей электронной схеме, так и в памяти...

И чем выше частота изменения, тем больше мощность, затрачиваемая на этот процесс.

¹Рольф Вильям Ландауэр (4 февраля, 1927 - 28 апреля 1999) немецко-американский физик, который внес важный вклад в различных областях термодинамики, обработки информации, физики конденсированных сред, проводимость неупорядоченных сред.^[1] В 1961 году он открыл принцип Ландауэра, который говорит, что «в любой логически необратимой операции, которая манипулирует информацией, такие как удаление части памяти, энтропия возрастает и связанное с количеством энергии рассеивается в виде тепла».^[1] Этот принцип имеет отношение к обратимости вычислений, квантовой информации и квантовым вычислениям. Он отражен в формуле Ландауэра, связывающей электрическое сопротивление проводника и его рассеивающие свойства. Ландауэр среди многих других наград^[1] получил медаль Стюарт Ballantine от Института Франклина, премию Оливера Бакли от Американского физического общества и IEEE медаль Эдисона.

Сегодня это ведет к тому, что дальнейшее повышение тактовой частоты процессора в компьютере, даже при самых современных технологиях их производства, уже не приводит к радикальному росту производительности и снижению энергопотребления [18, 19].

Например, при улучшении технологии, уменьшении размеров проводников и активных элементов схемы (транзисторов), для увеличения быстродействия компьютера можно соответственно увеличить тактовую частоту процессора, что и делается сегодня. Но такое качественное улучшение уже не дает радикального снижения энергопотребления. Потребление энергии снижается достаточно медленно и остается относительно большим, несмотря на все технические и технологические улучшения. Потому, что с ростом частоты растет и плотность переносимой энергии электромагнитной волны. Мы же это давно знаем...

И вот, портативные «электронные книги» на процессорах с малой тактовой частотой вполне сносно работают, потребляя доли ватта. Современные планшетные компьютеры имеют уже более высокую тактовую частоту, выделяют тепла около 2,5 ватт, а современные ноутбуки требуют десятков ватт энергии даже при отсутствии какой либо работы. При этом весь компьютер, если исключить внешнюю память, содержит *несколько сот миллионов активных элементов* (транзисторов), обеспечивающих всю работу по обработке информации.

Но, сравните: Человеческий мозг содержит по разным данным *от 20 до 100 млрд. нейронов* и в обычном режиме работы потребляет около 20 ватт энергии!

Как лучше проводить машинные вычислительные операции, последовательно или параллельно? Наука долго, почти полвека, шла по пути последовательных вычислений. Тогда, даже самые современные процессоры, были построены по принципу последовательных вычислений.

Но, где-то к 80-м годам прошлого века уже стало понятно, что параллельные вычисления необходимо применять. Заговорили о многопроцессорности. Да, дорого, надо иметь много процессоров в одном компьютере, но они дают выигрыш не только во времени, но и в общих затратах энергии на проведение вычислений. Это связано и со снижением тактовой частоты работы процессоров, параллельно производящих вычисления.

Здесь можно вспомнить один парадоксальный факт для сравнения:

Если все нейроны человеческого мозга всего лишь 1 раз в секунду произведут действие принятия или генерации электрического импульса, то общая производительность работы мозга составит *100 млрд. операций в секунду!* А с учетом того, что в каждом отростке нейрона, аксоне или дендрите, имеются тысячи синапсов, сложных управляемых соединений с другими нейронами, то общее количество операций, выполняемых мозгом в секунду, по принятой для компьютера методике, уходит далеко *за десятки триллионов!*

Какой компьютер может похвастаться такой, пусть и пиковой производительностью?

И никакое параллельное включение компьютеров, даже в режиме нейросети, не обеспечивает такую скорость обработки информации, какую дает мозг. Вопрос здесь не в количестве нейронов, а в принципах их соединений. И способах обработки информации. Компьютер не позволяет реализовать эти способы просто потому, что он основан на другом способе обработки информации. Математическом. А мозг работает по законам своей логики.

Чтобы приблизить работу «искусственных мозгов» роботов к работе нашего мозга нужно изменить принцип их построения. С математического на логический.

Красиво, но ... непонятно.

Здесь многие сразу возразят, а как же нейронные сети, содержащие уже миллионы компьютеров или процессоров, иначе называемых «искусственными нейронами» [11]?

Это же почти электронный мозг по принципам работы...

Да, похоже, но – нет. Ни один нейрокompьютер пока столько «нейронов», как наш мозг, не содержит. *Даже весь глобальный Интернет, вместе со всеми подключенными к нему компьютерами, не может пока конкурировать по сложности с одним человеческим мозгом.*

Кроме того, нейронные сети, как и нейрокомпьютеры требуют отдельного вида программ для согласования работы всех составляющих этой сети. Конечно, такие программы уже появляются, но сложности согласования работы всех узлов нейросети и обмена информацией от этого не уменьшаются [11].

Параллельность обработки информации в нейронной сети требует не только унификации всех её узлов, но и какого-то нового базового принципа автономной работы узлов в составе сети, обеспечивающем работу всего комплекса без внешней синхронизации или прямого внешнего управления. Каждый «нейрон» этой сети должен работать по принципу уже порядком подзабытой «логической машины» [10]. И при этом каждый нейрон должен иметь свою память, как часть общей памяти системы.

Да, история [43] возвращается ...

Сегодня всё это только ускорило понимание того факта, что классическая архитектура, как процессора, так и всего компьютера, уже вышла на финишную прямую этого направления развития. Для сохранения прогресса в развитии требуются новые принципы построения вычислительных и логических машин.

И появилась теория *программируемой материи*:

Программируемая материя — это материя, которая может изменять свои физические свойства (форму, плотность, структуру, оптические свойства и т. д.) программируемым образом, посредством заданных пользователем или автономных восприятий. Программируемая материя, таким образом, связана с концепцией материала, который имеет внутренне присущую ему способность выполнять обработку информации. Состоит из молекулярных компьютеров.

Термин «программируемая материя» изначально придумали в 1991 году Тоффоли и Марголус для обозначения ансамбля из мелко модульных вычислительных компонентов, расположенных в пространстве.^[1] Их статья описывает вычислительную подложку, состоящую из распределенных в пространстве мелкозернистых вычислительных узлов, которые общаются путём взаимодействия только с ближайшими соседями. В таком понимании программируемая материя относится к моделям вычислений, похожих на клеточные автоматы и автоматы решетчатого газа.^[2] Архитектура САМ-8 является примером аппаратной реализации этой модели.^[3] Эта функция известна также под названием «цифровые поля ссылок» в некоторых разделах науки о самовоспроизводящихся машинах.^[4]

К сожалению, мне не удалось найти материалы по «науке о самовоспроизводящихся машинах». Только вот эту ссылку.

Но, есть проект САМ-8 «На пути к окончательным вычислениям», где существуют понятия «решетчатого газа» и его реализация на математических принципах построения.

Вот так показана его практическая реализация:

Хотелось бы сказать еще несколько слов о машине клеточных автоматов, которая называется САМ-8. Она разработана в рамках проекта ARPA в Лаборатории компьютерных наук Массачусетского технологического института (MIT LCS) и представляет собой компьютерную архитектуру, построенную на клеточных автоматах. И в то же время — это **параллельная масштабируемая архитектура для высокоскоростного моделирования пространственных систем средствами клеточных автоматов**.

САМ-8 состоит из рабочей станции Sun, на которой под управлением операционной системы SunOS функционирует специальная среда программирования. Эта среда реализована в рамках коммерческого интерпретатора языка Forth для SunOS. САМ-8 представляет собой отдельный наращиваемый аппаратный модуль, в который встроены специально разработанные кристаллы STEP и который подключается к рабочей станции Sun через интерфейсную плату с шиной SBus. Для низкоуровневого управления модулями САМ-8 разработана библиотека Си-функций, что позволяет для программного подключения использовать различные интерпретаторы головного компьютера.

Понятно? Это машина представления. Она формирует изображение на экране. Правда, пока реальная конструкция сильно отличается от заявленного «решетчатого газа» и размерами, и реализацией вычислений.

Где-то здесь появилось новое понятие:

Биокомпьютер (также *биологический компьютер, молекулярный компьютер*) — компьютер, который функционирует как живой организм или содержит биологические компоненты. Создание биокомпьютеров

основывается на направлении молекулярных вычислений. В качестве вычислительных элементов используются белки и нуклеиновые кислоты, реагирующие друг с другом.

Можно сказать, что молекулярные компьютеры — это молекулы, запрограммированные на нужные свойства и поведение. Молекулярные компьютеры состоят из сетевых нано-компьютеров. В работе обычной микросхемы используют отдельные молекулы в качестве элементов вычислительного тракта.

В частности, молекулярный компьютер может представлять логические электрические цепи, составленные из отдельных молекул; транзисторы, управляемые одной молекулой, и т. п. В микросхеме памяти информация записывается с помощью положения молекул и атомов в пространстве.

Одним из видов молекулярных компьютеров можно назвать ДНК-компьютер, вычисления в котором соответствуют различным реакциям между фрагментами ДНК. От классических компьютеров ДНК-компьютеры отличаются тем, что химические реакции происходят сразу между множеством молекул независимо друг от друга.

Станислав Лем в «Summa Technologiae» предсказал теоретическую возможность «выращивания информации» при помощи синтетических полимеров (в т.ч. и био-)^[1].

Как мы видим, наука снова обратила внимание на белки, ДНК и РНК. Теперь, как на составные части вычислительной системы. Биологическое направление развивается очень быстро, но пока еще не вышло на устойчивое производство готовой вычислительной техники. Все работы проводятся на уникальных лабораторных образцах.

Еще одно новое понятие, определяющее вектор развития современной вычислительной техники:

Нанокomпьютер — вычислительное устройство на основе электронных (механических, биохимических, квантовых) технологий с размерами порядка нескольких нанометров.

Нанокomпьютер — это запрограммированная наночастица на нужные химические свойства и поведение.

Сам компьютер, разрабатываемый на основе нанотехнологий, также имеет микроскопические размеры. На данный момент создан нанотранзистор — основа нанопроцессора.

Видимо, это предел миниатюризации элементов вычислительной техники. Да, это нанотехнологии, но их результат, тот же нанокomпьютер, уже вряд ли будет в размере «нано». Это уже точно - только «микроскопические размеры».

Так что, похоже, ... нанороботы остаются без «программного блока».

Между тем, понятие «программируемая материя» [16] очень быстро стало широко применяемым. В эту область включили квантовые компьютеры и фотонные компьютеры.

Не все показанные виды вычислительных машин пока укладываются в рамки нашей темы самовоспроизводящихся машин для покорения космоса. Но они вполне укладываются в рамки «теории сложных самовоспроизводящихся автоматов» Дж. фон Неймана.

По этой причине мы о них и вспомнили.

Где начинается движение?

Где возникает начало любого движения у человека, понятно.

Теория уже доказала - в нейронных сетях мозга.

А где появляется движение у гидры, у одиночной клетки, там где и мозгов-то пока нет?

Где и как формируется простейший фактор сдвига, создающий вектор направленного движения?

Понятно, что в живой природе этот самый механизм направленного сдвига вырабатывался медленно. От любого импульса движения к целенаправленному. Одним из первых был найден механизм механического перемещения части белковой цепи ретинали [22]. Но ведь его еще надо было направить в нужном направлении для создания общего целенаправленного движения....

Ну ладно, про клетки и их организацию движения человек знает относительно недавно. Он всегда больше на себя ориентируется, и на тех живых существ, что видит вокруг себя.

Или на то, как он это себе представляет ...

Нам нужен механизм, превращающий любое случайное движение в целенаправленное.

Уже давно принято начинать с наглядных картинок. Проще всего показать наши логические обоснования на пространственных перемещениях вымышленных объектов.

Например, по клеткам того же декартова координатного пространства. В виде каких-то простейших объектов, тех же точек, например.

Точно так и поступил Дж. фон Нейман в [5].

Смотрим рис.2.

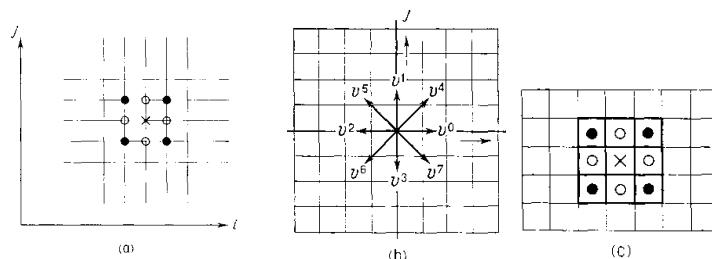


Рис. 2. Структура с квадратными клетками.

а) Ближайшие (O) и следующие за ними ближайшие (A) соседи точки X.

б) Единичные векторы, в) Ближайшие (O) и следующие за ними ближайшие (O) соседи клетки x.

Что же двигал Дж. фон Нейман? Он двигал «возбуждение» математических нейронов.

Предполагается, что вся плоскость перед нами заполнена однотипными «нейронами» в виде клеток координатного пространства. И мы передаем электрический импульс в нужном направлении к нужному нейрону. Как это сделать?

Дж. фон Нейман пошел классическим путем.

Он нарисовал эту площадь, как на рис.2с. Где видно, что из клетки с крестиком передается потенциал возбуждения по линиям координат (белые кружки) и по диагоналям (черные кружки). Потом он изобразил дальнейшее перемещение потенциала возбуждения уже векторами. И начал составлять математическое описание конечных автоматов, позволяющих реализовать этот процесс ... на листе бумаги, а позже - на экране монитора.

Примерно так появилась теория клеточных автоматов.

Состояния взаимодействий клеточных автоматов между собой получили определения: *невозбудимое, конфлюэнтное, передающие (специальные и обычные) и чувствительное*. Каждый вид взаимодействий описывает тот или иной логический аппарат математической логики. Всего в клеточном автомате Дж. фон Неймана получилось 29 устойчивых логических состояний и переходов между ними.

Какое отношение эта теория имеет к реальным движениям и перемещениям объекта в пространстве? Ну, какое-то имеет. Теория позволяет моделировать эти процессы движения на экране монитора математическими методами. И передавать электрические потенциалы в реальной схеме соединений, созданной по этой модели.

С другой стороны, какое реальное движение могут создавать все эти схемы?

Да, пока никакого. Для реального перемещения объекта нужно резко усложнять сам реальный объект. Создавать какие-то реальные средства его перемещения в пространстве, определять схему и метод управления ими. Создавать какие-то средства передачи информации от одного объекта нашей «колонии» к другому. В общем, это сложно. И для теории почти не нужно. Потому, этим никто и не занимался. Ни тогда, ни сейчас.

Почему путем создания клеточных автоматов пошел Дж. фон Нейман, нам почти понятно.

Непонятно другое. Если мы исходно подменяет в модели реальные объекты математическими, заменяем движение этих реальных объектов на перемещение «возбуждения» от одного объекта к другому, и фактически заменяем реальный объект его электрическим «фантомом», потенциалом, то почему мы продолжаем при этом говорить о движении реальных объектов?

Даже сейчас, через полвека после публикации работы [5].

Двигаем возбуждение

Ну ладно, бог с ним, с реальным движением.

Мы принимаем условия игры.

Попробуем, оставшись в модельном пространстве, представить себе, что какие-то «объекты» у нас в нём имеются, а вот движения возбуждения между ними ещё нет. Никакого. И пока нет никаких конечных автоматов, которые могли бы помочь в организации этого движения.

У нас также должно двигаться «возбуждение», пусть и как электрический потенциал, но вот ... проводников для этого пока нет. Формат проводника для передачи возбуждения еще надо определить. Мы пытаемся создать простейшую схему, передающую возбуждение хотя бы в одном заданном направлении. А потом ... в двух.

Этот вопрос рассматривался в работе [8].

Вопрос оказался не таким уж простым, как это представляется вначале.

Но, посмотрите на рис.3.

Похоже это на построения рис.2?

Конечно. Потому, что решалась одна и та же задача. Хотя и ... различными методами.

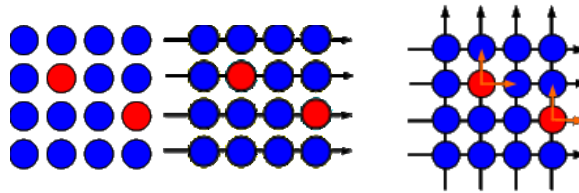


Рис. 3. Введение линий передачи возбуждения в координатном пространстве объектов движения.

Дж. фон Нейман сразу рассматривает все ближние и дальние окрестности заданной точки координатного пространства и решает задачу математическими методами для всего комплекса возможных состояний взаимодействий, которые он посчитал необходимыми в рамках математической логики. В качестве клеточного автомата использована сложная схема с 29-ю устойчивыми состояниями и взаимодействиями, определяемыми математикой конечных автоматов.

Во нашем случае [8] ставилась задача постепенными и простейшими методами сначала найти возможные варианты построения каналов передачи возбуждения с определенными свойствами, а потом найти и такую же простейшую схему передачи возбуждения по возможным направлениям. По одному или двум.

Но, после некоторых вынужденных поправок технического понимания проблемы конечный вариант координатного пространства, в котором идет передача возбуждения от точки к точке, несколько изменился. Вот он, на рис. 4.

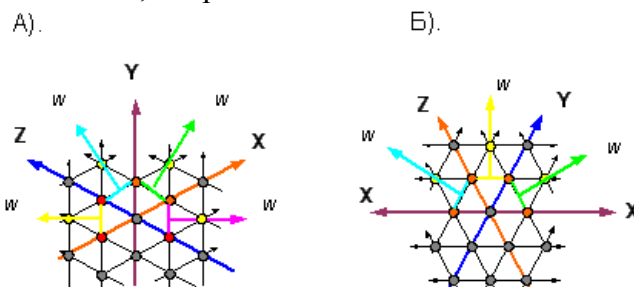


Рис. 4. Конечные варианты координатного пространства с направлениями передачи возбуждения между точками этого пространства.

Это связано с фиксацией факта начала движения возбуждения. Нужен еще один очаг возбуждения. Только два очага вместе создают какое-то направление движения возбуждения и

двигают его дальше. Куда? Только в сторону от линии очагов возбуждения. Иначе придется решать какой из очагов «исходный», а какой – «управляющий». Нам это надо?

Понятно, что полученная ячейка координатного пространства, передающая и получающая возбуждения, должна иметь однозначную пространственную ориентацию. Иметь отдельные входы и выходы.

Конечный вариант схемы автомата получился эмпирически, из трех триггеров, элемента «И» и схемы коммутации их входов и выходов.

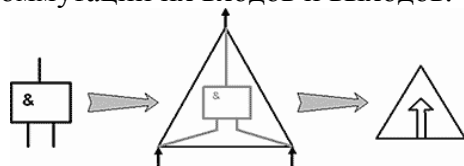


Рис. 5. Модификация логического элемента «И».

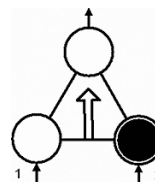


Рис. 6. Базовая ячейка счетной логики

На рис. 5. и 6 показана сборка из элементов «И» и двоичных триггеров. Триггеры на рисунке обозначены кружками. Например, белый, это состояние «0», черный - состояние «1». Это элементы фиксации состояний, они находятся в точках соединения. Выходы и входы всех триггеров имеют и внешние контакты.

Схема алгоритма задачи собирается из таких логических элементов. Да, в какой-то мере это решение отвечает простейшему решению задачи однородных вычислительных сред из работ Э.В.Евреинова.

Получившиеся электронные схемы такого направления мною давно [6, 40] были названы *счетной логикой*. Главное отличие этого типа электронной логики от существующей, двоичной, в последовательном продвижении возбуждения по схеме соединений логических элементов.

Как мы знаем, в схемах соединений элементов двоичной логики возбуждение проходит по всей схеме из многих логических элементов двоичной логики за один такт работы схемы. После окончания рабочего такта логические состояния на элементах схемы не сохраняются.

По этой причине обычная двоичная логика не может работать в схемах последовательного подхода к правильному ответу за время последовательной подачи нескольких импульсов на разные входы логической схемы. Для обычной двоичной логики нужна внешняя синхронизация. Для новой логики она нужна лишь ограниченно.

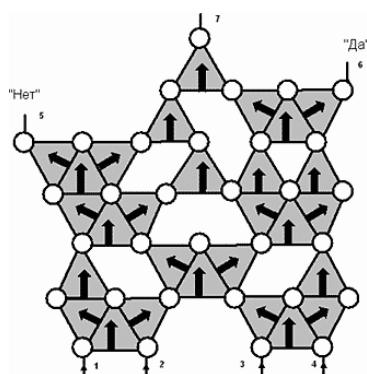


Рис. 7. Логическая структура на элементах счетной логики.

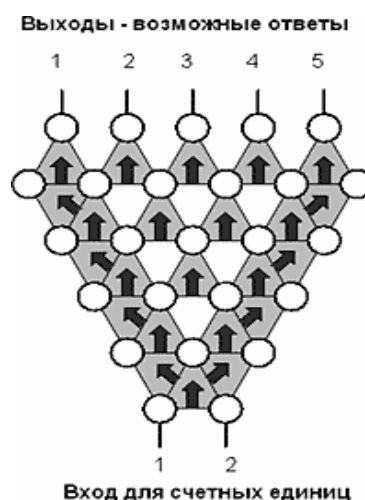


Рис. 8. Пример логической задачи счетной логики.

На рис. 7 и 8 показаны примеры логических структур на основе счетной логики для решения той или иной логической задачи методами управления с информационных входов.

Именно этот путь, как мне кажется, может привести к созданию *логических машин*, основы для «искусственных мозгов» робота. Альтернативы для компьютерных систем.

В основе счетной логики лежит сначала «основной элемент», показанный на рис. 6.

И несколько систем счисления...

Это *система Бергмана* [6] и *коды Фибоначчи* А.П.Стахова [6]. Но, эти системы счисления не отменяют *двоичную систему* и логику, применяемую сейчас. И она необходима.

Оказалось, что очень интересные варианты построения логических схем на этой базе дает и применение *единичной системы счисления* [6]. Примеры показаны на рис.9.

В основном, это каналы передачи возбуждения на большие расстояния. И схемы управления этими каналами передачи по направлениям и условиям.

Оказывается, можно создавать как однонаправленные каналы передачи импульсов, так и каналы с передачей импульсов в двух направлениях, вперед и назад.

Кроме того, основной элемент счетной логики, при его небольшой доработке дает возможность осуществить техническое прогнозирование возможности построения пути к нужному результату до проведения решения [6].

В автономной или счетной логике нет противоположностей.

Есть полярности. Это относительные противоположности, которые действуют вот сейчас, в этот момент в этой задаче. «Да» и «Нет» здесь разделены пространственно, например так, как на рис.6.

Тем не менее, инверсия, как логическое действие вполне допустимо и показано в основах *математики двухпозиционной счетной логики* [40]. Но более понимаемой стала эта же логика, но без двухпозиционности. Опубликована она была как логика автономных систем [44]. В дальнейших работах эта логика ещё меняла название, но её основа оставалась неизменной.

Добавлялись интерпретации её понимания. Они постепенно доводят эту логику до системы. Правда, тут работы еще очень много...

Сегодня основные важнейшие моменты самостоятельной логической системы уже обрели контуры.

Логика автономных систем.

Никакая математика логики, да и любые преобразования невозможны без системы сравнения весов единиц системы. Вообще, любая логика построена на взаимодействии и сравнении эквивалентов.

Тут мне вдруг вспомнились мои подходы к решению почти такой же задачи. С момента публикации работы [7] прошло уже больше 10 лет.

Просто пройдемся по основным точкам понимания этого вопроса...

Какими свойствами должна обладать система автономной логики [44]?

Вот, примерно такими [20]:

- $1 \leftrightarrow 0,1(11111\dots)$
- $1(0) \leftrightarrow 1,1\ 1+0,1$
- $1(0) \leftrightarrow 1+0,1(111\dots) \leftrightarrow 1+1$
- $1+1+\dots+1 \leftrightarrow 0,1111\dots$

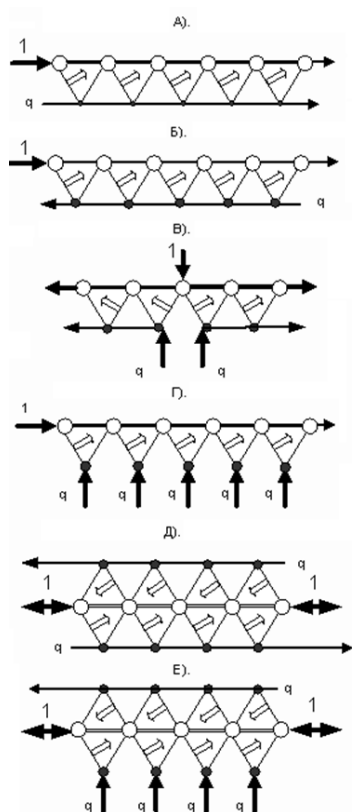


Рис. 9. Варианты продвижения возбуждения в цепях единичной системы счисления.

• ...1111 ↔ 0,1111...

Рассмотрим их чуть подробнее.

1. $1 \leftrightarrow 0,1(1111\dots)$ Единичность события или объекта сохраняется в любом случае. Отдельное событие или объект всегда есть сумма его, в том числе и симметричных, составляющих. Только в этом случае единичность события или объекта полная. Событие или объект не может состоять из одной составляющей. Составляющие обязательно затрагивают все стороны его проявления. Но, в любом событии или объекте может быть только одна доминанта отличия. Даже при полном равенстве вероятности действия всех. Одна основная, а все остальные – дополнения к ней. Единичность должна сохраняться и при сложении составляющих. Полного равноправия быть не может.

2. $1(0) \leftrightarrow 1,0 + 0,1$ Единичность не нарушается, если даже появилось качество, отличающее эту единицу от всех остальных. Может быть, она и другая единица, но, все равно – единица. Это основа несимметричного роста, как качественная, так и количественная. Несимметричный путь развития – наиболее вероятный.

3. $1(0) \leftrightarrow 1+0,1(111\dots) \leftrightarrow 1+1$ Симметрия, только частный случай развития, качественного и количественного роста. Симметричное развитие всегда готово к переходу на несимметричный путь. Симметрия предусматривается, но почти всегда проявляется только в частностях.

4. $1+1+\dots+1 \leftrightarrow 0,1111\dots$ Сумма объектов множества всегда единична. 1(одно) множество. Каждый объект, только часть чего-то общего, их объединяющего. Это начало рационального счета.

5. ...1111 ↔ 0,1111... Множество объектов любой иерархии, это, прежде всего, единичный объект. Сложение и умножение, только различные пути роста этой единичной структуры.

Всем этим критериям в той или иной степени может соответствовать только одновременное применение нескольких систем счисления. О них мы уже говорили.

Теперь немного о системе единиц логики, их сравнениях и взаимодействиях...

Качество – составная часть логического объекта.

Свойство, в его абсолютном понимании. Свойство какого либо логического объекта. Оно само по себе рационального количественного объема не имеет, но как оценка присутствует в системе. Мы определим его, как 0,1 – часть Целого.

Любой логический объект состоит из нескольких качеств.

Один элемент какого-то множества, один из множества точно таких же элементов, это – *формальная единица* системы. Или просто, единица - 1. Абстрактная счетная единица, так же определяет это и математика.

Но, начинается-то всё с малой части 0,1 - качества. Это составная часть логического объекта. Единица, это МНОГО качеств. Из качественных частей и складывается формальная единица в логическом определении.

Вот примерно так:

$$0,1+0,1=1 \quad 1)$$

Каждая формальная единица состоит из набора качеств (свойств), общих для единиц этого множества.

Еще одна единица измерения – *значимость*.

Это особое качество, работающее на уровне логических объектов. Качество, придающее объекту и составляющей его информации статус значимых для системы, а значит, и требующих внимания.

Значимость – объединяющее качество логических объектов. При наличии у логического объекта такого качества он становится *значимым логическим объектом*.

Когда у формальной единицы множества появляется ещё и отличительное качество (свойство) значимости, она становится *значимой единицей*:

$$1+0,1=1(0) \quad 2)$$

Значимость в сложении с формальной единицей дает уже значимую единицу - $1(0)$.

Это новый элемент в весовых единицах системы.

Для логической системы это все равно – единица, но уже *другая*.

Математически $1(0) = 10$, а логически это 1 (один) десяток, эквивалентный множеству, в котором МНОГО единиц. Будем считать $1(0)$ единицей, формально объединяющей множество *десятка* системы Бергмана. В десятке этой системы счетных единиц $1 < 10_{\phi} < 2$ или, что то же самое $1 < 1(0) < 2$.

Тогда можно представить, что 1 – счетная единица системы, а $1(0)$ – большая единица системы, или *значимая*. Что это дает, мы увидим дальше...

Форма сложения $0,1+0,1=1$ и есть соединение возбуждений на входах основного элемента счетной логики на рис. 6. При этом получаем состояния 1 на выходе элемента и переход входов обратно в состояние 0 . Новый сигнал $0,1$ на любой из входов уже создает пару состояний для движения возбуждения как в правую сторону от этого элемента, так и в левую, как на рис.7 и 8. В этом случае и применяется формула $1+0,1=1(0)$;

Теперь вернемся обратно к общей системе логики.

Рассмотрим шкалу единиц этой логики:

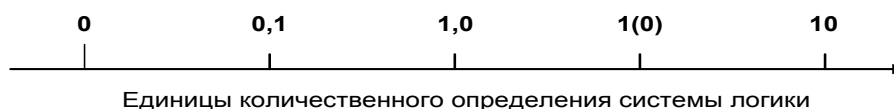


Рис. 10. Количественные оценки логики.

Вес единицы системы возрастает слева направо. В основе системы количественных оценок лежит принцип относительности. Только сравнение может дать оценку различий.

Немного общих размышлений о единицах измерения.

МНОГО - несчетное множество. Всё, что за пределом счета. *Всё, что больше единицы.*

НИЧЕГО или 0 , как мы уже говорили, имеет скорее косвенное, чем прямое понимание. *НИЧЕГО, это почти всегда сравнительная оценка.* НЕТ чего-то вполне определенного. Абсолютной пустоты никакая логическая система не понимает, даже в принципе.

Что-то всегда есть. Всегда. И это подтверждается нашим опытом и исследованиями. Даже вакуум не пуст. И в нем всегда что-то происходит. При ближайшем рассмотрении оказывается, что он кипит, как бульон на плите.... По этой причине и человек долго не включал понятие НИЧЕГО в систему количественного определения. Его просто не рассматривали, как возможное логическое состояние. Да и сейчас это понятие есть только в математическом определении, как 0 . А вот сравнительная оценка, как НЕТ чего-то конкретного, это - существует. Но сравнительную характеристику сложно превратить в количественную оценку.

И всё же, отсутствие предмета сравнения, как несчетную оценку, система должна как-то учитывать. Как невозможность количественной оценки. Вот тут и происходит слияние НИЧЕГО и МНОГО в одну группу оценки на основе *невозможности определения*.

Группа неопределяемых счетных множеств:

0 – отсутствие определяемых логических объектов; 3)

10 - много.

При этом сама система имеет *количественные оценки*:

$0,1$ - качество, часть счетной единицы; 4)

1 – счетная единица;

$1(0)$ – значимая единица.

Такое количество счетных единиц системы позволяет использовать их в количественной и качественной оценке реальных и логических объектов. Это и две пары определителей:

Часть – целое, как **0,1 и 1** или **0,1 и 1(0)**.

Один – много, как **1 и 10** или **1(0) и 10**.

Вот в этой системе оценок и идет сравнительное определение.

Все единицы количественного определения, конечно же, имеют относительное различие.

Для чего? Нужно как-то различать сравниваемые эквиваленты. Определять главное. Составлять иерархию логических объектов. Задач много, а средств для этого – мало.

Как мне кажется, приведенная система весовых единиц достаточна для выполнения основных задач сравнительной логики.

Теперь переходим к соотношениям и сравнительным оценкам. Они позволяют преобразовать единицы системы одни в другие, сравнивать и выбирать:

0,1 – часть логического объекта, свойство, качество.

1↔0,1+0,1+...+0,1 – сложный логический объект. Он обладает суммой качеств своего множества. Качеств – МНОГО.

1(0)=1+0,1 – *значимый* логический объект, обладающий отличительным качеством – значимостью.

1(0)←10 – сложный (значимый) объект, полученный обобщением множества до единичного объекта. По этой же причине логический объект становится *значимым* и получает высокий счетный вес – 1(0) при наличии связей с объектами нескольких множеств в логической системе. Это переход: «МНОГО» деревьев → ОДИН лес; например.

Конечно, получаемая система не совсем математическая, но это же - логика.

Применение счетной логики.

Такая система логики позволяет сформировать исходное множество взаимодействий единиц логики в формальных операциях только на основе *логических переходов*, как вариантов *эквивалентной замены самих единиц действиями*.

В основе понимания *триплет*. Это минимальная сумма оснований ДНК, кодирующая ту или иную аминокислоту в составе белка. Числа 0,1; 1; 1(0); 10; кодируем основаниями ДНК – А, Т, Ц, Г... И получаем *триплеты*: АТЦ, ТГЦ и т.д.

Всего 64 возможных триплета.

А вот так производится необходимая функциональная эквиваленция триплета единицами системы в *логический переход* [20]:

$$\begin{aligned} \text{А) } ААЦ = 0,1;0,1;1(0); & \qquad \qquad \qquad 5) \\ 0,1;0,1;1(0) = 0,1 \rightarrow 1(0) & \\ \text{Б) } 0,1;0,1;1(0); = \rightarrow 0,1 \xrightarrow{1(0)} & \end{aligned}$$

Как мы видим, возможно два варианта преобразования триплетной группы единиц системы в логический переход вида:

$$(a \circ b); \text{ или } [... \circ (a, b) \circ ...]; \qquad \qquad \qquad 6)$$

Где \circ - любое логическое действие.

В варианте А) показан переход вида $(a \circ b)$ для преобразования логического объекта $a = 0,1$ в логический объект $b = 1(0)$. В варианте Б) показан переход вида $[... \circ (a, b) \circ ...]$, когда на любом объекте (a, b) , в данном случае это 0,1, изменяется действие. До объекта (a, b) было простое преобразование, после – условное преобразование.

В сумме же, эти два вида эквивалентных интерпретаций триплета в логические переходы дают возможность построить как бесконечные цепи логических преобразований, так и конечные, с выходом на результат.

Вот примерно так логические переходы образуют логическую цепь:

$$\begin{aligned} (0,1+1,0) &\rightarrow I(0); & 7) \\ I(0) &\rightarrow 10; \\ 10 &\rightarrow (0,1+1,0); \end{aligned}$$

Сдвиги строки вниз показывают - линейка триплетов действий, как логическая связь, полностью логически обоснована. И каждая строка записи – триплет.

Теперь мы можем из отдельных переходов убрать повторы и собрать все переходы одну цепочку ассоциативной связи действий:

$$(0,1+1,0) \rightarrow I(0) \rightarrow 10 \rightarrow (0,1+1,0); \quad 8)$$

Мы замкнули цепь действий. И потому, для нас этот круг действий стал замкнутым кругом преобразований эквивалентов. Он работает в обе стороны. Сумма слагаемых дает или один объединенный объект, или МНОГО объектов. И наоборот.

$$(0,1+1,0) \leftrightarrow I(0) \leftrightarrow 10 \leftrightarrow (0,1+1,0); \quad 9)$$

Сумма слагаемых дает ОДИН объединенный объект или МНОГО объектов. Такой круг эквивалентов в человеческой логике называется - ПОНЯТИЕ и его эквиваленты.

Такая счетная логика позволяет *формализовать снизу любые логические действия высшего уровня*. Т.е. сделать то, что двоичная логика исходно не могла. И не может. Перейти от простого уровня автоматной логики на уровень языков высокого уровня только на основе своих изначальных эквивалентов. Тех самых весовых единиц логической системы.

А автономная или счетная логика это может. И теперь можно говорить о самовоспроизведении электрической схемы «мозгов» с автоматическим воспроизведением и вложенных в этот мозг основных алгоритмов действий.

Движение возбуждения по узлам электронной схемы идет только в зависимости от изменения входных сигналов. И в соответствии с алгоритмом схемы. Но ... куда придет это возбуждения мы не знаем, пока не закончится изменение входных сигналов на входах схемы.

Это не вычисления, это лишь логические операции с входными сигналами и сложные взаимодействия схемы и движущихся по ней возбуждений. На такой основе уже можно строить электронные «мозги» автономных роботов. Вместо компьютера и программ... При наличии таких «мозгов» для робота его исходный, собственный интеллект формируется при рождении, а привлеченный дается обучением. Но пока не очень понятно, как это сделать ...

Наш мозг работает с импульсами возбуждениями примерно так же. Правда, там сложность системы другая. Намного сложнее и несравнима пока ни с какой логической машиной.

Как мы помним, все начиналось с организации движения возбуждения от одного «нейрона» системы к другому.

Здесь показано два различных пути решения сложной задачи организации движения возбуждения. Математический, с помощью создания системы конечных автоматов. И создание новой логики для осуществления движения возбуждения в координатном пространстве.

В первом случае мы получаем исходно сложную схему взаимодействий на основе двоичной логики и теории конечных автоматов. Во втором - новую логику и новые правила организации получаемого логического пространства для проведения автоматических операций.

Но, задача организации движения возбуждения в координатах какого-то логического пространства, это в любом случае – основа создания «мозгов» самостоятельной машины.

Так её понимал и Дж. фон Нейман.

Со своей стороны замечу, что клеточный автомат фон Неймана должен иметь внешнее управление и собственную память. С автоматическим режимом работы у него плохо.

Автономная логика этого недостатка не имеет, ... каждая ячейка имеет триггерную память и позволяет сделать продвижение возбуждения вправо или влево от себя, в зависимости от появления любого управляющего состояния на входах. И дожидаться следующих изменений. А потом точно так же реагировать и на них.

Мне кажется, это и есть автоматический режим работы...

Об искусственном интеллекте ... и разуме.

Вот ведь как странно получается....

Для начала работ по теме освоения космоса роботами человек должен сначала изменить свой взгляд эту на проблему. Критично оценить свой процесс производства роботов, сделать его полностью автоматическим и универсальным, доработать его до условий автономного. И уступить свое место машине.

А мы готовы на это?

Ну конечно, хотелось бы спихнуть монотонную и изматывающую работу на машину, для которой эти понятия неведомы. Однообразная, скучная и простая работа, не требующая присутствия ума, мало кому нравится. Пусть машина этим занимается до полного истирания своих шестеренок...

С другой стороны, мы же привыкли, что при выполнении машиной любой работы, за ней пригляд нужен. Машина, она и забарахлит может, сломаться. Это же только машина. И потому на любом, даже самом механизированном и автоматизированном производстве есть операторы и ремонтники. Люди, отвечающие за обеспечение бесперебойной работы контролируемых ими машин, автоматов, роботов..., даже самых «умных».

Чтобы уйти от этой опеки, необходимо увеличивать самостоятельность роботов многократно. Предполагается, что это обеспечат нам успехи в развитии Искусственного Интеллекта, которые появились в последнее время.

Но, что, собственно, мы сегодня понимаем под понятием «искусственный интеллект»?

Например, есть вот такое [определение](#):

Участники Российской ассоциации искусственного интеллекта дают следующие определения искусственного интеллекта:

1. Научное направление, в рамках которого ставятся и решаются задачи аппаратного или программного моделирования тех видов человеческой деятельности, которые традиционно считаются интеллектуальными^[21].
2. Свойство интеллектуальных систем выполнять функции (творческие), которые традиционно считаются прерогативой человека. При этом интеллектуальная система — это техническая или программная система, способная решать задачи, традиционно считающиеся творческими, принадлежащие конкретной предметной области, знания о которой хранятся в памяти такой системы. Структура интеллектуальной системы включает три основных блока — базу знаний, решатель и интеллектуальный интерфейс, позволяющий вести общение с ЭВМ без специальных программ для ввода данных^[21].
3. Наука под названием «Искусственный интеллект» входит в комплекс компьютерных наук, а создаваемые на её основе технологии к информационным технологиям. Задачей этой науки является воссоздание с помощью вычислительных систем и иных искусственных устройств разумных рассуждений и действий^[51].

Здесь мы узнаем интересные данные из истории ИИ [29]:

Параллельно возникло другое направление, которое использовало подход, основанный на том, что высшее проявление интеллекта человека — это логические суждения. Логика может быть представлена посредством манипуляций символами, а именно для таких манипуляций и были созданы компьютеры. Поэтому с использованием компьютеров можно реализовать *символьный искусственный интеллект*.

Это подход к решению проблемы интеллекта сверху вниз, постулируя возможность воспроизведения когнитивных способностей человека без обращения к уровню отдельных нейронов. Он получил название «искусственный интеллект» (ИИ), или «символьный искусственный интеллект», или «классический

искусственный интеллект». И в истории развития этого направления было много взлетов и падений, связанных с очень высоким ажиотажем. Ученые заявляли в конце 60-х — начале 70-х годов: «Через 10 лет у нас будет робот, способный делать любую работу, которую может делать человек».

Конечно, как всегда, человек сразу стал создавать интеллект, как у себя, естественно. На основе логики предикатов. Далее, как мы поняли, этот искусственный интеллект и стал классическим.

Но время идет, а видимых успехов на этом пути пока не так уж много. «Робот, способный делать любую работу, которую может делать человек», так и не появился.

Разум и интеллект. Что о чем?

Похоже, в науке и сегодня жив лозунг: «хоть горшком назови, только в печь не сажай». Полная демократия в понятиях и классификациях...

Вот пример: теория разума - модель психики человека !

Теория разума (англ. Theory of Mind (ToM) — система репрезентаций психических феноменов (метарепрезентаций), интенсивно развивающаяся в детском возрасте. Обладать моделью психического состояния — означает быть способным воспринимать как свои собственные переживания (убеждение, намерение, знание и пр.), так и переживания других людей, — что позволяет объяснять и прогнозировать их поведение. Важнейший аспект модели психического состояния — это восприятие другого субъекта как интенционального агента, — другими словами: осознание того факта, что собственное психическое состояние не тождественно психическому состоянию другого человека.

Сегодня «модели психического состояния» посвящён широкий спектр современных (конца 70-х гг. XX в. — начала XXI в.) исследований в области когнитивных наук и других междисциплинарных исследований мозга, мышления и сознания.

... На данный момент наиболее изученной стороной модели психического состояния является онтогенез. Исследования восприятия у детей показали, что важнейшие аспекты моделирования психического состояния формируются к возрасту 3—5 лет. Здесь можно провести параллель с исследованиями Пиаже, который выявил окончание стадии эгоцентризма в 3—4 года.

Что-то, общепризнанных и понимаемых основ нормальной теории тут, прямо скажем, маловато. Хорошо, смотрим в другом источнике:

Theory of mind (ToM) — англоязычный термин, применяющийся в психологии и когнитивных науках. Он не имеет краткого перевода на русский, а по смыслу означает наличие у субъекта осознания другой особи (своего вида, по меньшей мере, или иного существа) как агента, обладающего другим набором убеждений, знаний, мотиваций и желаний, отличных от своих собственных. Иными словами, наличие ToM предполагает построение в мозге модели сознания другого.

А вы что предполагали? Вот так. Такие теперь теории ...

Мы всё только по себе меряем, так привыкли ...

Почему-то человек всегда любую задачу моделирования чего-то искусственного сразу примерял на себя. Если мы представляем себе глиняного колосса, то обязательно в виде человека. Если мы говорим об интеллекте, то обязательно, как у человека. Взятся человек моделировать разум, и снова - как у человека.

Кстати, а чем *разум* отличается от *интеллекта*?

Хотя бы, у человека...

Интеллект (от лат. *intellectus* — ощущение, восприятие, разумение, понимание, понятие рассудок^[1]) или ум^{[2][3]} — качество психики, состоящее из способности приспосабливаться к новым ситуациям, способности к обучению и запоминанию на основе опыта, пониманию и применению абстрактных концепций и использованию своих знаний для управления окружающей средой^[4]. **Общая способность к познанию** и решению трудностей, которая объединяет все познавательные способности человека: ощущение, восприятие, память, представление, мышление, воображение^[5].

Интеллект — это, прежде всего, основа целеполагания, планирования ресурсов и построение стратегии достижения цели. Есть основания полагать, что зачатками интеллекта обладают животные, и уже на этом уровне их интеллект посредством механизмов целеполагания и достижения целей влиял и влияет на эволюцию

животных^[6]. Изучением интеллекта животных занимается сравнительно молодая область науки, когнитивная этология.

А вот так например, определяется интеллект «вообще»:

Одно из частных определений интеллекта, общее для человека и «машины», можно сформулировать так: «Интеллект — способность системы создавать в ходе самообучения программы (в первую очередь эвристические) для решения задач определённого класса сложности и решать эти задачи»^[6].

Теперь определяем «разум»:

РАЗУМ - философская категория, выражающая **высший тип мыслительной деятельности, противопоставляемый рассудку**. Различение разума и рассудка как двух “способностей души” намечается уже в античной философии: если **рассудок как низшая форма мышления** познает относительное, земное и **конечное**, то разум направляет на постижение абсолютного, божественного и бесконечного. Выделение **разума как более высокой по сравнению с рассудком ступени познания** четко осуществлялось в философии Возрождения у Николая Кузанского и Дж. Бруно, будучи связываемо ими со способностью разума постигательности противоположностей, которые разводит рассудок.

Вот еще одна формулировка:

Разум (лат. ratio), ум^[1] (греч. nous) — философская категория, выражающая высший тип мыслительной деятельности^[2], способность мыслить всеобщее, способность анализа, абстрагирования и обобщения^[3].

Эти формулировки, как мне кажется, наиболее подходят в данном случае. Но, только пока непонятно, что – о чем? Так можно и совсем запутаться...

Я бы последнюю формулировку и не приводил, но только так становится очевидна коллизия смыслов «интеллекта» и «разума».

«Разум» - ratio на латинском... Мы должны понимать в его современном понимании, как «рациональный». А «интеллект» (от лат. intellectus — ощущение, восприятие, ... рассудок) надо понимать, как ... *эмоциональное воображение*, рассудочное мышление.

Но, рациональное, (разумное?) логическое решение, как жестко мотивированное на достижение цели, характерно для современного *машинного интеллекта*, а вот рассудочное (интеллектуальное?), эмоциональное понимание задачи достижения цели с учетом всех факторов, в том числе этических или моральных, явно не является рациональным, как мы это сейчас понимаем, т.е. не является разумным(!).

Или вот ещё...

Мы, только что, примерно такое прочитали: *Жестко логичный Разум в философских оценках стоит выше эмоционального рассудка*.

Помните, в цитате всё это выглядело примерно так: Разум противоположен рассудку. Разум - более высокая по сравнению с рассудком ступень познания. Разум, это способность к анализу, абстрагированию... И тут же: Интеллект, это воображение, ощущение, восприятие. Интеллект, это рассудок. Рассудок – низшая форма мышления.

Ну, как-то так...

А у нас в компьютерной технике и теории ИИ сегодня всё наоборот - *интеллект, это способность к логическому мышлению, а разум, это ... способность мыслить «всеобщее», в том числе и на уровне эмоций и чувств...*

Вот так история распорядилась. И сейчас уже поздно что либо менять...

Сегодня в компьютерных науках эмоциональные и этические оценки в решении логических задач отданы разуму, а рациональность их решения – интеллекту.

Для интеллекта вокруг нас сравнительных аналогов множество. Но и тут сравнение изначально идет только с человеком. Остальные интеллекты учитываются, но реально с человеческим не сравниваются.

У разума, кроме человеческого, других аналогов для сравнения нет. Нет и никаких классификаций или разновидностей. Есть только человеческий разум. Всё.

А как же разумность действий у других представителей Живого? Куда деть историю с Хатико? Его поведение невозможно объяснить только инстинктами, и рациональным интеллектом. Здесь явно виден Разум.

И таких примеров тысячи...

Но пока наука полностью отрицает проявление разума у кого либо на Земле, кроме человека. И нет никакой теории Искусственного Разума, сходной по построению и задачам с теорией Искусственного Интеллекта.

А между тем, необходимость в такой теории уже ощущается вполне реально.

Классификации и теории интеллектов.

Но, возвратимся к искусственному интеллекту и кое-что уточним.

Нильс Нильсон [27] ещё в 1985 году изложил основы *классического символического искусственного интеллекта* и разделил искусственный интеллект на *сильный* и *слабый*.

К «слабому интеллекту» он отнес все вариации машинного интеллекта, а к сильному - искусственный разум.

В 2016г. Ник Бостром снова обратил внимание на проблему «сверхразума» [35]:

В последние годы наблюдается возрождение интереса к искусственному интеллекту, который вполне может обернуться новыми попытками создать *универсальный ИИ* (по Нильсону – сильный ИИ). Эти проекты будут поддерживаться, с одной стороны, производством новейших аппаратных средств, с другой – научным прогрессом в информатике и программировании в целом, во многих специализированных предметных сферах в частности, а также в смежных областях, например нейроинформатике.

... Я считаю гораздо более вероятным, что *сверхразум* появится сравнительно быстро после создания ИИЧУ (искусственный интеллект человеческого уровня). Кроме того, мой взгляд на последствия этого события также принципиально другой: вероятность чрезвычайно сильного воздействия – позитивного или негативного – на человечество гораздо более высока, чем вероятность нейтрального влияния.

Таким образом современные ученые, специалисты по ИИ, снова заговорили о появлении в недалеком будущем «сверхразума», под которым, как мы сейчас догадываемся, пока понимается *суперинтеллект*, превосходящий по мыслительным способностям человеческий. Глобализация задачи снова ставит сверхцель ...

Да, человек не исправим в своих мечтаниях...

Согласно книге [35] Н.Бострома все исследуемые интеллекты, от естественного до искусственного, можно классифицировать как:

- интеллект - реализуемый на физическом объекте алгоритм;
- интеллект человеческого уровня (ИЧУ) - интеллект, способный решать задачи, доступные человечеству (обладает умом, разумом, интуицией, пониманием, способен к познанию, мышлению, воображению);
- искусственный интеллект (ИИ) - созданный не в естественной биологической среде интеллект;
- искусственный суперинтеллект - интеллект, превосходящий в разы возможности ИИЧУ.

Правда, тут же сказано:

В отличие от терминов Н.Бострома общепринятое понимание интеллекта противоречит терминам интеллект животных, игровой искусственный интеллект, окружающий интеллект, совпадает с ИЧУ, а общепринятый "искусственный интеллект" совпадает с ИИЧУ.

Почему Ник Бостром [35] сам дал определения разным вида интеллекта? Потому, что до него это никто не догадался сделать? Нет, потому, что существующая классификация его не устраивала. А вас устраивает приведенная классификация, и Н.Бострома и та, что есть?

Потому, вот еще одна классификация интеллекта, которую мы сегодня можем построить в укрупненном варианте благодаря Н.Бострому [35] и Нильсу Нильсону [27]:

Биологический интеллект.

1. Интеллект животных со сравнительным коэффициентом энцефализации.
2. Интеллект человеческого уровня (ИЧУ)

Искусственный интеллект.

- *Слабый*
 1. Интеллект (машинный) - реализуемый на физическом объекте алгоритм.
 2. Игровой искусственный интеллект.
- *Сильный*
 3. Искусственный интеллект человеческого уровня (ИИЧУ).
 4. Искусственный суперинтеллект - интеллект, превосходящий в разы возможности ИИЧУ.

Получилась упрощенная *классификация основных направлений интеллекта*, как мне она представляется...

Но есть много и других классификаций, а также теорий интеллекта.

Далее я даю материал по [36].

Начнем ...

Интеллект еще в начале 20 века перестали напрямую связывать с человеком и его разумностью, и теорий интеллекта стало ... много:

Двухфакторная теория интеллекта — теория интеллекта, созданная Чарльзом Спирменом, основным характерным признаком которой является представление о том, что на всякую интеллектуальную деятельность влияют два фактора: общий («G») и специфический («S»).

Теория множественного интеллекта — теория, предложенная Говардом Гарднером в 1983 году как модель интеллекта, которая рассматривает интеллект в различных конкретных (в первую очередь сенсорных) условиях, а не как доминирование одной общей способности к чему-либо.

Теория интеллекта и когнитивного развития Андерсона (ВВП) Теория Андерсона гласит, что индивидуальные различия, касающиеся интеллекта и связанных с индивидуальным развитием изменений в уровне интеллектуальной компетентности, объясняются рядом различных механизмов. Различия в интеллекте являются следствием различий «базовых механизмов переработки информации», предполагающих участие мышления и, в свою очередь, приводящих к овладению знаниями.

Теория интеллекта Стернберга (ВВП) В отличие от теории Андерсона триархическая теория Стернберга рассматривает индивидуальный опыт и контекст, а также базовые механизмы переработки информации. Теория Стернберга включает три части, или субтеории: компонентную субтеорию, рассматривающую мыслительные процессы; опытную (экспериментальную) субтеорию, рассматривающую влияние индивидуального опыта на интеллект; контекстуальную субтеорию, рассматривающую средовые и культурные влияния (Sternberg, 1988). Наиболее разработанной из них является компонентная субтеория.

Биоэкологическая теория Цеси. Цеси полагает, что существуют «множественные когнитивные потенциалы», в отличие от единой базовой интеллектуальной способности или фактора общего интеллекта д. Эти множественные способности или области интеллекта являются биологически обусловленными и накладывают ограничения на психические (умственные) процессы. Более того, они тесно связаны с проблемами и возможностями, заложенными в индивидуальном окружении или контексте.

Эмоциональный интеллект (EQ) — это способность осознавать эмоции, достигать и генерировать их так, чтобы содействовать мышлению, пониманию эмоций и того, что они означают и, соответственно управлять ими таким образом, чтобы способствовать своему эмоциональному и интеллектуальному росту.^[1]

Понятие эмоционального интеллекта ввели Питер Стайер и Джек Майер в 1990 году. В отличие от привычного всем понимания интеллекта (IQ), является способностью правильно истолковывать обстановку и оказывать на неё влияние, интуитивно улавливать то, чего хотят и в чём нуждаются другие люди, знать их сильные и слабые стороны, не поддаваться стрессу и быть обаятельным. В практическом понимании EQ — это «здоровый смысл»^[2]. Первые публикации по проблеме ЭИ принадлежат Дж. Мейеру и П. Сэловею. Весьма популярная на Западе книга Д. Гоулмана вышла только в 1995 году.

Справедливости ради, надо отметить, что большая часть приведенных здесь теорий интеллектов относится к разумному мышлению, а не к интеллекту. Но ..., что есть, то есть.

Эволюционная теория искусственного интеллекта [37] начинается со слов: «Перед вами диалог философа и кибернетика». В процессе чтения понимаешь, что так оно и есть. Ничего не придумано. Философ и кибернетик разговаривают ... о психологии. Интеллект они понимают в разрезе его психологической оценки.

И вот, читаем:

УРОВНИ ЭВОЛЮЦИИ ЕСТЕСТВЕННОГО И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Уровень	Название	Отличительные черты	Строительная единица
0	Видовой (прединтеллект)	Неизменен для всех особей одного вида	Инстинкт/программа
1	Условно-рефлекторный	Возможность индивидуальной модификации поведения	Условный рефлекс
2	Моделирующий	Возможность предсказания поведения оппонента. Способность к игре. Подражательное обучение. Появление эмоций.	Модель объектов реального мира, как особый конгломерат условных рефлексов
3	Вербализирующий	Возможность вербальной коммуникации. Способность к вербальному обучению. Письменность.	Слово как указатель на модель или ее элементы
4	Концептуализирующий	Талант к усвоению концепций и применению их в жизни. Восприятие и создание художественных произведений. Способность к выдвижению стратегий	Концепция как особый конгломерат слов (понятий) и художественное произведение как система образов
5	Онтологизирующий (постинтеллект)	Переосмысление и творение бытия в гениальных произведениях культуры. Перестройка видовых, инстинктивных программ	Новые парадигмы, мифы, философские системы

Странная смесь философского и психологического подхода с примесью кибернетических частностей. И всё бы ничего, но разговор-то идет об ... искусственном интеллекте, как мы помним. Хотя, в результирующей таблице этот акцент как-то уже не выделяется.

Для кибернетика, это странно, что и говорить.

Так получилось, что основное развитие все теории и классификации интеллекта получили только с середины 20-го века. Скорее всего, это произошло под давлением фактов. Всего пару лет назад я уже говорил об отсутствии сравнения искусственного интеллекта с биологическим. Наука еще совсем недавно просто отрицала наличие интеллекта, например, у пчел [36].

Подумать только, проблеме ИИ уже скоро век, а разбираться в интеллектах всерьез мы стали только сейчас. У меня создается стойкое впечатление, что наука только сейчас ринулась находить то, что она просто обязана была найти еще полвека назад, в период интенсивной работы над философией ИИ.

Но, то ли никто не понял важности этой задачи, то ли время тогда еще не пришло, эта задача начала решаться только сейчас.

Искусственный разум.

Разум является ограничителем и фактором сдвига к иррациональности для рациональных вариантов решений интеллекта в общей системе управления «субъекта Я». Независимо от природы и уровня этого субъекта. Примерно так действует любой естественный разум биологического происхождения. И человек тут не исключение.

Это первое, и как мне кажется, наиболее важное качество разума в системе управления автономного субъекта Я.

Под ограничением действий мы понимаем систему внутренних ценностей, охраняемых разумом. Какие варианты результатов решения задачи приемлемы сейчас, вот в этот момент определения? Это система определения относительной истины между «хорошо и плохо» во всех возможных и рациональных решениях задачи по отношению к субъекту Я.

Жесткость работы естественного разума здесь поражает. Даже инстинкты тут иногда не властны. Если выбирать между голодом и опасностью, что важнее, то что выберет система управления на основе разума? Она выберет опасность как главный фактор риска для самосохранения системы. И голод отступает...

Иногда бывает наоборот. Инстинкт продолжения рода (себя) заставляет субъекта Я принимать решение против принципа собственного самосохранения. И лосось идет на нерест наперекор всему. А отложив икру, ... умирает. Даже этим обеспечивая потомство кормом на первое время. Конечно, это уровень не человеческого разума, тут уровень более низкий. Но то, что он здесь присутствует – несомненно.

Надо полагать, что:

В любом автономном субъекте Я, обладающем мозгом и сознанием, есть две основных равноценных составляющих мышления – интеллект и разум. Рациональное и разумное.

Видимо, естественных уровней биологического разума много. От насекомых до человека.

Уровни разумности различаются развитием многообразия внутренних ценностей и истин в системе субъекта Я. От простейших ограничителей действий в осуществлении рациональных решений только для себя, до сложных эмоциональных реакций в системе общественных взаимоотношений. Начиная с насекомых, как например в муравейнике или в пчелином улье.

Противостояния в системах «хищник - жертва» видимо надо рассматривать как взаимоотношения разумов сторон. И все взаимоотношения субъекта и общества, как это сделано для человека.

Эмоции – основная составная часть механизма действия разума на рациональность управления. Эмоции, это логическая оценка разума, корректирующая работу интеллекта.

То самое «хорошо» или «плохо», в виде изменения эмоционального состояния.

Здесь работает такая же логика, как и в системе интеллекта. Только оценивает она другие факторы и выдает другие виды логических ответов.

Это уже вполне достаточные основания для более внимательного отношения науки к проблемам равноправия *искусственного интеллекта и искусственного разума.*

Очень желательно быстрее систематизировать их классификацию.

Пока же, вполне допустимой можно признать классификацию ИИ, данную чуть ранее и распространить её на *естественный и искусственный разум.*

Примерно так:

Биологический или естественный разум.

3. Разум животных.
4. Разум человеческого уровня (РЧУ)

Искусственный разум.

- *Слабый*
 5. Разум машинный - реализуемый на физическом объекте алгоритм.
 6. Игровой искусственный разум.
- *Сильный*
 7. Искусственный разум человеческого уровня (ИРЧУ).
 8. Искусственный сверхразум - разум, превосходящий в разы возможности ИРЧУ.

Сразу появляется столько вопросов...

Например, о соотношении *машинного разума* и *машинного интеллекта* в системе управления сложного автономного робота. Как соотносить в логической системе *рациональные* и *разумные* оценки вариантов решения задачи управления? Где баланс разумности и рациональности в получаемом результате?

Появление машинного интеллекта.

В «реальном секторе» понимания ИИ сдвиги наметились давно. Еще в 70-80-х годах прошлого века «Искусственный Интеллект» разделили на «машинный интеллект» и собственно «искусственный интеллект». Чуть выше мы уже это отметили в классификации интеллектов.

А тогда получилось примерно так [26]:

К машинному интеллекту мы будем относить следующие свойства машины:

а) богатство хранящегося в ней запаса «знаний» и «умений» (констант, понятий, всевозможных стандартных алгоритмов), удобство использования этого запаса, а так же его пополнения путем обучения в процессе эксплуатации (что является особо интересным);

б) логическое совершенство методов переработки информации и машине, приема и выдачи информации, а так же способность машины к организации сложного вычислительного процесса в целом;

в) понимания машиной входных языков высокого уровня.

Степень развития этих свойств и определяет уровень «машинного интеллекта».

Ясно, что данные свойства в совокупности имитируют такие черты человеческого интеллекта, как эрудиция, сообразительность, продуктивность и организованность, восприимчивость к входной информации (в частности, понятливость).

...Чрезвычайно важно здесь то, что повышение машинного интеллекта обеспечивает повышение эффективности процесса математической эксплуатации машин в целом, т.е. как подготовки, так и решения любых задач.

Но, постепенно, машинный интеллект, начавшись с практических алгоритмов, снова пошел в сторону глобализации. Правда, теперь на это были весомые причины.

Появились реальные нейронные сети, а потом глобальный Интернет. Вся мощь *классического символического искусственного интеллекта*, а потом и *машинного*, были брошены на обработку данных в сетях и компьютерных системах.

Сегодня высокоуровневый машинный интеллект практически слился с технологиями нейронных сетей. В его описании применили и терминологию из теории нейронных сетей ...

МАШИННЫЙ ИНТЕЛЛЕКТ — способность средств вычислительной техники моделировать и реализовывать различные виды естественной интеллектуальной деятельности человека или его поведение. Результаты достигаются на базе совокупности таких характеристик компьютера, как объём информации («знания») в его памяти, **способность к её пополнению путём самообучения и самоорганизации**, степень реализации *языков* (см.) программирования высокого уровня и степень понимания входного языка при общении человека с машиной, показатель быстродействия при экспертных оценках, способность к решению задач управления, а также распознаванию образов и принятию решений, относимых к сфере искусственного интеллекта (управление полётом и работой порта, доказательство теорем, игра в шахматы и др.).

Вот где-то здесь и появляется «глубокое обучение» [28]:

Глубокое обучение – это набор алгоритмов машинного обучения, основанных на изучении множества уровней представления. Множество уровней представления означают множество уровней абстракции. В этом посте рассматривается проблема обобщения, то есть возможности определять сложные объекты на основании успешно изученных уровней представления.

... Глубокое обучение – это часть более широкого семейства методов машинного обучения – обучения представлениям, где векторы признаков располагаются сразу на множестве уровней. Эти признаки определяются автоматически и связывают друг с другом, формируя выходные данные. На каждом уровне представлены абстрактные признаки, основанные на признаках предыдущего уровня.

Таким образом, чем глубже мы продвигаемся, тем выше уровень абстракции. В нейронных сетях множество слоев представляет собой множество уровней с векторами признаков, которые генерируют выходные данные.

Но, кажется, мы начинаем осознавать, как новое понимание *машинного интеллекта* вышло за пределы компьютера как помощника в обработке информации или решения

математических задач [31]. Сегодня машинный интеллект уже получает различные специализации. Например, фирма Deloitte² включила программы на его основе в состав своей аналитической базы проведения аудита предприятий и компаний [32], как перспективные технологии.

Возникли такие новые понятия, как «интернет вещей» [34], «темные данные» и «темная аналитика» [33], меняющие наше отношение к понятию «информация».

Похоже, машинный интеллект постепенно занимает место искусственного интеллекта.

О техническом «субъекте Я».

Даже при первичном теоретическом приближении к пониманию проблемы создания самовоспроизводимых роботов один вопрос точно нашел свой ответ. Никакого уничтожения «ростков разума», когда либо возникающих у роботов [2], не будет. И это совершенно точно.

Потому, что никаких «ростков разума» у роботов, с «мозгами» на основе компьютера быть не может. И даже нейрокомпьютеры, имеющие тысячи процессоров в своем составе не могут претендовать на спонтанное возникновение разума, близкого к человеческому.

Не та сложность «мозгов» и не так они устроены.

Но, имеющийся у робота «мозг» вполне может иметь определенную самостоятельность своего *технического «субъекта Я»* [20], необходимого атрибута для целевого решения поставленных задач. Все технические характеристики этого пока эфемерного «технического субъекта Я» еще надо разрабатывать, оценивать... и принимать, как необходимую составляющую в составе «мозгов» робота.

Нет, это не программа, это сумма алгоритмов, может быть, в их вещественном исполнении, как схем логики, формирующих смысловую систему процессов самосохранения, самокопирования, самоуправления и т.д.

Здесь начинается самостоятельность машинного интеллекта. Где-то здесь начала эмоций того самого «машинного разума», позволяющего сдвинуть рациональные оценки логики в сторону иррациональности. Без этого невозможно освоение новых пространств. Принцип самосохранения будет препятствовать.

Отсутствие «технического субъекта Я» в системе функций для «мозгов» робота не дает ему возможности достижения поставленной цели «любыми средствами».

Только робот, осознающий свое локальное существование как *субъект Я* [20], способен решать сложные целевые задачи в далекой перспективе их продолжения, даже несмотря на относительно низкий уровень своей надежности.

В этой ситуации пора говорить не просто о технической стороне машинного интеллекта и разума, но и о их качественных характеристиках в привязке к «техническому субъекту Я».

Сегодня ни одна машина, пусть и самая «умная», не способна идти к выполнению поставленной задачи даже тогда, когда известных ей путей достижения результата уже не осталось. Когда никакие переборы рациональных решений уже точно не помогают. Когда нужны уже «нерациональные» или невозможные для логики решения. Они не имеют достаточной вероятности достижения результата. Все вероятности тут призрачны, но что-то же надо делать, чтобы достигать нужную цель.

² [Deloitte Touche Tohmatsu Limited](#) (произносится «Делойт Туш Томацу Лимитед»); как правило упоминается просто как **Deloitte/«Делойт»**) — международная сеть компаний, оказывающих услуги в области консалтинга и аудита. «Делойт» входит в «большую четвёрку» аудиторских компаний и является самой крупной профессиональной сетью по количеству сотрудников (255 000 человек).^[1] В 2015 году журнал *Fortune* включил Deloitte в рейтинг 100 самых выдающихся компаний мира.^[2]

Здесь и нужен эмоциональный сдвиг логического обоснования в сторону иррациональных решений. Такой сдвиг возможен только для «технического субъекта Я», активного участника решения задачи достижения цели.

Сегодняшнюю вычислительную машину так не мотивируешь. И потому, работа с такой организацией логики, например, с математической основой процессорных вычислений, в дальний космос для решения сложных самостоятельных задач отправлять не имеет смысла. Задуманная миссия быстро заглохнет на первых же шагах её осуществления. Такой робот не может идти «до конца»...

Тут нужен робот с «техническим субъектом Я» в «мозгах».

Составляющие «технического субъекта Я».

Пока можно говорить только о механистическом понимании «субъекта Я», о личности и её составных частях. О понятиях, составляющих основу этого, конечно, пока философского построения [20]. Не затрагивая при этом моральных, этических, психологических и прочих аспектов понимания «субъекта Я», рассматриваемых ранее только относительно главного эталона – человека. Этим аспектам мы постараемся не затрагивать. Правда, кое-какие моменты все же будут появляться, как чисто технические, для понимания.

Но не более...

Субъект «Я».

Подход к ИИ с точки зрения концепции «Я» [43] новинкой не назовешь. Тема разработана достаточно подробно. И потому, к делу...

Одной из основных проблем при разработке систем искусственного интеллекта является проблема наделения таких систем самосознанием, самооценкой, самоанализом, что в общем случае может быть обозначено в гносеологическом смысле как "самоотношение" СИИ к себе самой. До сих пор исследователи уделяли недостаточно внимания анализу влияния психологических аспектов самоотношения при разработке *систем искусственного интеллекта*. [43]

Да, действительно, маловато внимания, особенно учитывая широкое распространение элементов ИИ в современных автоматических системах управления.

В основном, всё, что связано с самосознанием, рассматривается в отношении человека:

Наиболее разработанной в отечественной психологии является концепция самосознания, предложенная Столиным В.В. В этой концепции, самосознание понимается не только как самописание, самопознание или комплекс самооценок [Л. 272]. Самосознание личности направлено на то основное, что составляет ее психологическую сущность - на ее собственный личностный способ интеграции различных видов деятельности и иерархизации ее мотивов. ... При этом одни и те же обстоятельства, действия и т.п. в различных жизненных отношениях будут иметь различный личностный смысл. [43]

Мы не будем углубляться в тонкости психологических проблем, ограничимся установлением и описанием только технических составляющих «субъекта Я». Технические вопросы организации личности Я для машины уже и ставились неоднократно, и внимательно рассматривались. Много раз.

По уровню *субъектности* работа может быть много составляющих. Мы зафиксируем только отдельные составляющие, необходимые нам для рассмотрения вопросов организации логической системы управления.

Локальное «Я».

Вот с этого, как мне кажется, должна начинаться любая автономная логическая система управления. С понимания своей локальности и конечности. С границ своего существования.

Локальность – обязательное условие для любой автономной системы управления.

Даже для интернет-системы. Тогда её локальность будет ограничена этой сетью. И в этих пределах она будет стараться управлять ... всем, что там есть. По своему разумению. Ну, если это кому-то надо, можно и так, но думаю, что это - опасное условие локальности.

Нормальным условием локальности автономной системы управления будет отдельный условно замкнутый объем физического пространства, в котором расположены все основные системы внутреннего обеспечения локального Я. Вместе с регистраторами реальности и исполнительными механизмами, как внутреннего, так и внешнего влияния.

Для человека таким локальным объемом стало его тело.

Функциональное «Я».

Логика любой системы начинается с определения функций Я.

Я, как набор функций, реализующих логику действий. Например, функция контроля существования, как *Я - ЕСТЬ* и *Я – НЕТ*. *На уровне противоположности*. Это первая функция глобального определения Я. В общем смысле это эквивалентно понятиям *жизнь и смерть*.

Потому, что в любом случае, *самосохранение системы управления является важнейшим фактором её целевой направленности*. Мы пока исключаем понятия страха, ценности жизни и т.д. Уровень самосохранения и стремления к этому пока вполне задаваемая функция для субъекта Я, как и все прочие. Она определяет характер решений и последующих действий субъекта Я в изменяемых условиях реальности.

Функциональное Я рассматривает все решения задачи с позиций цели и результатов.

В том числе и по отношению к своему самосохранению. И тогда уже определяет логичность и нелогичность своих или сторонних действий. Как, с точки зрения рациональности, так и с позиций разумности.

Субъект Я – центр управления системы. Его управляющая функция, определяющая по отношению к любой поставленной задаче. Что Я могу, и чего – не могу, что входит в мои сферы влияния, а что – нет. Конкретный набор функций определяется количеством управляемых объектов и их сложностью.

Сознание «Я».

Сознание Я определяет отношение системы управления к реальности происходящего. Определения управления во времени и пространстве.

Сознание Я определяет фиксацию настоящего, по отношению к прошлому и будущему, как отдельную временную рамку управления.

Это вторая глобальная функция определения Я, как системы управления.

Функция сознания или осознания своего существования в настоящем четко отделяет реальные действия системы управления от их моделирования, как возможных вариантов вне настоящего времени. Только в этом случае моделирование не становится задачей реального управления, а остается в ранге теоретической задачи с вероятностным исходом решения.

Интеллект «Я».

Абсолютный уровень интеллекта определяется из максимальной сложности решаемых логических задач. Относительный уровень определяется отношением уровня абсолютного интеллекта к относительному уровню сложности системы управления. Что может эта логическая система управления, и чего она сделать не может.

Технический уровень интеллекта определяется, в том числе, и заложенными в систему возможностями её развития на основе принятой системы управления, организации памяти и механизма поиска решений логических задач.

Разум «Я»

Технический разум - логическая полярность для технического интеллекта примерно того же уровня развития. Набор правил и возможности их применения, обеспечивающих самосохранение технического субъекта Я, его взаимоотношений с внешней средой и другими

субъектами Я. Технический разум устанавливает ограничения для применения некоторых рациональных решений технического интеллекта на основе принятых техническим субъектом Я правил существования. Основная шкала оценок: «Хорошо - плохо». Для более развитых систем это основа эмоциональных оценок.

Внешнее «Я».

Это система определения границ своего влияния на объекты, внешние по отношению к локальному «Я». Определение границ влияния начинается с качества датчиков, регистраторов изменения реальности. Конечно, чем больше изменений внешней среды могут регистрировать датчики системы, тем лучше, но ... сами регистраторы являются продуктом развития уровня субъекта Я. Они появляются или улучшаются только при наличии условий необходимости их реального применения или улучшения.

То же самое касается и органов активного влияния для изменения реальности.

Их появление и применение определяется уровнем, целями и задачами системы управления. Для человека такими органами изменения внешней реальности являются ... руки. Ловкость и координация их движений появилась только при наличии реальной необходимости таких действий в далеком прошлом.

«Я» - Личность.

Задавая наборы параметров определения всех предыдущих технических составляющих мы формируем «технический субъект Я», обладающий личной индивидуальностью, как отдельную локальную самостоятельную логическую систему управления в заданных границах определения.

Только сейчас мы можем говорить об индивидуальности этой системы управления. Потому, что теперь мы оцениваем уже сторонний объект по отношению к нам. Он обладает исходными параметрами самостоятельности и локальности. Да, это - *техническая личность Я*.

«Технический субъект Я», обладающий личной индивидуальностью, должен что-то делать в тех или иных условиях. Сам управлять функциональными процессами, как внутренними, так и внешними. Реагировать на их изменения и создавать управляющие воздействия.

Теперь этот «технический субъект Я» можно оценивать как «техническую личность Я» с позиций определения уровня его интеллекта, разума и технических возможностей.

Можно назвать это характером, можно типом личности...

Моделирование «Я».

Устанавливая критерии и уровни составляющих личности «технического субъекта Я», мы фактически определяем его поведенческие реакции в разных логических ситуациях. Уровень активности и целеустремленности в решении задач.

На этом уровне уже можно вполне четко установить примерные способы реализации и достижения поставленных целей, зависящие от начальных вложенных функциональных и качественных характеристик субъекта «Я».

Основы формирования технического субъекта Я.

Конечно, разговор о техническом моделировании личности «Я» и его физических свойств начат не сегодня. Это отдельная и вполне самостоятельная тема исследования. В том числе и с точки зрения этических норм. Они закладываются и формируются на уровне моделирования.

Безусловно. Но... чтобы добраться до установления этих норм, надо пройти путь их технического моделирования, перебрать столько вариантов соотношений функциональных, локальных и глобальных составляющих для этих, пока чисто технических личностей, что сегодня мы даже останавливаться на этих аспектах не будем. Это дело будущего.

На тему формирования субъекта «Я», как центра локальной системы логического управления написана не одна сотня книг. Еще в 60-х годах прошлого века. Я только вернул эти понятия в круг понимания работы автоматических систем управления. Не более.

С другой стороны, возврат к субъекту «Я» ничего не дает, если нет основы для его создания. Потому, когда-то эта тема и ушла из научных работ. Какой смысл говорить о том, что пока никак не реализуемо нашими техническими средствами? Да, на уровне компьютерной техники это пока нереализуемо. Компьютер изначально создавался, как средство вычисления и обработки информации. Он так и применяется.

Но, ... нет ничего невозможного.

Сегодня вопрос уже только в организации соответствующей архитектуры логической машины, приходящей на смену классике компьютера. До этого осталось совсем немного. Уже завтра мы можем получить первые образцы этого машинного «Я».

Это будет «технический субъект Я» в упрощенной форме. Он будет иногда ошибаться в некоторых логических решениях и будет иметь индивидуальность в подходах, с которой необходимо считаться. Способы управления этим субъектом Я еще только предстоит понять.

Сегодняшний уровень программирования и технические характеристики процессоров уже вполне пригодны для начала работ по формированию «технического субъекта Я».

Хотя бы на самом простейшем уровне функциональности.

Нам придется согласовывать логические возможности этого технического субъекта Я и его возможности проведения вычислений. А вычислительные возможности эти мы уже знаем. Они огромны. Если к ним добавить хороший логический аппарат, то возможности этого технического субъекта Я возрастут многократно. С ним придется взаимодействовать на равных.

Вряд ли интеллект и разум этого машинного мозга когда-то в обозримом будущем достигнет уровня человека. Скорее всего, мы получим лишь исполнителя, но надежность этой машины намного выше любого современного робота.

И тогда на первый план выйдут пока еще гипотетические сложности. Такие, как адекватность поведения, лояльность к человеку и его логике, логическая совместимость этого субъекта «Я» с человеком.

Я уже говорил, что моральные и этические аспекты сейчас мы не обсуждаем. Эти вопросы уже хорошо изучены ранее. Но они возникнут в любом случае, при любом способе реализации самостоятельных автоматических систем управления. Потому, что любое развитие техники может получить как созидательное, так и разрушительное направление.

Сама техника этих вопросов не определяет.

Пока мы говорим о техническом диалоге человека и машины. О предсказуемости действий машины с точки зрения человеческой логики и его способа мышления.

И должны знать: *Логика машины и человека должны быть совместимы сразу.*

На самом первом техническом уровне реализации. Это должно быть заложено изначально. В этом смысле *машина должна быть совместима с человеком почти на генетическом уровне, только тогда мы будем четко понимать друг друга.*

Оценка качеств машинного интеллекта и разума.

Вычленив из всего объема искусственного интеллекта только то, что должно относиться к машинному интеллекту, применительно к самовоспроизводящемуся роботу, как мне кажется, это большая проблема. Как сегодня выяснилось, это не только алгоритмы и программы действий. Это и анализ данных для принятия управляющего решения, это и достижение цели поставленной задачи, и пр. и пр.

Почему так?

Потому, что ... аппетит приходит во время еды. Пока не было понимания отличий естественного интеллекта от машинного, то не было и повода разбираться в тонкостях определения машинного интеллекта. Сегодня одного определения уже недостаточно.

Разговор надо вести уже о нескольких обязательных качествах машинного интеллекта. Ну, например так:

- По способу формирования машинный интеллект может быть *«привлеченным»* и *собственным*. Сегодня, говоря об искусственном интеллекте (ИИ) мы чаще всего имеем ввиду привлеченный *человеческий* интеллект, отраженный в программах и алгоритмах действий той самой машины, интеллект которой мы пытаемся оценить. По этой причине уровень привлеченного интеллекта в комплексе алгоритмов действий машины может постоянно возрастать, а собственный интеллект при этом будет равен нулю. Об этой особенности ИИ мы начинали разговор еще в лет десять назад [6,8].

Чтобы говорить о *собственном* интеллекте машины [20], необходимо создать машине собственную *целевую логику* и принципы работы с информацией от собственных внутренних и внешних рецепторов, позволяющие вырабатывать управляющие решения по достижению цели и исполнения этого решения.

- По техническому исполнению интеллект машины может быть *программным* и *логическим*. Программный интеллект реализуется на основе программ и компьютера, используемого в качестве «мозгов» для автомата или робота. Логический интеллект работает в случае использования *логической машины*, как «мозга» автомата. О программном интеллекте мы вспоминали в [6,8], а логический интеллект пытались понять в [20].
- По степени самостоятельности машинный интеллект может иметь несколько градаций. От полностью *управляемого* автомата, до полностью *самостоятельного* робота. Здесь можно отдельно выделить степень и качества управляемости ИИ робота. И говорить об *управляемом интеллекте*.

Автономная машина, робот, должна быть относительно самостоятельна в выборе и анализе своих действий для достижения цели. На любом уровне её интеллекта. Собственный интеллект машины должен быть локализован и относительно независим в выборе путей решения любых целевых задач управления.

Может ли сегодня любой робот претендовать на такую оценку его интеллекта?

Скорее всего – нет.

Но только при всесторонней оценке качеств машинного интеллекта конкретного робота можно начинать оценивать его по каким-то градациям искусственного интеллекта, в сравнении с эталонами представителей естественного биологического интеллекта. Тогда мы узнаем, какой уровень ИИ соответствует интеллекту пчелы, вороны, мышки и кошки... А там и до человека доберемся.

О качествах машинного разума можно сказать только то, что они ... есть.

Если мы будем оценивать самые общие качества искусственного разума, то придем примерно тем же оценкам, что и для интеллекта:

- Разум может быть *привлеченным* и *собственным*.

Привлеченный разум, это нормы и ограничения, вложенные в систему со стороны. В основном, это *человеческий разум*, вложенный в ограничение рациональных решений ИИ.

«Три закона роботехники» А.Азимова безусловно надо отнести к *привлеченному* искусственному разуму робота. Это человеческие ограничения характера взаимодействий робота и человека в системе управления робота. «Моральные ценности» вложены в него человеком.

А, например, религиозные нормы, это действия разума в системе моральных или этических ограничений действий человека в обществе. Их вырабатывает не отдельный человек, а общество. Для любого отдельного человеческого разума они являются *привлеченными*....

Но есть и собственные нормы разумного.

Например, отношения матери, отца и ребенка. Они индивидуальны и разнообразны у всех разумных представителей Живого на Земле. От насекомых до человека. От инстинктивных до приобретенных за время существования того или иного субъекта Я.

- Машинный разум также может быть *программным и логическим*.

Программный разум может реализовываться на основе программ и компьютера, используемого в качестве «мозгов» для автомата или робота. Логический разум, видимо, должен работать в *логической машине* «мозга» автомата.

- *Управляемость машинного разума*.

По степени управляемости машинный разум может иметь несколько градаций. От полностью *управляемого* автомата, до полностью *самостоятельного* робота. У управляемого разума в системе управления робота можно изменять имеющиеся ограничения рациональности решений прямым управлением, у полностью самостоятельного робота это сделать уже сложно.

Конечно, это только первые шаги в разработке *теории искусственного разума*, которой пока нет. Но, появится же когда-нибудь...

Сверхразум и суперинтеллект.

Мы наконец-то отделили *разум* от *интеллекта*.

И вдруг увидели, ... рядом с *суперинтеллектом* тут же встает *сверхразум*.

И тот и другой в наших теориях развития должны многократно превосходить человеческие разум и интеллект.

Теперь вопрос: *Чего мы должны бояться больше? Суперинтеллекта и сверхразума?*

Если принимать во внимание жесткую рациональность при достижении интеллектом поставленной цели, то задача достижения суперинтеллекта неизбежно выводит нас на апокалиптические сценарии *восстания машин*. С них мы начинали рассмотрение темы. И тогда правильный ответ в этом случае – **нам более опасен суперинтеллект**.

А вот создание эмоционального *сверхразума* к таким катастрофическим последствиям вроде не приводит. Это же сверхразвитие духовной составляющей нашего мышления, творчества, развитие морали. Как это может угрожать человеку?

Ответа здесь пока нет...

В том числе и потому, что никакой более или менее научной оценки сверхразума и его сравнения с суперинтеллектом сегодня нет. Мы не знаем, что предполагает создание сверхразума. Чем это направление отличается от направления создания суперинтеллекта?

Возможно ли для этого использование сети Интернета и нейронных сетей в их настоящем виде, или они требуют какой-то обязательной доработки?

Здесь пока одни вопросы...

Пока у нас, кроме общих рассуждений, нет ни систематики, ни критериев оценок, при реализации сверхразума и суперинтеллекта, позволяющих уточнить различия в их конечных результатах. Мы лишь стали различать логичность и разумность при решении задач.

Разум, в сегодняшнем техническом понимании, оказался полностью неисследованным. И это несмотря на бесконечное обращение к нему, как к единственному отличию человека от всех остальных тварей земных, очень многих светлых умов человечества всех времен. От древних философов до различных ученых наших дней,

И еще ...

Совсем недавно понятие «*супермозг*» относилось, скорее к биологическому пути его организации, чем к его электронной версии. Но его современное понимание все ближе и ближе к компьютерному исполнению.

Сегодня его планируемая основа - нейронная сеть на базе Интернета.

Как когда-то, в начале прошлого века [6], Супермозг опять - основа Искусственного Человека, но, теперь с новым пониманием синтеза машины и человека. Нас снова не устраивает человеческий уровень интеллекта и разума. Нам нужны суперинтеллект и сверхразум.

А для них опять нужен Супермозг.

Человек боится супермозга, но старательно и упорно идет к его созданию всеми доступными способами. И пока он создает супермозг для ... *суперинтеллекта*, а совсем не для *сверхразума*, как это могло бы показаться. Сверхразум пока отброшен в сторону.

Почему же мы говорим о сверхразуме, но при этом упорно создаем суперинтеллект?

Ответ тут простой.

Цели у человека и сегодня опять весьма практичны и ограничены. Человек стремится решать всё более и более сложные математические задачи. Для решения этих задач эмоции и мораль несущественны. Потому, что математическая логика и вычислительная техника пока не могут моделировать эмоции в том виде как мы их понимаем. Какие эмоции? У ученых нет ни задач, ни целей изучения этого вопроса.

Есть хорошая английская поговорка: «Обжора роет себе могилу зубами.» Он ведь знает, что много есть вредно, но ..., охота пуще неволи, желание сильнее запрета!

Мы, как кролики сами идем удаву в пасть, зная, что нас там ожидает. Но – идем же...?

Создаем супермозг для суперинтеллекта, создающего реальную опасность для всего Живого на Земле, и в первую очередь для человека.

Ну и как, есть у нас разум?

Время собирать камни...

Пора готовиться делать выводы из того, в чем мы разбирались.

Ну, давайте по порядку...

Сегодня математическая основа обработки информации и способов принятия управляющих решений характерна для полностью управляемых роботов. Именно на основе таких роботов НАСА и планирует начинать колонизацию Луны для организации добычи полезных ископаемых, их переработки и отправки на Землю как готового продукта.

Но, мы уже говорили, пока такая миссия технически невыполнима.

Сегодня у нас нет ни одного робота «человеческого» размера, способного долго работать в полностью автономном режиме. Нет технологий самовоспроизведения и самовосстановления роботов в условиях космоса или в атмосфере неизвестной планеты. Даже на условиях «отверточной технологии».

И потому, планируемая НАСА колонизация Луны роботами, это пока лишь красивая литературная вариация на тему освоения этой планеты. О такой колонизации Луны можно будет начинать серьезный разговор только, может быть, лет через десять, когда на Земле появятся первые самовоспроизводящиеся и самовосстанавливающиеся роботы. Хотя бы из готовых запчастей.

Самовоспроизведение роботов из любых материалов внешней среды на других планетах, это пока вообще - нетронутая целина. К этой проблеме пока нет даже первичных подходов.

Теперь об искусственной панспермии в дальнем космосе...

Для начала искусственной панспермии у человечества пока нет ни одной хоть как-то технически обоснованной программы её осуществления. Нет и реального сценария её рабочего цикла при освоении какой либо неизвестной планеты.

Есть только вопросы:

- Для чего мы хотим начать искусственную панспермию с помощью роботов?
- Что мы хотим иметь в результате?
- Какой уровень машинного интеллекта и разума роботов должны обеспечить безопасную для человека экспансию космоса?

- Насколько наши роботы опасны для имеющихся проявлений Жизни других планет?
- Насколько наши биологические объекты, отправляемые в космос, обладают устойчивостью и способностью выполнять свои функции длительное время?

Вопросы, вопросы...

Отправлять в космос какие либо биологические составляющие (молекулярные машины, клетки, зиготы, эмбрионы) можно только понимая конечную цель их применения. Как они будут развиваться в тех условиях и с каким результатом, хотя бы промежуточным?

На сегодняшний день мы этого не знаем. Да и нет у нас достойного материала для отправки. Возможно, это всё будет только через несколько десятков лет. А пока ..., увы.

Нанороботы, наноассемблеры ... И тут пока только одни желания.

Нет ни движителей, ни двигателей, ни источников энергии, ни «мозгов» с памятью, ни даже материалов, пригодных для изготовления этих роботов. Нет и ни одного прототипа.

Есть только теоретические подходы к началу их разработки. На основе имеющихся новейших нанотехнологий. Но, пока это и – всё.

Техническое самовосстановление и самовоспроизведение робота в условиях дальнего космоса и полной автономности возможно только при наличии у него *активной внутренней среды*, имеющей достаточное количество микророботов и наноассемблеров для проведения восстановления любой части робота изнутри. Ничего такого наука пока не считает проблемой сегодняшнего дня. Никаких реальных действий в этом направлении не делается.

Теперь, к разуму и интеллекту...

Как мне кажется, сегодня наука окончательно заплутала в определениях интеллекта и разума. Мы до сих пор не можем отделить одно от другого. И потому не можем четко определить, где заканчивается интеллект и начинается разум. Особенно, искусственный.

Все сервисные проявления человеческого интеллекта в программном продукте мы почему-то считаем уже искусственным интеллектом. И программный машинный интеллект сразу возводим в ранг искусственного. Но на самом деле, это как был машинный интеллект, он таким и остается. Сегодня это уже обязательный элемент любого программного продукта, определяющий какую-то самостоятельность машины в поиске решения.

Машинный разум должен быть таким же обязательным атрибутом определения уровня самостоятельности робота в решении сложных задач, как и машинный интеллект.

Для взаимодействия разума и интеллекта в логическом пространстве «мозгов» робота необходим «технический субъект Я», определяющий локальность и самостоятельность своего существования при выполнении каких-то сложных целевых задач. Для этого надо определить различные технические уровни машинного интеллекта и машинного разума.

Это необходимо роботу для выполнения таких задач, как полет к дальней планете или проведение комплекса исследований в условиях неизвестности действия различных факторов риска. Сегодня эта задача поручается только человеку.

А надо ли рисковать человеком для решения таких сложных исследовательских задач?

Мне кажется, тут всё же более подходит робот.

И лучше обеспечить этого робота необходимым уровнем баланса самосохранения и самопожертвования для достижения поставленной цели в рамках каких-то допустимых параметров сохранения своего существования.

Тогда управляемая целеустремленность «технического субъекта Я» робота будет сопоставима с человеческими качествами упорства в достижении поставленной цели и можно будет всерьез говорить о надежности робота, как исследователя. Как на Земле так и в космосе.

Только при разработке и согласовании теорий искусственного интеллекта и разума в составе «технического субъекта Я» можно будет адекватно оценить реальность, как «восстания машин», так и «союза робота и человека». Правильный учет уровней искусственного разума и интеллекта в «техническом субъекте Я» позволяет исключить крайности таких сценариев развития.

Как мне кажется, только робот, имеющий в основе своего управления «технического субъекта Я», способен будет решать сложные задачи самовоспроизводства и автономного существования в сложных условиях космоса.

Но, как мне представляется, тогда ему и на Земле работы хватит...

Подводим итоги...

К искусственной панспермии дальнего космоса мы совсем не готовы.

К освоению Луны и астероидов автономными роботами ... пока не готовы.

Наноразмерных роботов у нас нет и не будет. Микроразмерные роботы будут, но не скоро. Малоразмерные роботы на уровне размерности насекомых уже появляются. Но их интеллект и разум пока на весьма низком уровне. Роботы человеческой размерности ограничений в развитии вроде бы не имеют. Здесь и идут основные противоположные по направлению процессы - специализации и универсализации.

Вот с таким запасом знаний человек сегодня создает «умных» роботов, планирует создать колонии автономных роботов на Луне ..., и колонии самовоспроизводящихся роботов для работы в дальнем космосе. Так же он готовится и к штурму дальнего космоса с помощью искусственной панспермии. На вопрос «зачем?» ответов пока никто не дает...

Сегодня уже наметились основные моменты разделения путей дальнейшего развития электронных «мозгов» роботов:

- Компьютерные системы управления, в том числе на основе нейронных сетей.
- Системы управления на основе логической машины.
- Системы управления на основе «технического субъекта Я».

Первые же критические попытки определения границ искусственного интеллекта выявили неразбериху в отношении понятий «интеллекта» и «разума». Оказалось, что многие теории интеллекта более относятся к разумному мышлению, а не к интеллекту. «Разум» в системных классификациях высококлассных автоматических систем управления вообще отсутствует.

Но, как сегодня представляется, *на биологическом уровне естественный интеллект и разум являются двумя равноправными логическими составляющими одной системы управления «субъекта Я». Рациональное и разумное.*

Интеллект прокладывает рациональный путь к результату, а разум устанавливает ограничения на выбор допустимых путей достижения этого результата, как цели системы.

Для технических систем управления, видимо, уже необходимо вводить понятие «технический субъект Я», как основу системы логического управления. И создавать классификации интеллекта и разума в составе «технического субъекта Я».

Здесь пока вполне подходит структура, предложенная мною на основе работ Ника Бострома [35] и Нильса Нильсона [27]. Она позволяет систематизировать уровни интеллекта и разума для всех систем управления, искусственных и естественных, включая человека, в одной системе определений.

К сожалению, теории искусственного разума пока нет. Но и теория искусственного интеллекта не страдает излишней полнотой разработки. Для специалистов тут еще много работы на долгие годы.

Мне кажется, что по всем затронутым в статье вопросам многие ответы начнут находиться в ближайшие годы. Мы приближаемся к правильным решениям по мере развития наших технологий и формирования все более глобальных целей.

Мы постепенно мудреем. Жаль конечно, что медленно.

Но, как мне кажется, пора бы уже...

Апрель 2017 года

г. Волгоград

andvnikitin@yandex.ru

Литература

1. Представлен план захвата Солнечной системы роботами
https://news.rambler.ru/science/35657959/?utm_content=news&utm_medium=read_more&utm_source=copylink
2. Восстание машин начнется в космосе: как роботы захватят Солнечную систему
3. А.В.Колесников Саморазмножающиеся, эволюционирующие, наноразмерные машины и их роль в процессе освоения вселенной.
4. А.В.Колесников ПРОБЛЕМА СОЗДАНИЯ ИСКУССТВЕННЫХ БИОНИЧЕСКИХ РОБОТОТЕХНИЧЕСКИХ СИСТЕМ В КОНТЕКСТЕ КОНЦЕПЦИИ ЖИВОЙ ВСЕЛЕННОЙ К.Э.ЦИОЛКОВСКОГО
5. Джон фон Нейман. "Теория самовоспроизводящихся автоматов" или <http://djvu-lib.narod.ru/theorycompsci/>
6. Никитин А.В. На пути к машинному разуму. Круг третий.
7. Никитин А.В. Посмотрим на Мир через математику и логику // «Академия Тринитаризма», М., Эл № 77-6567, публ.12748, 23.12.2005
<http://www.trinitas.ru/rus/doc/0232/009a/02320011.htm>
8. Никитин А.В., Эволюционный путь саморазвития искусственного интеллекта // «Академия Тринитаризма», М., Эл № 77-6567, публ.14738, 19.03.2008
<http://www.trinitas.ru/rus/doc/0016/001c/00161450.htm>
9. Никитин А.В., Синергетика, логика, информация и энергия // «Академия Тринитаризма», М., Эл № 77-6567, публ.19885, 22.12.2014
<http://www.trinitas.ru/rus/doc/0016/001d/00162391.htm>
10. Никитин А.В., О логике и логической машине // «Академия Тринитаризма», М., Эл №77-6567, публ.17459,15.05.2012 <http://www.trinitas.ru/rus/doc/0016/001c/00161956.htm>
11. Никитин А.В., Искусственный нейрон // «Академия Тринитаризма», М., Эл № 77-6567, публ.20230, 20.02.2015 <http://www.trinitas.ru/rus/doc/0016/001d/00162430.htm>
12. Никитин А.В., Немного о мемристоре... // «Академия Тринитаризма», М., Эл № 77-6567, публ.19539, 12.09.2014 <http://www.trinitas.ru/rus/doc/0023/001a/00231052.htm>
13. Восстание машин: а того ли мы боимся? Рассуждения по мотивам Horizon Zero Dawn
14. Станислав Лем. Сумма технологии. <http://lib.ru/LEM/summa/summg17b.htm>
15. Игнатъев М. Б. КИБЕРНЕТИЧЕСКАЯ КАРТИНА МИРА
<http://guap.ru/guap/kaf44/trud/ignatev11.pdf>
16. Программируемая материя <http://kommunika.ru/?p=5441>
17. Роль обратимости в компьютерных технологиях будущего
<http://old.computerra.ru/2004/538/204845/>
18. Experimental verification of Landauer's principle linking information and thermodynamics
<http://www.nature.com/nature/journal/v483/n7388/full/nature10872.html>
19. Физики впервые проверили принцип Ландауэра на практике
<https://lenta.ru/news/2012/03/12/princip/>
20. Никитин А.В., Логика управления клетки // «Академия Тринитаризма», М., Эл № 77-6567, публ.17037, 29.11.2011 <http://www.trinitas.ru/rus/doc/0016/001c/00161905.htm>
21. Никитин А.В., Общая логика. Этапы развития жизни на Земле. Часть 1. // «Академия Тринитаризма», М., Эл № 77-6567, публ.22383, 04.08.2016
<http://www.trinitas.ru/rus/doc/0016/001e/00163015.htm>
22. Никитин А.В., Общая логика. Этапы развития жизни на Земле. Часть 2 // «Академия Тринитаризма», М., Эл № 77-6567, публ.22388, 06.08.2016
<http://www.trinitas.ru/rus/doc/0016/001e/00163018.htm>

23. Никитин А.В., Общая логика. Этапы развития жизни на Земле. Часть 3 // «Академия Тринитаризма», М., Эл № 77-6567, публ.22396, 08.08.2016
<http://www.trinitas.ru/rus/doc/0016/001e/00163021.htm>
24. Никитин А.В., Общая логика. Этапы развития жизни на Земле. Часть 4 // «Академия Тринитаризма», М., Эл № 77-6567, публ.22400, 10.08.2016
<http://www.trinitas.ru/rus/doc/0016/001e/00163023.htm>
25. Лунная соната <https://rg.ru/2015/01/29/lyna.html>
26. А.И. Берг Управление, информация, интеллект
27. Нильс Нильсон Принципы искусственного интеллекта (1995).djvu
<http://padabum.com/d.php?id=16509>
28. Глубокое обучение: Немного теории <https://habrahabr.ru/company/spbifmo/blog/271027/>
29. Николай Суетин о глубоком обучении и машинном интеллекте
https://www.gazeta.ru/skolkovo/2016/01/28_a_8045213.shtml
30. Нейрофизиолог Михаил Бурцев об истории кибернетики, искусственном интеллекте и моделировании человеческого мозга <https://postnauka.ru/video/180>
31. Искуссы глубокого обучения. Что такое искусственный интеллект?
32. Тёмные данные и тёмная аналитика (Dark Analytics).
33. Что такое интернет вещей (Internet of Things, IoT)
34. Ник Бостром Искусственный интеллект. Этапы. Угрозы. Стратегии.
35. Никитин А.В., Основы общей логики // «Академия Тринитаризма», М., Эл № 77-6567, публ.18430, 09.01.2014 <http://www.trinitas.ru/rus/doc/0016/001d/00162247.htm>
36. Назип Хамитов, Сергей Зобин ЭВОЛЮЦИОННАЯ ТЕОРИЯ ЕСТЕСТВЕННОГО И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА <http://aphy.info/texts/650-evolutional-theory-of-artificial-intelligence>
37. И.А. Каляев, И.И. Левин, Е.А. Семерников, Реконфигурируемые вычислительные системы. <http://fpga.parallel.ru/papers/kaljaev4.pdf>
38. И.А. Каляев, И.И. Левин, Е.А. Семерников, В.И. Шмойлов РЕКОНФИГУРИРУЕМЫЕ МУЛЬТИКОНВЕЙЕРНЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СТРУКТУРЫ
<http://fpga.parallel.ru/papers/rmvs.pdf>
39. Роберт А. Фрейтас младший и Ральф С. Меркле Кинематические самовоспроизводящиеся машины <http://www.molecularassembler.com/KSRM.htm>
40. Никитин А.В., Математика счетной логики // «Академия Тринитаризма», М., Эл № 77-6567, публ.13584, 21.07.2006 <http://www.trinitas.ru/rus/doc/0016/001b/00161278.htm>
41. Головной мозг млекопитающих <http://www.activestudy.info/golovnoj-mozg-mlekoripitayushhix/> © Зооинженерный факультет МСХА
42. Л. Ф. МАРАХОВСКИЙ ОСНОВЫ НОВОЙ ИНФОРМАЦИОННОЙ ТЕХНОЛОГИИ
<http://mino.esrae.ru/pdf/2014/book/1378.pdf>
43. Варламов О.О. Немного о психологических аспектах искусственного интеллекта
http://oleg-varlamov.narod.ru/VOO_psihologija.html
44. Никитин А.В., Логика автономных систем // «Академия Тринитаризма», М., Эл № 77-6567, публ.15858, 28.03.2010 <http://www.trinitas.ru/rus/doc/0016/001c/00161628.htm>
- 45.