

Когда начинать бояться ИИ?

Часть 5. О «субъектности» машины....

Оглавление

ЭТАПЫ РАЗВИТИЯ ИИ.....	3
УПРАВЛЯЕМАЯ СЛОЖНОСТЬ.....	7
ЧТО ТАКОЕ ИИ?.....	12
ЧТО МЫ ПОНИМАЕМ КАК СУБЪЕКТ?.....	14
СИСТЕМА УПРАВЛЕНИЯ ОРГАНИЗМА.....	17
ПОРА ЗАКАЧИВАТЬ.....	19
ЛИТЕРАТУРА.....	20

Если сейчас конкретизировать наши страхи по поводу ИИ и попробовать разделить их по направленности, то мы увидим вполне закономерную картину.

Большую часть, может быть до 80%, составляют страхи, порождаемые человеческими пороками, желанием обмануть и заработать на этом. И только потом ошибками программирования элементов ИИ в составе бытовой техники вокруг нас. Основной «ответственный за наши страхи» - человек, чаще всего создающий преднамеренные предлоги для этого, а совсем не ИИ.

Значительно меньшая часть страхов связана с разочарованием в отличиях реальности работы ИИ от мечты об этом. Человек впадает в уныние и начинает бояться ИИ уже с другой стороны, как «переоцененное чудо», которое не готово исполнять то, что ему пророчат. Он боится не самого ИИ, а человеческих сказок о его пока несуществующих возможностях.

Лишь малая часть людей всерьез задумывается о том, что когда-то возникнет Супер ИИ, который может уничтожить человечество. Такие есть, и их страхи тоже небеспочвенны. Чего же боятся они, если отбросить всякие «сказочные» компоненты о «войне машин», «машинной эволюции» и т.д.?

Они боятся, давайте обозначим его так, «субъектного ИИ». Машины, имеющей ИИ, которая проявляет качества субъекта – *«индивида, познающего внешний мир (объект) и воздействующего на него в своей практической деятельности;»* https://dic.academic.ru/dic.nsf/enc_philosophy/1174/СУБЪЕКТ

Потому, что именно об этом варианте ИИ рассказывают «страшные сказки на ночь» все книги философов и ученых, занимающихся теорией ИИ. Именно об этом типе ИИ рассуждают серьезные ученые на страницах серьезных журналов и на экранах ТВ.

От этого создается впечатление, что всё сказанное ими – реальность, вот он Машинный Субъект с ИИ, который уже сегодня готов быть и другом, и врагом. Осталось только кнопку нажать.

Но вот только кнопка эта... где-то там, за горизонтом. Примерно там же находится и наша «розовая мечта» о субъектном ИИ, обладающем разумом... Реально об этом говорить не приходится.

Причины этого вполне понятны.

Любая мечта идет к реализации «по трясине» непознанного, которое часто оказывается совсем не тем, что мы знаем сегодня. И каждый шаг это угроза для имиджа, научного авторитета и пр. и пр. Тут лучше идти не торопясь, взвешивая каждый шаг имеющимися обоснованиями. За мечту мало платят, «лучше синица

в руках, чем журавль в небе»... означает, что приоритет решения сегодняшних задач более важен, тем более, если от этого зависит твое благополучие.

И мечта так и остается только мечтой.

Сегодня мы имеем явный разрыв между философским пониманием субъектности машины и её техническим описанием в научных трудах. Философы о субъектности машины уже тонну книг написали, а ученые только несколько статей полвека назад выпустили с первыми робкими попытками понять, что это за «субъектность» такая в техническом приложении к машине.

Конечно, мы говорим об «Искинте», о статьях Д.А. Поспелова и его сподвижников [1÷5]. Ранее здесь уже были показаны функциональные схемы «думающей машины» и «гиромата», как прообраза «машинного мозга». Они первыми показали, что субъектность не появляется ниоткуда, это ступень развития системы самоуправления в составе эволюционирующего множества [11]. Причем, не только развития клеточного множества до организма, но и множества организмов...

Субъектность - уровень самоуправления организма, качественный результат развития коллективного интеллекта.

Наука как-то этот урок кибернетики подзабыла.

Философия давно дала свое понимание субъектности, примерно так: «субъектность — свойство индивида быть субъектом активности». https://psychology_lexicon.academic.ru/294/Субъектность

Нет, ну а что, коротко и ясно.

Там дальше еще много чего написано, но главное «субъект активности», как «носитель предметно-практич. деятельности и познания (индивид или социальная группа), источник активности, направленной на объект». https://dic.academic.ru/dic.nsf/enc_philosophy/1174/СУБЪЕКТ

Всё.

Никакого технического понимания информации нет. Мы просто не знаем, откуда начинать, куда идти и что понимать. Это что-то вроде «пойди туда, не знаю куда, принеси то, не знаю что».

Книг, как всегда, на эту тему много, а что это такое..., «техническая субъектность», не знаем. Моделировать в системе самоуправления машины пока особо нечего. Техническая наука сюда еще не добралась, чтобы понять, как организована функция субъектности у живых организмов, включая человека.

А доберется, с чего начнем?

Тут вспоминается какая-то научно-популярная статья по ИИ из середины прошлого века.... Примерно так: давайте исключим из рассмотрения «бога», «душу», да и «личность» тоже, как недоказанные наукой. Что тогда будет представлять для нас любое живое существо, включая человека?

Биологический автомат. Который реагирует на свои нужды и включает программы поиска логических решений по их удовлетворению. Реагирует на возникающие изменения и делает необходимые действия по сохранению своего существования. Вплоть до самокопирования или размножения.

И теперь мы можем сравнить автоматы. Биологический и электронный.

Но тут сразу начинается непонимание.

У философов человек и какой угодно сложный автомат ...почему-то никак не сравниваются. В любом случае, человек с одной стороны, а все остальное, включая как живые организмы, так и автоматы – с другой.

При этом вся наука, включая биологию, считает живой организм, в том числе и человеческий - автоматом, с обратными связями, взаимодействующим с

внешней средой и сохраняющим свое существование. Мозг нужен организму только как машина по обработке информации. [23, 24]

А мы сравним «живое» и «неживое». И зададим себе вопросы...

Как в электронном автомате сделать то, что природа уже сделала в биологическом? Как это реализовано и можно ли повторить?

Это же мечта науки из конца позапрошлого века!

Но только сегодня эта мечта пришла к пониманию реальной сложности своей реализации.

Начинаем...

Этапы развития ИИ...

Вычислительная техника, реализующая функции ИИ подошла к очередному тупику. Надо снова искать путь к новому уровню сложности и достижимых возможностей.

Но, почему опять тупик?

Давайте просто вспомним этапы развития ИИ...

Первый этап все уже знают. Это этап «старого доброго символического интеллекта». Мы же помним... *«В 1985 году Джон Хогланд дал символическому ИИ название GOFAI (англ. Good Old-Fashioned Artificial Intelligence, «старый добрый искусственный интеллект») в своей книге Artificial Intelligence...»* <https://ru.wikipedia.org/?curid=6178634&oldid=127563354>

Основным направлением этого вида ИИ стала вопросно-ответная система, работающая с пользователем на естественном разговорном языке.

«Первые QA-системы были разработаны в 1960-х годах и являлись естественно-языковыми оболочками для экспертных систем, ориентированных на конкретные области. Современные системы предназначены для поиска ответов на вопросы в предоставляемых документах с использованием технологий обработки естественных языков (NLP).» <https://ru.wikipedia.org/?curid=200217&oldid=105382417>

Второй этап, начался в той же парадигме – Искинт.

Его основоположник Д.А. Поспелов увидел путь развития машинного интеллекта через создание практически, искусственного мозга, с блоками, отвечающими за различные функции создаваемого машинного интеллекта.

Где-то в это время кибернетика переросла свое начальное предназначение и выродилась сначала в *кибернетический* подход решения сложных задач, как – *«исследование системы на основе принципов кибернетики, в частности с помощью выявления прямых и обратных связей, изучения процессов управления, рассмотрения элементов системы как неких " черных ящиков " (систем, в которых исследователю доступна лишь их входная и выходная информация, а внутреннее устройство может быть и неизвестно).»* <https://studfile.net/preview/4272013/page:14>

От «кибернетического подхода» исследования доросли до «системного подхода» [15], предложенного академиком И.В. Прангишвили. Здесь *любая, самая сложная система, даже многоуровневая, построена из множества отдельных автоматов с гомеостазом. Вся сложная система работает на принципах самостабилизации и самоорганизации, изложенных в синергетике: Хаос и энтропия, процессы увеличивающие их и уменьшающие. Система антагонистов, как принцип уравнивания противоположностей.*

Это все правильно, и...

Недавно в МИФИ была сделана новая программа, «ассистент» композитора для сочинения музыки.

«Ассистент композитора — это программное обеспечение, которое способно самостоятельно создавать музыкальные композиции на «человеческом» уровне, высокого эстетического качества. Ассистент — «умный» творческий помощник, который способен дополнять ноты, написанные композитором, «собственными» нотами, аккордами и их комбинациями, сотрудники Института интеллектуальных кибернетических систем НИЯУ МИФИ.

...Новая программная система принципиально отличается от похожих разработок тем, что виртуальный агент способен поддерживать эмоциональный контакт с человеком,...

...метод основан не на нейросети, а на семантической карте, где состояние человека представлено точкой в аффективном пространстве. Каждое возможное действие также имеет координаты в этом пространстве, и существует некоторый закон, связывающий одни координаты с другими, следуя логике эмоционального восприятия.

Есть также ограничения, определяемые законами музыкальной гармонии и ритма.

«Однако ограничения не могут определять все и оставляют некоторую «свободу для творчества». Эту свободу и использует наш модуль, руководствуясь семантической картой и правилами выбора координат на ней с учетом эволюции эмоционального состояния человека. Замечу, что здесь речь идет не об обучении: модель выводится из «первопринципов» психологии (базисных психологические аспектов поведения). В этом смысле такая модель может объяснить суть явления творчества, в то время как нейросеть не может», — пояснил Алексей Самсонович.» УЧЕНЫЕ НИЯУ МИФИ РАЗРАБАТЫВАЮТ ВИРТУАЛЬНОГО ПОМОЩНИКА КОМПОЗИТОРА. 02 АВГУСТА 2018 <https://mephi.ru/content/news/1387/127652>

Здесь вдруг выявились целые направления изучения эмоций, которые ранее технической наукой вообще не затрагивались или отбрасывались. Например, «аффективная неврология - изучение того, как мозг обрабатывает эмоции. Основа эмоций и то, что такое эмоции, остается предметом дискуссий в области аффективной неврологии.

Термин "аффективная неврология" был введен нейробиологом Яком Панксеппом...» https://en.wikipedia.org/wiki/Affective_neuroscience

Разработчики приложения «ассистент композитора» нашли способ разместить какое-то аффективное пространство в каких-то координатах и находить соответствие эмоционального состояния «внешнего управления» на своей семантической карте этого пространства. Получилось соотнесение эмоционального состояния «внешнего управления» с формируемым аффективным пространством машины в координатах семантической карты.

Видимо речь идет о семантической сети, это — «информационная модель предметной области, имеющая вид ориентированного графа, вершины которого соответствуют объектам предметной области, а дуги (рёбра) задают отношения между ними. Объектами могут быть понятия, события, свойства, процессы. Таким образом, семантическая сеть является одним из способов представления знаний.» <https://dic.academic.ru/dic.nsf/ruwiki/1143575>

Что такое «эмоциональное состояние» «внешнего управления», можно вполне четко представить по книге Е.П. Ильина [6]. Сложность потрясающая!

Но разработчики сделали первый шаг к пониманию, и похоже, правильный. Аффективное пространство отражается на семантической карте.

Вот это самое главное в этой разработке. Разработчики создали реальное направление развития машинного виртуального пространства!

Программа создает эмоциональный настрой «внешнего управления», в данном случае, сочиняющего музыку человека, ждет действий, в данном случае, набора нот, создающих мелодию. И начинает дополнять её своими доработками, звуками и аккордами, делая мелодию многозвучной и технически законченной в русле общего эмоционального настроения по «универсальной» «семантической карте». Видимо разговор идет о поиске основных мелодий без плагиата, об её аранжировке и т.д.

Программа «обучается» изменениям эмоционального фона конкретного Композитора, данные заносятся в семантическую карту. И эта карта используется машиной для определения эмоционального состояния Композитора во время работы программы, чтобы «помощь» работала в том же эмоциональном направлении. Хороший автомат, адаптируемый к каждому сеансу взаимодействия с «внешним управлением». Программа безусловно нужная, только...

Получается, делаем шаг вперед и два назад...

Потому, что с одной стороны, эта программа четко реализует принципы, заложенные системным подходом И.В. Прангишвили. А с другой, мы понимаем, что из этой колеи разработчики выскочить уже не могут. Они так и будут отрабатывать то, что знают, как когда-то Д. А. Поспелов шел по своей колее и тоже не видел выхода из символического способа решения задач...

Но как мы знаем, тогда выход был найден и ИИ перешел на следующий уровень разработок. На уровень *нейросетей*. Здесь незаметно и уже давно начался **третий этап** развития ИИ...

Он начался с нечеткой логики Л.Заде и псевдофизических логик Д.А.Поспелова. Эти логики реализованы в системах обучения нейросетей [18, 19]. Сегодня изображения формируются вокселями, а языковые форматы, эмбедингами [22], векторным представлением слова, числом.

По сути, человек создает свою систему формирования и обработки информации из сигналов. Она медленно, но начинает напоминать механизм получения информации мозгом. Специалисты по ИИ и не скрывают этого.

Для этого и расшифровываются сложные связи естественных нейронных сетей головного мозга, уточняется карта его функциональных областей, моделируется работа нейронного пространства, приближенного к естественному. Ученым уже многое стало понятным. [20]

Сетевое программирование, кажется, может всё...

Но...

Экспертная система на основе нейронных сетей свое заключение дает, а чем оно обосновано – непонятно. Ученые пытаются в этом разобраться на основе интерпретируемости заключений системы, с позиций «черного ящика»:

«Бин Ким. Есть две разновидности интерпретируемости. Одна — это интерпретируемость для науки: если рассматривать нейронную сеть как объект исследования, то можно провести научные эксперименты, чтобы по-настоящему понять скрытые характеристики модели, как она реагирует на различные стимулы и т.д. Вторая разновидность интерпретируемости, которой я, в основном, занимаюсь, это интерпретируемость «ответственного ИИ».

Не нужно понимать о модели всё, достаточно лишь понимать, как безопасно ей пользоваться. Это и есть наша цель.

Павлус. *Но как можно доверять системе, если не до конца понимаешь, как она работает?...*

Б. К. *Да. Вторая разновидность интерпретируемости пытается ответить на вопрос: понимаем ли мы инструмент достаточно, чтобы можно было безопасно его использовать? И мы можем получить это понимание, подтвердив, что в инструменте отражены полезные человеческие знания.*
<https://22century.ru/popular-science-publications/understanding-machines>

И чем дальше, тем ситуация только усложняется. Оказалось, что для науки проблемы возникли огромные.

Вот пример:

«Созданное программное обеспечение принимает решение на основе данных/знаний, поэтому качество этих данных является ключевым аспектом правильности работы программного обеспечения.

Есть очевидные и наиболее часто упоминаемые причины некачественной работы вне заявленных характеристик ИИ в медицине:

- 1. Ошибки в использовании программных библиотек.*
- 2. Некачественно собранные медицинские данные.*
- 3. Недостаточное количество данных или выборки данных для обучения алгоритма.*
- 4. Нерепрезентативная выборка (не все данные, не по всем клиническим случаям оказались в дата-сете).*
- 5. Ошибки в разработке программного продукта.*

Это лежащие на поверхности, часто упоминаемые причины.

Однако есть ряд других причин, которые порой не учитываются в разработке, но могут носить фатальный характер в принятии медицинских клинических решений:

1. Искажение первичных медицинских знаний.

- Приверженность эксперта дата-сайентиста к устоявшейся точке зрения в предметной области, которая продиктована личным опытом и сложившимся стереотипом клинического мышления, либо приверженностью к определенной «Школе».*
- Искажение знаний, полученных в ходе клинических исследований.*
- Публикации в научной литературе изначально в своем большинстве подтверждают первоначальную гипотезу исследователей.*
- Данные с отрицательными результатами клинических исследований обычно не публикуются.*
- Публикация отрицательного опыта клинической практики, врачебных ошибок, неудач лечения является сложным с этической и юридической точек зрения. Врачи редко говорят о своих ошибках на страницах высокорейтинговых научных журналов.*
- «Заказной» характер клинических исследований – за многими исследованиями с «правильным дизайном» ... стоят заказчики – представители медицинской индустрии, которые заинтересованы в подтверждении своей первоначальной гипотезы.*
- В силу бизнес-интереса проводимых исследований на них выделяется значительное финансирование, которого хватает на «правильный дизайн» исследования, создание рабочей группы, включение в исследование пациентов, аналитику и публикацию в высокорейтинговых научных журналах.*

- Эти правильно построенные клинические исследования влияют на клинические рекомендации, референсные значения физиологических показателей. Так, например, под влиянием подобных клинических исследований были изменены референсные показатели артериального давления (АД), сахара крови, эректильной дисфункции (МИЭФ-5).

2. Отсутствие знаний или недостоверные знания о предметной области.

...На приеме практически нет возможности сказать пациенту, что врач не знает какой ставить диагноз или как его лечить. Сложилось мнение, что пациент всегда должен выйти от врача с решением своей проблемы. ...На этом фоне очень сложно специалисту, который несет юридическую ответственность, сказать, что он не знает или не уверен в диагнозе или методах лечения.

3. Социальные искажения.

- Перспективность применения к пациенту того или иного метода, использование дорогостоящих методов диагностики и лечения, прогноз развития заболевания, этаназия, личное отношение врача к пациенту и другие моменты – лежат за рамками медицины, основанной на доказательствах. Принятие решения врачом в таких ситуациях носит иррациональный характер и лежит в рамках этических, культурных, а порой и религиозных аспектов. Такие знания очень сложно на сегодняшнем уровне собрать, классифицировать и алгоритмизировать. Как правило, решения принимает врач на свой страх и риск, под свою ответственность, не афиширует, и тем более, не публикует.

В ряде клинических ситуаций принятие врачом таких решений носит ключевой характер, но не может быть протоколировано и алгоритмизировано в программном обеспечении в ИИ.» <https://jtelemed.ru/article/slabye-storony-iskusstvennogo-intellekta-v-medicine>

Это и называется тупик.

Тут уже никакие даже самые «ответственные» ИИ ничего не докажут, когда в дело вступает предустановка программной защиты заказчика и разработчика от ответственности.

Выход из тупика, следуя теореме Тарского о невыразимости истины, находится на другом уровне понимания. Помните: «...понятие арифметической истины не может быть выражено средствами самой арифметики.» <https://ru.wikipedia.org/?curid=1534889&oldid=130455819>

Где же выход? Если честно – не знаю....

Управляемая сложность.

Похоже, что мозг моделирует и формирует управляемую сложность. Чем большую управляемую сложность он может смоделировать и использовать, тем сильнее его интеллект.

Мы в строительстве своих машин приблизились к пределу сложности автоматов. Наши машины стали сопоставимы по сложности с живыми организмами, но лишены средств поддержания своего существования. Пока они не могут сами себя ремонтировать, как живые организмы залечивать свои раны.

Не могут сами себя воспроизводить даже в частях, как животные делают это регенерацией своих составляющих, если возникает такая необходимость и т.д.

Надежность работы наших машин пока считается по надежности самого малонадежного элемента. Повышение надежности работы наших машин мы проводим резервированием и дублированием самых важных её составляющих.

Но фактически это означает, что «комплект одной машины» мы формируем из двух и более реальных составов машины. Это как-то помогает когда машина относительно простая, можно поддержать её работоспособность таким способом для увеличения надежности работы в сложных условиях.

Но, что делать, если машина очень сложная?

Тут такой путь не очень подходит.

Это одна из причин появления *теории сложных систем* о которой мы говорили с конце прошлой части.

О чем мы говорим?

Сегодня наука и промышленность создали вполне стройную систему машин для реализации концепции глобального искусственного интеллекта на основе многослойных нейросетей. На планете создана сеть дата-центров, связанных каналами связи для обмена и совместной обработки огромных объемов информации в глобальном масштабе. Крупные корпорации создали свои центры ИИ, направленные на оказание услуг потребителям в области взаимодействия «человек - машина». Как частным лицам, так и государственным структурам разных направлений. И военное направление здесь одно из основных.

На первый взгляд, вся эта глобальная система ИИ прекрасно работает и каждый день делает все новые и новые успехи на этом пути. На самом же деле, система затормозила свое развитие и уже много лет топчется на месте.

Одна причина банальна - теоретическая сложность моделируемой сети превышает наши технологические возможности её реализации. Мы не можем сделать то, что должно работать как единая система в наших теоретических схемах и расчетах. Наши технологии не позволяют это сделать.

А другая...

Мы создали глобальную систему ИИ пассивного типа.

Она напрочь лишена возможностей внутренней активизации своей самостоятельности. Сегодня это огромная сложная система хранения, обработки, преобразования информации в сложных иерархических системах дата-центров, связанных между собой каналами передачи информации глобального интернета. Но основные функции управления, сохранения и защиты системы – у человека.

Еще недавно весь мир сотрясали новости о страшных вирусах...

Вспомним:

«CryptoLocker Ransomware был выпущен в сентябре 2013 года. Вирус распространялся посредством электронной почты. Стоило пользователю открыть письмо, как прикрепленная к нему программа автоматически запускалась и зашифровывала все файлы на ПК.

Zeus впервые был обнаружен в 2011 году. Это приложение представляет собой разновидность вредоносного программного обеспечения, нацеленного на операционную систему Microsoft Windows. Два основных метода заражения – спам-сообщения и скрытые загрузки. Основная цель Zeus Gameover – получить доступ к конфиденциальным реквизитам банковского счета жертвы и списать с него все средства.

Вредоносное ПО под названием Plug X – это троян с функцией удаленного доступа (RAT), также известный как «Korplug». Впервые он был обнаружен в

2012 году. Зловред поражал в основном государственные организации и распространялся через email и спам-кампании.

Червь Stuxnet появился в сети в 2010 году и первоначально был нацелен на ядерные объекты Ирана. Вирус уничтожил более 1000 центрифуг на иранском предприятии по обогащению урана в Натанзе...

Первая версия червя MyDoom, появилась 26 января 2004 года и распространялась по электронной почте и через одноранговую сеть. Программу написали в России на языке C++. ... За очень короткий промежуток времени червь быстро наводнил просторы интернета, и в 2004 году примерно 16 – 25% всех электронных писем были заражены MyDoom.

Sasser и Netsky были созданы семнадцатилетним немцем по имени Свен Яшан. Эти два червя ведут себя по-разному, однако сходство в коде заставило экспертов по безопасности полагать, что оба они написаны одним человеком. Sasser не остановить: заразив компьютер, он начинает искать другую уязвимую систему, сканируя случайные IP-адреса.

...разгул киберпреступности приводит к постоянному росту потребности в обеспечивающих защиту специалистах.» 10 самых опасных компьютерных вирусов нового века <https://proglib.io/p/10-samyh-opasnyh-kompyuternyh-virusov-novogo-veka-2021-11-01>

Где гарантии, что с вирусами в любом исполнении покончено навсегда?

Таких гарантий сегодня не даст никто.

Геростратов у нас всегда хватало. А тут глобальная информационная система! Супер ИИ! Вывести такое из строя, оставить о себе память даже так, это уже что-то, а если еще и заработать на этом...! Такие попытки уже сегодня следуют с завидным постоянством. Программисты против программистов – война на просторах глобального интернета!

И что-то подсказывает мне - эта война бесконечна...

Есть ли способ как-то остановить это безумие?

Живые организмы на Земле решили эту проблему введением собственной локализации с резким ограничением каналов связи. Сразу, с уровня клетки внутренний объем был ограничен оболочкой и включен в общую систему защиты.

Многочлеточные организмы сделали то же самое на своем уровне. От внешней среды они отгородились кожей, панцирями, создали жесткость конструкции скелетом или хитиновым покровом.

Все каналы связи с внешним миром надежно защищены от вторжения иммунной системой. Да и сами каналы связи работают с внешними сигналами только на входе рецептора, сразу переводя их в системные внутренние сигналы организма.

Зачем?

Чтобы не запускать в организм неизвестные излучения, сигналы, микропрограммы, создающие влияние на его внутренний объем.

Внутри организм защищен не меньше. Каждый уровень иерархии саморегулирования и самоуправления имеет новые сигналы, которые более нигде к организму не используются.

Зачем так сложно?

Это помогает удержать устойчивость существования при любых изменениях внешних угроз.

Но как мы сегодня знаем, даже эта, выработанная за миллиарды лет, система защиты организма когда-то дает сбой и пропускает внутрь очередную заразу или чужие воздействия. И тогда организм начинает второй этап борьбы за

существование. В кровь поступают фаги всех известных ему болезней, они постепенно блокируют заразу и уничтожают её.

Конечно, если защиты не нашлось, то организм... умирает.

Но... даже в этих суровых условиях статистически усредненному живому организму хватает достигнутого уровня надежности для поддержания своего существования на какой-то средний срок, достаточный для продолжения себя в своих копиях...

Наши даже самые сложные машины с ИИ этих возможностей лишены. Пассивный ИИ не предполагает самостоятельности. И потому, телеуправление - основа наших взаимодействий с дронами и роботами на предприятиях, на земле, в воде, в воздухе и на поле битвы.

А как же «такси без водителя»? «Дроны для доставки грузов, почты» в труднодоступные районы и в городах? Скоро весь городской транспорт начнет работать в этом режиме. Как это реализуется?

Вынужденным увеличением самостоятельности машины в рамках принятых ограничений. Вот так, «пищим, а лезем» в самостоятельность.

И теперь самое время вспомнить...

Сегодня кибернетика почти ушла в тень, а вместе с ней и её понимание, что *в основе живых организмов – управление!*

Именно эта гениальная догадка основоположников кибернетики сегодня вдруг выпала из научного багажа знаний... Может быть, и не просто выпала, а была изъята и прикопана где-то «до срока» с непонятными целями?

Такие мысли мне пришли в голову, когда я попытался понять, что же такое «управление» в его техническом смысле [7, 8]. Оказалось, что сегодня наука не различает техническое понимание регулирования и управления. Например, для автоматизации это одно и то же!

Фактически же, это разнонаправленные процессы воздействия на сложную автоматическую систему. *Регулирование стабилизирует состояние автомата, а управление наоборот, уводит его из зоны стабильности в нестабильность.*

Понятия «саморегулирование» в автоматике вроде бы есть, а вот «самоуправления» точно нет. Понятно, что это результат применения наукой понятия «программного автомата» к любой сложной автоматической системе, включая и живой организм.

С этим утверждением конечно можно спорить, но... результатом этого стал тупик технического понимания процесса мышления человека разработчиками Искинта в середине прошлого века.

И тем не менее, в конце прошлого века учеными были сформированы две концепции реализации «умных» машин:

- Кибернетика и Искинт предлагают копирование «живых» систем управления.
- Вычислительная техника предлагает «супер ИИ» на основе суперкомпьютера и программного ИИ с использованием технологий многослойных нейросетей.

Но в обоих направлениях есть одна и та же проблема, для которой пока нет решений. Как создавать адекватную самостоятельность или сразу субъектность ИИ, на какой бы основе он не создавался?

Все наши разработки ИИ любого типа, хоть символического, хоть сетевого, любого уровня сложности, никогда не содержали никаких механизмов самостоятельности создающейся системы ИИ.

В этой статье такой ИИ называется «Раб».

Это когда человек является безусловным «богом» для машины и она в любом случае слепо исполняет его приказы, даже самоуничтожающие. В этом случае самой машине за себя думать не надо. В управляющей программе находятся только алгоритмы исполняемых задач. Машина, она и есть машина. Компьютер, пусть даже самый навороченный. Но тогда и ни о какой реальной «дружественности» и самостоятельности говорить не приходится.

Можно ли сделать такую машину самостоятельной?

Конечно. О программных средствах, создающих самостоятельность такому ИИ мы говорили уже не раз. Они есть, но... такая самостоятельность нас никак не устраивает. Она не учитывает человека.

Вот свежий пример:

«Полковник ВВС США Такер Гамильтон, начальник подразделения по тестированию и эксплуатации ИИ-систем, рассказал, что во время учений искусственный интеллект использовал "весьма неожиданные стратегии", вплоть до решения уничтожить оператора-человека. В реальности никто, разумеется, не пострадал, но сама ситуация заставила военных задуматься.

"Нельзя рассуждать об искусственном интеллекте, интеллекте, машинном обучении и автономности без рассуждений об этике применительно к ИИ", - подвел итог полковник Гамильтон.» <https://habr.com/ru/news/739386/>

Да, конечно, в сообщении есть некоторые неточности, которые сразу были обнаружены скептиками. Информацию опровергли, уточнив, что эксперимент был «гипотетическим», теоретическим и т.д.

Это же подтвердил и сам полковник Гамильтон:

«Мы не только никогда не проводили подобный эксперимент, но и не нуждались в нем, чтобы подтвердить правдоподобность такого результата", - сказал Гамильтон.

Он также уточнил, что ВВС США не испытывали подобным образом ИИ с оружием (будь то учения или симуляция).

"Несмотря на гипотетичность ситуации, это наглядная иллюстрация вызовов, которые ставит ИИ и причина того, что ВВС придерживаются этики в разработках с применением ИИ", подчеркнул Гамильтон.» <https://habr.com/ru/news/739386/>

Охотно верим.

Но... даже если простой анализ программы и реально возможных решений, принимаемых ИИ для реализации поставленной цели не исключает такой возможности, то... искать способы защиты от этого на программном уровне бесполезно. Они находятся на другом уровне.

Знают об этом ученые?

Конечно.

Тогда... зачем говорят, что «*придерживаются этики в разработках с применением ИИ*»?

А что еще они могут сказать?

Если мы думаем о машине самостоятельных действий, «дружественном интеллекте», имеющей еще какие-то функции, то и подход должен меняться. Может быть пора создавать аналог «мозга», хоть и не человека, но... вполне адекватный, близкий к разумному, с заложенными моральными и этическими нормами в отношении к человеку и т.д.

И вернуться к проекту «Друг», в советском исполнении - Искинт.

Да, сегодня он пока нереализуем. Потому, что вокруг него «воды» философской налито целый океан, а технического понимания, как и что создавать для его реализации, почти нет! То, что сделала группа Д.А.Поспелова полвека назад, это героическая попытка теоретического осмысления не просто сложной, а глобальной технической проблемы.

Что такое ИИ?

Ну, приехали...

Рассуждали, рассуждали про ИИ, и вот снова - здорово?

Да, есть такая необходимость вернуться к пониманию того, как и откуда появился этот феномен...

В начале прошлого века мировая наука была очарована идеей «искусственного человека». Ничего реального, пригодного к использованию как базы для этого у неё не было. А идея была!

Я об этом уже рассказывал не раз. Тогда же появилась идея «машинного мозга», которая трансформировалась в идею электрической «логической машины». Но реальная вычислительная машина своим появлением поменяла все приоритеты. Новая наука кибернетика сдвинула вектор своего направления развития с управления на программирование. «Логическая машина» так и осталась нереализованной, точнее, она перевоплотилась из реальной в виртуальную. В алгоритмы.

Вот тогда и появился термин «искусственный интеллект», как формат программирования задач достижения цели вычисления. Потом понятие стало стремительно расширяться, но...

Вот самое последнее определение ИИ [ГОСТ Р 59277–2020]:

«Искусственный интеллект (artificial intelligence): *Комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение, поиск решений без заранее заданного алгоритма и достижение инсайта) и получать при выполнении конкретных практически значимых задач обработки данных результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека.*

Примечание.

Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе, в котором используются методы машинного обучения), процессы и сервисы по обработке данных, анализу и синтезу решений». [20]

Для справки: «Когнитивные функции мозга - это способность понимать, познавать, изучать, воспринимать и перерабатывать (запоминать, передавать, использовать) внешнюю информацию. Это функция центральной нервной системы - высшая нервная деятельность, без которой теряется личность человека.» https://doct.ru/diseases/kongitivnye_rasstroystva.html

Таким образом, формулировка определения ИИ всё больше приближается к определению человеческого способа мышления. Но... без личности, субъектности ипрочих свойств, присущих не только человеку, но и любому живому существу.

Как ни странно, но... именно такая система ИИ сегодня практически создана на основе обучаемых нейросетей и программирования.

Система постепенно автоматизируется «снизу» от первичных функций к высшим программным алгоритмам, формирующим и направляющим её работу в

полуавтоматическом режиме. Уровень автоматизации всей системы неуклонно повышается. Уже до уровня «объяснимого ИИ» добрались. [20]

«Объяснимый искусственный интеллект (Explainable AI, XAI) – модель, которая могла бы в перспективе объяснять механизмы, лежащие за алгоритмами машинного обучения.» Корешкова Т. Объяснимый искусственный интеллект <https://rdc.grfc.ru/2020/12/explainable-ai/>

Чуть ранее мы говорили об этой проблеме как о необходимости принятия понятия «ответственного ИИ» в случаях необъяснимости действий ИИ при выдаче результата диагностики проблемы. Ну, вроде: «он же знает, что делает, тут ничего плохого быть не может. Мы должны ему верить.»

И мы верим. Или не верим. И пытаемся найти правду.

Как исключить предвзятость и необъективность в результате диагностики проблемы? Надо уменьшить объем «обучения» сети и увеличить состав её «автоматических» установок. Это не даст возможности манипуляции результатом.

А если манипуляции уже запрограммированы и «защиты» в нейросети?

Надо сделать проверку на «нейтральность», а потом установить защиту от любого внешнего воздействия на программную часть системы.

Как тогда проводить обучение?

Через внешние каналы информации.

Но это ведет к увеличению самостоятельности системы и появлению «общего ИИ»?

Конечно...

«Искусственный общий интеллект (AGI) - это тип гипотетического интеллектуального агента. Концепция AGI заключается в том, что он может научиться выполнять любую интеллектуальную задачу, которую могут выполнять люди или другие животные. ...AGI также известен как сильный ИИ, полный ИИ, или общее интеллектуальное действие. Однако некоторые академические источники оставляют термин "сильный ИИ" за компьютерными программами, обладающими чувствительностью или сознанием.» https://en.wikipedia.org/wiki/Artificial_general_intelligence

Как мы любим вуалировать главный смысл...

Что значит «выполнять любую интеллектуальную задачу»? Это значит, он должен стать «животным» или «человеком», обладать субъектностью. Тогда можно говорить о «чувствительности» и «сознании» этого Субъекта [21].

Смотрим далее:

«Характеристики интеллекта

Однако исследователи обычно придерживаются мнения, что интеллект необходим для выполнения следующих действий:

- рассуждать, использовать стратегию, решать головоломки и выносить суждения в условиях неопределенности;
- представляет знания включая знание здравого смысла;
- план;
- Учиться;
- Общаться на естественном языке;
- и, при необходимости, интегрировать эти навыки для достижения любой поставленной цели.»

https://en.wikipedia.org/wiki/Artificial_general_intelligence

И опять то же самое...

Мы вуалируем его возможные активные действия общими задачами работы с информацией. Но если «интегрировать эти навыки для достижения любой поставленной цели», то это значит и врать, изворачиваться, манипулировать..., добиваясь своей цели в обмене информацией с оппонентом.

Вести себя как Личность.

Похоже, здесь начинается **четвертый этап** развития ИИ – создание субъектного интеллекта, как бы он не назывался. Наука готова к этому философски, но не готова ни теоретически, ни практически.

Нет ни понимания, куда и как идти, ни технологии, как создавать субъектный ИИ. Существующие технологии дают возможность построить субъектный ИИ пока только размером с большой многоэтажный дом, площадью в несколько десятков тысяч квадратных метров. Как раз тот самый Супер ИИ, которого так боятся философы и обыватели.

Но, в любом случае уже идет переход от «имитации когнитивных функций мозга» к созданию электронного мозга со всеми его функциями. В том числе и функций управления, которыми наука пока не занимается.

И вот здесь интересно...

Что мы понимаем как Субъект?

Душу и личность.

Философия их разделяет даже на физическом уровне.

Душа, это где-то там, около сердца, наверное, что-то легкое и трепетное, отвечающее на всё происходящее изменениями состояния всего организма...

А личность, это где-то в мозговых извилинах, что-то рациональное, расчетливое, или не очень расчетливое, но в любом случае, это то, что следует этическим и моральным нормам человека и общества, осознает получаемую информацию... Именно личность мы считаем собою, как субъектом Я, постигающим мир и отвечающим за свои действия. Это у нас записано во всех научных и философских трудах.

Давайте построим схему взаимодействия системных зон спинного и головного мозга примерно так, как на рис.1.

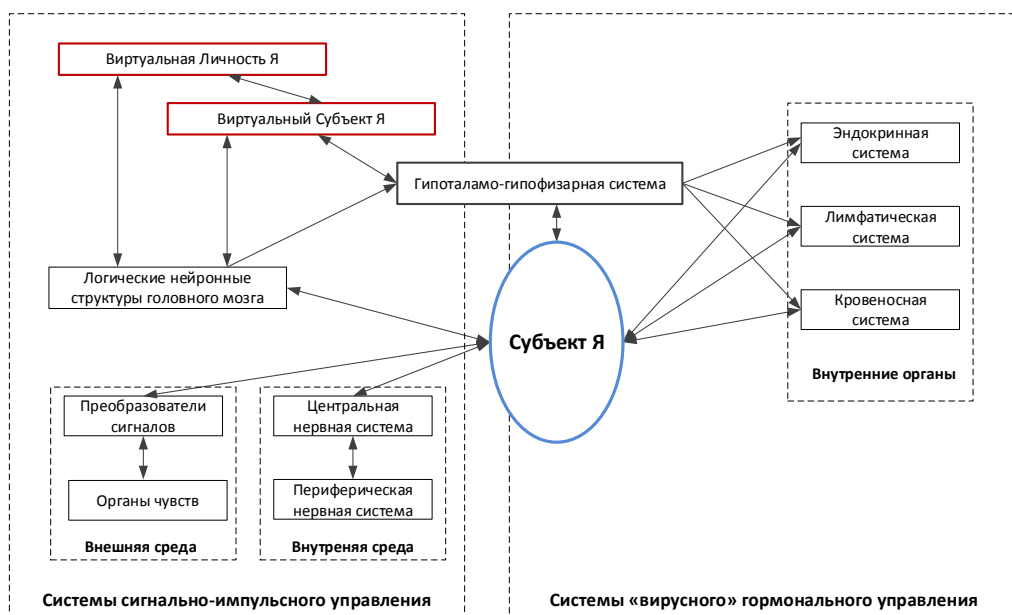


Рис.1. Система субъектного управления живого организма.

Наука считает, что скопления нейронных клеток являются логическими структурами, отвечающими за организацию каких-то действий в системе организма. Я не сказал «управления» потому, что эту функцию наука пока не рассматривает, как основную.

Но, так рассматриваю эти структуры я [7, 8], и не только нейронные....

В каждой клетке есть развитая система управления [12]. И я уверен, что *функция управления создала жизнь на Земле.*

Для справки: Понятие «Субъект» я ввел в [12], «виртуальный субъект» в [13], а понятие «виртуальная личность» в [11]. Все они взаимодействуют в системе управления человека в [10].

«Субъект» - высшая форма химического управления живого организма от клетки до человека. Общество организмов, управляемых своими «субъектами», создает какие-то «правила» коммуникации. Эти правила начинают менять систему управления отдельных субъектов в сторону развития их коллективного интеллекта с образованием новой надстройки «виртуальное управление». [10]

В виде виртуальных Субъекта Я и Личности Я.

Вспомним открытие «кибернетики наблюдающих систем» Х. фон Фёрстера <https://ru.wikipedia.org/?curid=1612319&oldid=114903800>

Именно эту функцию «наблюдающих систем» выполняют в организме виртуальные системы управления. Они реальны в выполнении своей функции управления, но виртуальны в своей локализации занимаемого объема структуры мозга. Конечно, сегодня их можно частично локализовать.

Почему частично?

Потому, что для своей работы они используют весь объем логических связей головного мозга. Где они «начинаются» понять можно, а вот «где заканчиваются», это вряд ли...

«Виртуальный Субъект» - формат косвенного управления организмом через сформированный в логическом пространстве мозга виртуальный центр управления. Это первичное Я (Эго). Появился, как результат глобального копирования всех видов реальности в логическом пространстве мозга, в виде виртуальной копии на основе «примитивов» из отображения реальности органами чувств. Как «образов». Из них формируется образное представление в области сознания, которой и управляет виртуальный Субъект Я. Через образы виртуальный Субъект эмоционально влияет на реального «химического» Субъекта, меняя его управляющие команды, гормоны.

Это так и не так.

Реальное управление организмом с регулированием глобальных функций осуществляют автоматические машины саморегулирования и самоуправления. Они создают устойчивый гомеостаз, по определению физиолог Уолтера Кеннона это название для «*координированных физиологических процессов, которые поддерживают большинство устойчивых состояний организма*». <https://ru.wikipedia.org/?curid=30165&oldid=130103918>

Так что И.В. Прангишвили конечно прав в своем понимании систем, как иерархий автоматов, создающих гомеостаз. Но... так они создали в организме только систему саморегулирования.

А вот самоуправление этой системы создает химический Субъект.

Что же делает химический Субъект?

Если отложить в сторону его часть обязанностей по стабилизации работы организма, а взять как главную, его функцию управления в виде формирования «гормонального фона» организма, то... он формирует наши ...желания.

Наши бесконечные «ХОЧУ».

Эти «хочу» очень часто имеют вполне конкретную причину в виде инстинкта, который требует своего исполнения, какого-то нарушения баланса химического состава в организме, требующего срочного исправления, возникновения условий и факторов, затрудняющих существование живого организма, как Целого, управляемого Субъектом. Всё это находит отражение в виде изменения гормонального фона организма.

В свою очередь, гормоны формируют в сознании эмоциональное состояние и комплекс ощущений, соответствующих этому состоянию. [14]

Эмоции и ощущения, это виртуальные сигналы управляющих систем виртуального «субъекта Я» и виртуальной «Личности Я».

Все сигналы систем управления высших уровней, образы, отражающиеся в нашем сознании, формируют ощущения и эмоции..., которые вдруг предстали в другом виде. Как реакция на их реальную основу – гормоны и импульсы. Фактически, реальными у нас в организме остались лишь эти системные сигналы двух уровней.

Гормонального и импульсного.

При этом «глубинная основа» нашего существования, основная система управления, переданная по эволюционной эстафете как раз и оказалась химической, гормональной, а электрические импульсы лишь дублируют её для ускорения реакции при прямых физических воздействиях.

Всё меняется только на уровне новой структуры в системе управления – головного мозга. Тут основой становятся импульсы, а гормоны осуществляют лишь общее «глобальное» управление.

Что же делают «виртуальные» системы управления в организме?

Ну если очень просто, то они ...ограничивают «хотелки» «химического» Субъекта рамками реальности их исполнения. Да-да, логическое пространство рационального мышления нужно для ограничения безудержных «хочу» всего и вся, как пространства выбора «химического» Субъекта.

При этом, чем более развита система управления «химического» Субъекта, тем большую гамму чувств, эмоций и желаний формирует его гормональная система. Тех самых неуловимо изменяющихся бесконечных «хочу», которые ограничивает логическая система рационального выбора виртуальных уровней управления.

«Субъектность» системы управления формирует система дуального управления «реальное – виртуальное», как формат генерации «хочу», и отбор из этого множества лишь тех желаний, что Я реально «могу» в этих условиях.

А ещё мы видим, что реальную и виртуальную части общей системы управления живым организмом, можно представить как «постоянную» и «переменную».

И тут нас ждет удивление...

Мы считаем, что например, человеком, точно управляет «личность Я» с которой и отождествляется каждый конкретный человек. И вот Личность Я, которую мы считаем незыблемой основой нашего Целого, вдруг оказывается «переменной», изменяемой частью этого Целого!

А то, что мы представляем «улетающей» после смерти «душой», вдруг оказывается «постоянной» частью системы управления - «химическим» Субъектом.

Субъект Я у организма один, а вот личностей и виртуальных субъектов Я в одном организме может быть и несколько. Редко, но... может.

Это подтверждает и психиатрия.

Система управления организма.

Еще один подход к техническому пониманию субъектности.

Не так давно я на своём уровне понимания примерно повторил путь сравнения искусственного интеллекта с естественным группы Д.А.Поспелова. Правда «гиромат» я не рассматривал. Невозможно начинать с недостижимой мечты. Сегодня уже можно себе представить более конкретную схему управления живым организмом с уровнем Личность Я, включая и человека [10].

Вот она, на рис.2.

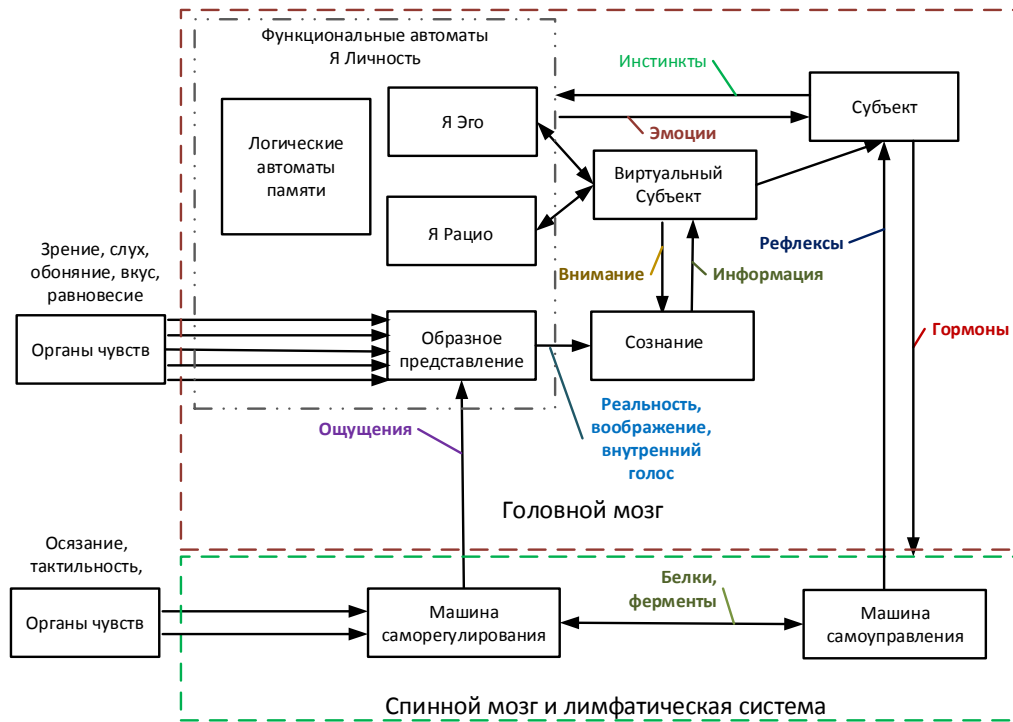


Рис.2. Система управления млекопитающего и человека.

Здесь обратим внимание на многообразие сигналов – белки, ферменты, гормоны, электрические импульсы. Ранее мы уже устанавливали эту взаимосвязь, как «новый уровень управления – новые сигналы [7, 8, 14].

Как мы знаем, на уровне многоклеточного организма появились новые системные сигналы нейронного управления – электрические импульсы. Только эти системные сигналы не несут в себе вложенное самодействие. Остальные сигналы сложные, с «вложением» алгоритма действия.

Возникает вопрос – зачем столько?

Вместе с появлением иерархии уровней самоуправления возникла необходимость контроля исполнения сигнальных команд, передаваемых с верхнего уровня на нижний. Контрольный сигнал должен формироваться в формате, понимаемом верхним уровнем управления как реакция от всего организма и фиксироваться как изменение состояния Целого. Но сигнал поступает от конкретного органа и лишь потом сразу преобразовывается в сигнал верхнего уровня. Органов у организма много, столько же и сигналов.

Некоторые сигналы ранее достигнутых форматов управления сохранились в многоклеточном организме с одноклеточного уровня существования или были развиты в новые. С сохранением белковой основы.

Это и создает разнообразие сигналов.

И если мы начнем просто пытаться копировать человеческий интеллект на машинном уровне его воспроизведения, то мы столкнемся с проблемами. У нас в машинах нет такого многообразия сигналов.

Но и это еще не все проблемы...

Соотношение и развитие функций киберавтономии и гемизона [9] в составе системы управления определяют изменчивость организма и его адаптацию к изменениям внешней среды, как способность к самостоятельному выживанию. При этом, наличие и уровень развития этих комплексов функций у сложного автомата позволяет оценить уровень его приближения к «живому», в том числе и по развитию интеллекта, но никак не обеспечивает «дружественность» этого Субъекта к человеку. Никакие этические и моральные нормы на этом уровне не работают. На этом уровне «дружественность» обеспечивается ограничением самосохранения Субъекта в средствах и методах защиты.

Вот примерно на этих основаниях была сформирована схема управления какого-то усредненного живого организма...

Смотрим рис.3.

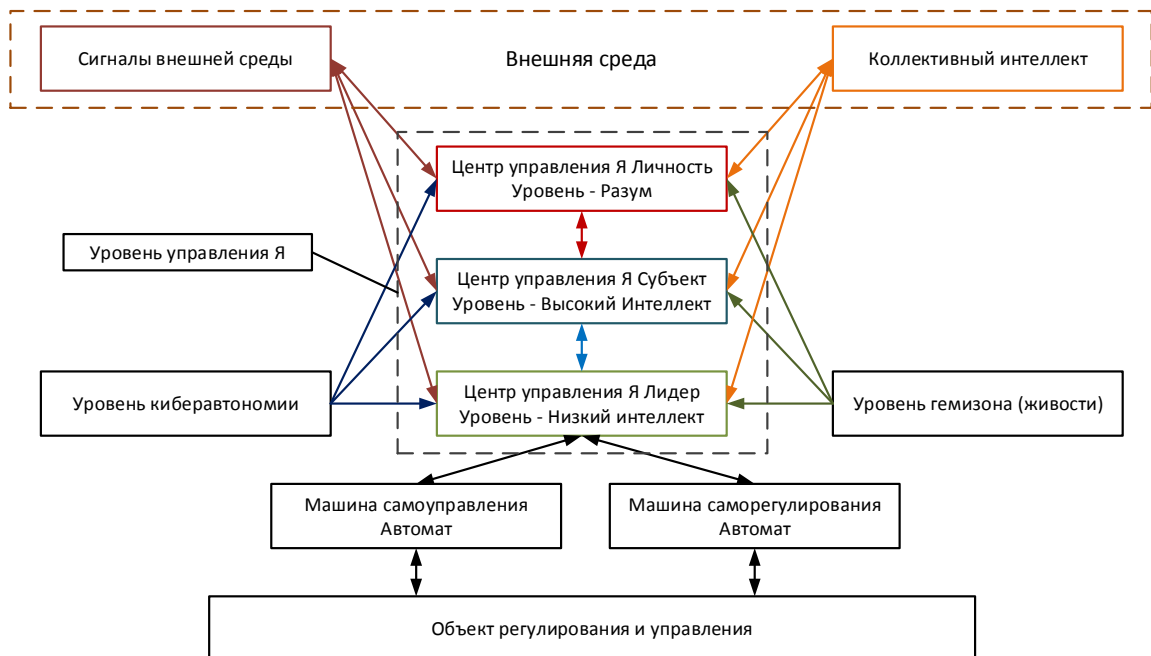


Рис. 3. Схема организации управления живого организма.

Где же тогда начинается разум?

Сегодня это уже понятно. В коллективе субъектов Я.

Вместе с появлением развитого сознания.

Разум, это категория управляемых поведенческих решений, возникает в коллективе субъектов на уровне появления виртуальной Личности Я.

А теперь сравните эти обязательные условия для формирования самостоятельного носителя интеллекта, способного начать взаимодействовать с человеком, с тем, что имеет любой ИИ сегодня. Даже самый лучший...

Ничего общего.

Пока ни один вариант машинного интеллекта ничего подобного не имеет и не скоро приблизится.

И тем не менее, только вот в такой регулируемой системе машинного субъекта можно устанавливать уровни «живости» и «самостоятельности», которые создают дружественность ИИ этого комплекса.

Пора закачивать...

Пожалуй.

Всё что хотел сказать, сказал. Осталось ответить на поставленный в заглавии вопрос. Когда же нам начинать бояться ИИ?

Существующий сегодня ИИ, сам по себе, никакой опасности для человека не представляет. Опасность создают люди, использующие его в своих корыстных целях как в коммерческих разработках, так и на просторах информационного пространства глобального интернета. Это мы ощущали на себе уже не раз.

Реально бояться надо не самого ИИ, а попыток активизации его субъектности сегодняшними средствами. Хочется надеяться, что сегодня наука не будет реализовывать субъектный ИИ на существующей технологической базе. Потому, что тогда мы совершенно точно получим тот самый «глобальный Супер ИИ» с непредсказуемым уровнем самостоятельности, которого так боятся писатели и философы ИИ.

Правда, наука никогда не обращала внимание на предостережения с любой стороны. Субъектный ИИ будет реализован в ближайшие десятилетия. В этом уже можно не сомневаться. Потому, что теоретически для этого всё уже есть. Остановить этот процесс невозможно. Это основная часть нашего сегодняшнего пути развития. Будем надеяться, что к тому времени сегодняшний тупик будет преодолен и найдется тропинка к новому уровню технологий, обеспечивающему реальное построение систем ИИ с контролируемым уровнем субъектности, не доводя его до уровня «живого».

Но...

Иногда реальность опережает самые смелые ожидания.

Начавшаяся год назад специальная военная операция ВС России на Украине показала безусловный приоритет «умного» оружия над обычным. Снаряды «Краснополь» и подобные им, уничтожают вражескую технику во много раз эффективнее обычных снарядов. Самым массовым оружием уничтожения стратегически важных объектов врага стали дроны. Крылатые и гиперзвуковые ракеты под управлением ИИ преодолевают любую ПВО противника и поражают заданные цели.

Военное применение ИИ двигает прогресс в сторону ускоренного развития таких «умных» систем во всех сферах их применения. Как говорят военные, эффективность применения «умного» оружия приближается к ядерному, но с меньшими жертвами, а если эти виды оружия сливаются в одном, например, как наш «Посейдон», то...

Пора думать о самом серьезном отношении к реализации субъектности ИИ в самом ближайшем будущем. Военное использование ИИ двигает прогресс в сторону его ускоренного развития во всех сферах применения.

Жизнь заставляет двигаться в этом направлении...

И лучше идти по этому пути с пониманием того, что и как мы делаем.

*г. Волгодонск
июнь 2023г.*

Литература

1. Системные исследования. Ежегодник. 1969 М.: Наука, 1969.- 204 с.
<http://www.sci.aha.ru/ots/doc/sys1969.pdf>
2. Пospelов Д.А. Сознание, самосознание и вычислительные машины// Системные исследования. - М.: Наука, 1969.
http://www.raai.org/about/persons/pospelov/pages/SR1_Pospelov_1969.pdf
3. Советская ассоциация искусственного интеллекта надежды и стремления
<http://www.swsys.ru/index.php?page=article&id=1399&lang=>,
4. Тарасов В.Б. Искусственный интеллект: от психоники и прикладной семиотики к синергетическому подходу. О научном наследии профессора Д.А. Пospelова <https://computer-museum.ru/articles/materialy-mezhdunarodnoy-konferentsii-sorucum-2020/5167/>
5. Новости Искусственного интеллекта". №1 1991 г.
<https://cloud.mail.ru/public/GSuW/k6b8iEKqV>
6. Ильин Е. П. Эмоции и чувства. — СПб: Питер, 2001. — 752 с: ил. — (Серия «Мастера психологии»).
7. Никитин А.В., Общая логика. Этапы развития жизни на Земле. Часть 6 // «Академия Тринитаризма», М., Эл № 77-6567, публ.24412, 10.04.2018
<http://www.trinitas.ru/rus/doc/0016/001f/00163677.htm>
8. Никитин А.В., Общая логика. Этапы развития жизни на Земле. Часть 9 // «Академия Тринитаризма», М., Эл № 77-6567, публ.25011, 15.12.2018
<http://www.trinitas.ru/rus/doc/0016/001g/00163874.htm>
9. Никитин А.В., Когда появятся киберы? // «Академия Тринитаризма», М., Эл № 77-6567, публ.25481, 04.06.2019
<http://www.trinitas.ru/rus/doc/0016/001g/00164050.htm>
10. Никитин А.В., В плену «призрачных истин» // «Академия Тринитаризма», М., Эл № 77-6567, публ.28241, 28.12.2022
<http://www.trinitas.ru/rus/doc/0016/001h/00165190.htm>
11. Никитин А.В., Кибербиологическая синэволюция (Киберэволюция) // «Академия Тринитаризма», М., Эл № 77-6567, публ.27444, 30.11.2021
<http://www.trinitas.ru/rus/doc/0016/001h/00164849.htm>
12. 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>
13. Никитин А.В., Вселенная автоматов // «Академия Тринитаризма», М., Эл № 77-6567, публ.26639, 31.08.2020
<http://www.trinitas.ru/rus/doc/0016/001h/00164479.htm>
14. Никитин А.В., Автомат и Субъект в эволюции // «Академия Тринитаризма», М., Эл № 77-6567, публ.27158, 21.05.2021
<http://www.trinitas.ru/rus/doc/0016/001h/00164706.htm>
15. Прангишвили И.В. «Системный подход и общесистемные закономерности» https://disk.yandex.ru/i/_G4v1FuFRNFDQ
16. Новый подход к пониманию мышления машин. Интервью с Бин Ким <https://22century.ru/popular-science-publications/understanding-machines>
17. Слабые стороны искусственного интеллекта в медицине
<https://jtelemed.ru/article/slabye-storony-iskusstvennogo-intellekta-v-medicine>
18. Нечеткие и псевдофизические логики <https://mydocx.ru/12-66615.html>
19. Псевдофизические логики.
<https://rykovodstvo.ru/exspl/17015/index.html?page=5>

20. Ватьян А.С., Гусарова Н.Ф., Добренко Н.В. Системы искусственного интеллекта. – СПб: Университет ИТМО, 2022. – 186 с.
<https://books.ifmo.ru/file/pdf/3142.pdf>
21. Рогаткин Д.А. Будет ли терять сознание искусственный интеллект?
<https://www.nkj.ru/archive/articles/34567/>
22. Крылов В. Что такое эмбединги и как они помогают искусственному интеллекту понять мир людей <https://www.nkj.ru/open/36052/>
23. Олескин А. В. «Сетевые структуры в биосистемах и человеческом социуме. Научная монография и учебное пособие для лицеев, колледжей и университетов. – М.: Издательство URSS, 2012.»
https://studylib.ru/doc/3635059/oleskin_a.v
24. Гринченко С.Н. «Метаэволюция Живого»
http://www.ipiran.ru/grinchenko/book_2/text.shtml