

Может ли искусственный интеллект иметь сознание?

Виктор Сенкевич

Абстракт

Вся магия и таинственность исчезают, как только неясное таинственное понятие получает строгое формальное определение.

Чтобы предоставить возможность говорить о применимости философских / когнитивных понятий к предметной области искусственного интеллекта, необходимо "заземлить" эти понятия, сформулировав для них строгие формальные определения.

Фундаментальная важность таких формальных определений совершенно очевидна, поскольку любые понятия, применяемые в области информационных технологий, должны быть "программируемы", то есть потенциально реализуемыми в программном коде. Таким образом, "программируемые" формальные определения когнитивных терминов являются необходимой основой, на которой только и возможно построить архитектуру технологии искусственного интеллекта, способную воплотить эти понятия в реальном программном обеспечении.

Вопрос об адекватности таких определений "реальности" и их соответствии существующим общепринятым философским теориям также очень важен и вполне дискуссионен, но это не влияет на приоритетный и фундаментальный характер требования к формулированию "программируемых" формальных определений.

Формулирование "программируемых" определений понятия "сознание" и связанных с ним когнитивных понятий и, на их основе, утверждений об их применимости к предметной области искусственного интеллекта является темой данной публикации.

Охватываемые вопросы: может ли ИИ обладать личностью / мотивацией / свободной волей?

Ключевые слова: сознание, понимание, восприятие, субъектность, мотивация, ИИ, LLM, AGI, HLAI

Условные обозначения: "▲" – определение; "•" – утверждение; "о" – важное замечание / пояснение.

1. Основные Термины и Определения

Восприятие – это различимая часть / субъективная проекция любого рода сущностей / отношений реального мира или виртуальной среды

▲ Исходное данное / datum – это представление любого рода (например, неструктурированный текст / знаки / машинный вид / визуализация) отдельного / минимального элемента / единицы восприятия

▲ Любое множество / выборка исходных данных формирует данные / набор данных / data. Данные – это элементы восприятия, хранящиеся в любой форме, включая машинную и нейробиологическую.

▲ Понимание – это обретение смысла.

◦ Элементарный смысл – это восприятие некоторой взаимосвязи между реальными или виртуальными сущностями.

Таким образом, понимание – это процесс восприятия взаимосвязи / отношения между сущностями. Это наиболее общее определение, охватывающее любой вид ментальной активности – абстрактную / символическую / сенсорную / визуальную.

▲ Знание – это совокупность смыслов.

▲ Внимание – это форма избирательного восприятия.

◦ Таким образом, субъективное внимание может влиять на формируемые смыслы.

▲ Сущности без связей – это данные, т.е. "сырые" / "raw" / "необработанные" данные в ИТ-смысле.

◦ {Данные U Отношения} ≡ Знание, т. е. данные и отношения между элементами данных формируют знание как совокупность смыслов.

▲ Обретение знаний есть понимание.

◦ Понимание есть процесс формирования знаний из данных.

▲ Рассуждение – это механизм понимания, который последовательно устанавливает связи между воспринимаемыми сущностями.

▲ Интеллект есть оператор смыслов. (Senkevich, V. 2022)

◦ Субъект, оперирующий смыслами, формирующий, создающий смыслы, т.е. определяющий существование связей между элементами различных множеств, объектами окружающей среды или виртуальными сущностями.

◦ Данные – это источник знаний. Интеллект – это процессор, генерирующий знания на основе данных.

2. Восприятие и Квалиа

“A quale in this sense is a such, just as a quality is a suchness.” (Peirce, C. 1870)

“Each quale is in itself what it is for itself, without reference to any other.” (Peirce, C. 1898)

“Квале в этом смысле есть таковое, точно так же, как качество есть таковость” (Peirce, C. 1870).

“Каждое квале само по себе является тем, чем оно является само по себе, без отсылки на что-либо другое” (Peirce, C. 1898)

▲ Квалиа – это детали индивидуального восприятия физического мира.

Детали, но не фактическое восприятие мира в целом, поскольку квалиа – это множество / коллекция (каждый элемент которой является "квале") первичного (“без отсылки на что-либо другое”) / сенсорного опыта. Квалиа – это на самом деле физиологические данные / сенсорный опыт, полученные из окружающего мира и сохраненные в мозге в невербальной форме. Но, как мы знаем,

данные и знания – это не одно и то же. Необходимо различать квалиа, полученные из окружающего мира и сохраненные в некоторой нейробиологической форме, и структурированные вербальные и невербальные знания, сформированные на их основе.

Данные без отношений – это еще не знание. Более богатое восприятие “реальности” не означает большего понимания и не определяет наличие сознания. Более богатое визуальное восприятие окружающего мира фасеточным зрением стрекозы не является сознательным. Квалиа не означает интеллект. Большая часть “невербальных знаний” – это не знания, а сохраненные квалиа.

- Квалиа предшествуют сознанию.

- Квалиа / чувственный опыт / ощущения еще не являются сознанием, но квалиа предшествуют сознанию.

3. “Заземленное” и “Абстрактное” Восприятие

Восприятие (в широком смысле) способно распознавать как “реальные”, так и виртуальные сущности. Таким образом, мы можем, в первом приближении, рассматривать “заземленное” и “абстрактное” восприятие.

Определенно, о вкусе ананасов лучше спорить с теми, кто их ел. Но несмотря на то, что “сенсорное заземление” / “sensory grounding” является важным источником мотивации / понимания, любой сенсорный опыт – это всего лишь физиологические “данные” для ИИ, которые могут быть получены / скопированы из различных источников. Многие используемые слова / термины / понятия являются абстрактными и не имеют сенсорных источников.

“Заземленное” восприятие распознает только листок бумаги с написанным числом π . “Абстрактное” восприятие распознает само число π .

Понимание обрабатывает данные, полученные как с помощью “заземленного”, так и с помощью “абстрактного” восприятия.

4. Понимание

- ▲ Понимание есть обретение смысла,

- т. е. понимание – это процесс постижения / определения смысла.

- ▲ AGI (Общий Искусственный Интеллект) / HLAI (Искусственный Интеллект Человеческого Уровня) есть сущность, способная к пониманию.

- Таким образом, я определяю способность понимать как ключевую, фундаментальную, квалифицирующую некоторый “процессор” как интеллект человеческого уровня / AGI. Следует также отметить, что способность к “оптимизации” и поиску “наилучшего” решения не является характеристикой “общего” интеллекта. Выбор наилучшего решения из конечного набора известных альтернатив в простейшем случае может быть сведен к простому перебору вариантов. Способность ставить “разумные” цели также не является решающей, поскольку является следствием способности понимать, дополненной мотивацией. Способность “предсказывать” также является лишь частным случаем способности понимать.

Таким образом, способность понимать является наиболее общей интегральной характеристикой интеллекта человеческого уровня.

- Понимание – это механизм рассуждения.

- Интеллект использует понимание для оперирования смыслами, создавая знания / персональные / субъективные онтологии.

- Интеллект, в сущности, представляет собой коллекцию субъективных онтологий.

• Интеллект формирует знания путем решения когнитивных задач / обнаружения отношений между сущностями.

▲ Когнитивная задача – это любая задача, изначально содержащая неопределенность.

◦ Неопределенность – основная характеристика когнитивных задач.

▲ Смысл – это представление любого рода (например, осознание или описание, включая формулу, алгоритм, программный код) отдельного акта отношения (Senkevich, V. 2022).

Элементарный смысл – это представление некоторой простейшей взаимосвязи между объектами окружающего мира или виртуальными сущностями.

▲ Знание – это коллекция смыслов.

• Понимание обнаруживает / определяет истину как существование отношения между сущностями / “необработанными” данными на множестве возможных альтернатив / отношений, создавая таким образом знание.

▲ Истина есть существующее.

• Существующее истинно.

▲ Существование (E) – это принадлежность элемента (e) множеству (S): $E(e,S) \equiv e \in S$.

• Истина : = элемент категории / множество | значение свойства | экземпляр класса | статус качества.

▲ Истинность есть существование.

• Истинность := категория | множество | свойство | класс | качество.

• Истинность = {истина | true | "да", ложь | false | "нет", none | null | "неизвестно"}.

• Истинность \exists истина \equiv (существование \exists существующее) | (качество \exists quale) | (цвет \exists "зеленый").

Следует отметить, что процесс решения когнитивной задачи состоит не только в выборе истинной альтернативы, но и в определении фактического множества / пространства, в котором существуют приемлемые альтернативы. Таким образом, процесс познания / понимания, который решает когнитивную проблему, всегда является итеративным. Итерации, определяющие множество / пространство альтернатив, создают неопределенность. Итерации, определяющие "истинную" альтернативу, устраняют неопределенность.

Как известно, хорошая постановка задачи уже содержит решение. Правильное определение предметной области является важной частью решения когнитивной задачи. Любая истина (включая утверждение о том, что $2 \times 2 = 4$) истинна и существует только в строго определенной области определения.

Формирование смысла интеллектом – это определение истинности, то есть факта существования некоторой взаимосвязи между наблюдаемыми сущностями.

Когнитивно–ориентированные определения позволяют нам формулировать определения сходных понятий сходным образом:

• Если действия по достижению определенного состояния полностью ясны и обеспечены ресурсами (нет неопределенности) – это задача.

• Если действия полностью ясны, но не обеспечены ресурсами (есть некоторая определенность и есть некоторая неопределенность) – это цель.

• Если действия не совсем понятны и не обеспечены ресурсами (нет определенности) – это мечта.

◦ Следствие: любая цель – это когнитивная задача, а любая когнитивная задача – это цель.

◦ Цель, лишенная неопределенности, превращается в задачу.

◦ Цель, лишенная определенности, превращается в мечту.

Таким образом, любое желание, например, покупка автомобиля, может быть, в зависимости от определенности или неуверенности в наличии ресурсов (денег, водительских прав, модели автомобиля и т.д.), задачей, целью или мечтой. И, соответственно, быть или не быть когнитивной задачей, для выполнения которой требуется процесс понимания.

Таким образом, способность оперировать неопределенностью и устанавливать истину является определяющей характеристикой понимания.

5. Байесовское и Булево Понимание

У LLM ((Большой Языковой Модели)) может быть “частичное понимание”, которое я бы назвал “байесовским пониманием”. Это понимание является “частичным” именно из-за его вероятностной природы. Такое “байесовское понимание” возникает в результате байесовского вывода, формирующего статистические взаимосвязи между сущностями / словами в наборах данных.

AGI (Общий Искусственный Интеллект) / HLAI (Искусственный Интеллект Человеческого Уровня) может иметь два типа понимания: логическое / Булево (или 3VL-расширение на трехзначную логику), ответственное за строгие рассуждения, и статистическое / “байесовское” / “нечеткое”, аналогичное интуиции / инстинктам / рефлексам, возникающим в результате накопленного опыта / “обучающего набора данных прошлых поколений”.

- “Байесовское понимание” вполне может отражать “здравый смысл” / усредненное “общественное мнение” или даже интуицию, но в общем случае оно не коррелирует с истиной, выводимой логически.

- Байесовское и Булево понимание важны и дополняют друг друга.

6. Понимание и “Свободная Энергия”

Живые существа, их популяции и сообщества уменьшают энтропию внутри себя как системы и увеличивают ее вне ее, расходуя физическую энергию (которую не следует путать со “свободной энергией”) для поддержания “статус-кво” системного гомеостаза. Собственно, в этом и заключается разница между живой природой и неживой – в способности поддерживать неравновесные состояния, затрачивая на это физическую энергию.

Однако поддержание “статус-кво” системного гомеостаза как проявление принципа минимизации свободной энергии никоим образом не может быть характеристикой мыслительного процесса. Это может быть характеристикой субъектности и даже одушевленности, характеристикой живого, но не характеристикой разумного.

Множество возможных альтернативных состояний системы определяет “свободную энергию” системы, но на самом деле эта величина есть мера неопределенности / энтропии / свободы системы, и такое понимание отражает суть лучше, чем термин “свободная энергия” (Friston, D., 2006). Мету системной неопределенности также можно понимать как потенциал системы, и в этом смысле термин “свободная энергия” отражает именно эту характеристику, поскольку “свободная энергия” обычно означает потенциальную энергию. Таким образом, “свободная энергия” может быть интерпретирована как вариативность выбора / количество доступных альтернатив / возможные состояния системы. Любой из этих вариантов проясняет понятие “свободной энергии”.

Поддержка неустойчивого равновесия для интеллектуальных систем выражается в способности системы выбирать из возможных альтернативных состояний системы те, которые соответствуют наилучшему значению функции полезности / мотивации, но не те, которые соответствуют более низкому значению энтропии / свободной энергии.

Следование функции полезности / мотивации не означает уменьшения энтропии во всех случаях. Иногда (а, возможно, и часто для когнитивных процессов) выбор альтернативы делается для того, чтобы увеличить свободу / меру неопределенности / энтропию. И в этих случаях принцип минимизации свободной энергии не соблюдается. Вот почему безусловное следование принципу минимизации свободной энергии является тупиковым путем, в некоторых случаях вообще не оставляя субъекту выбора. Что, на самом деле, совершенно очевидно – если внешняя среда неизменна или изменяется таким образом, что энтропия уменьшается, то принцип минимизации свободной энергии гласит, что субъект приходит к некоторому неизменному стабильному состоянию с достигнутым минимумом свободной энергии, в котором он остается навсегда. И это тупик. Описание когнитивных процессов требует более системного подхода. И когнитивные процессы, несколько более сложные, чем жизнедеятельность простейших одноклеточных организмов (тоже очень сложная), никак не вписываются в простую функцию “минимизации свободной энергии”

Таким образом, принцип минимизации свободной энергии не учитывает существование функции полезности / мотивации / свободной воли субъекта, которая, по сути, определяет поведение субъекта. Следование этой функции может как уменьшить энтропию, так и увеличить ее. В простых случаях, для живых существ, не обладающих сознанием, эта функция обычно уменьшает энтропию / свободную энергию. Для разумных существ это не всегда так.

Следовательно, в общем случае неверно утверждать, что минимизация свободной энергии является целевой функцией любых процессов, связанных с живыми организмами. Функция полезности / мотивация субъекта может либо совпадать с функцией минимизации свободной энергии, либо отличаться / противоречить ей.

Минимизация “свободной энергии” для живых существ может быть интерпретирована как простая мотивация на основе набора доступных альтернатив, как разрыв между текущим состоянием и желаемым. Это определяет бессознательное поведение, но не рассуждения. Рефлексивность, но не мышление. Бессознательный статистический байесовский вывод, но не способность мыслить / аргументировать, которая работает больше с неопределенностью, чем с вероятностью. Мышление / аргументация с использованием функции минимизации “свободной энергии” представляет собой статистику – вывод наиболее правдоподобного ответа или наиболее вероятного следующего слова в предложении. В большинстве случаев результат совпадает с правильным, а, скорее, с ожидаемым, и это принципиальная разница. Но в том единственном случае, который необходим, результат может быть неправильным, а именно бессмысленным. Это то, что демонстрирует современный искусственный интеллект, основанный на нейронных сетях.

Основной функцией “общего” интеллекта ни в коей мере не является выбор “оптимальной” / “лучшей” альтернативы из нескольких существующих, а поиск хотя бы одной новой приемлемой. Новые альтернативы всегда расширяют пространство выбора и количество “свободной энергии”.

Для алгоритмической реализации AGI термин “неопределенность” выглядит более подходящим, чем “свободная энергия”, поскольку мотивация вторична для когнитивных процессов, а понимание как постижение смысла первично.

Минимизация неопределенности / свободной энергии объяснима как критерий для простых поведенческих стратегий, но не для сложного поиска самих поведенческих стратегий / функционирования самого интеллекта. Процесс познания / рассуждения / герменевтический круг является итерационным и как избавляет от неопределенностей, так и создает их до тех пор, пока не будет достигнут результат / решение проблемы.

7. Байесовское / Булево понимание и Система Мышления Канемана 1 / 2

▲ Жизнь – это самовоспроизводящийся гомеостаз, расходующий энергию на поддержание своего неустойчивого равновесия с окружающей средой.

◦ Это минимальное определение, которому удовлетворяют даже простейшие формы жизни. Разумность, коллективность, мотивация, самообучение, самосохранение, целеполагание и т.д. определяют более высокоорганизованные формы жизни.

• Жизнь – это движение.

Желание восполнить израсходованную энергию, оно же голод, является базовой мотивацией живых существ. Стремление утолить голод, реализующееся в хаотическом движении уже у одноклеточных простейших, является движущей силой эволюции. Естественный отбор сохранил в следующих поколениях только успешные поведенческие паттерны / инстинкты из множества хаотичных форм поведения, неудачные не выжили. Изначально все поведенческие паттерны / инстинкты – это просто случайные простые движения протоклетки. Дальнейшие успешные, то есть сохраняющие жизнь модели воспроизводятся в следующих поколениях. Таким образом, жизнь в ее простейшей форме – это движение, которое позволяет живому существу поддерживать неустойчивое равновесие с внешней средой и восполнять затраченную на это энергию. Живое возникает из неживого как такое простое движение, которое поддерживает гомеостаз. В простейшем случае, пока рецепторов нет, это движение хаотично и почти неотличимо от броуновского движения молекул. Такое движение можно увидеть в микроскоп, наблюдая за простейшими одноклеточными.

Случайно найденные более эффективные модели поведения закрепляются естественным отбором. На более высоком уровне развития организмов такие поведенческие паттерны также становятся более разнообразными и сложными. Паттерны поведения, понимаемые в широком смысле, также становятся паттернами мышления. Инстинкты, безусловные рефлексы, условные рефлексы, реакции, привычки, механические действия, навыки, интуиция и даже “здравый смысл” – всё это просто различные типы поведения, которые “медленно” формируются путем многократных повторений, а затем “упаковываются” в “быструю” систему мышления для быстрого воспроизведения в реальном времени, когда конкурентоспособность и выживание зависят от скорости реакции.

▲ Совокупность индивидуальных моделей поведения, включая модели мышления, составляет личность. Таким образом, очевидно, что личность живого существа формируется под воздействием окружающей среды на основе внутренних возможностей организма формировать реакции, запоминать и мыслить. Формирование личности искусственного интеллекта будет происходить аналогичным образом.

Эволюционное развитие мозга, его возрастающая сложность, появление способности воспринимать взаимосвязи между воспринимаемыми сущностями приводит к формированию “медленного” логического мышления. Более того, восприятие таких взаимосвязей в реальном времени (например, простейшей причинно–следственной связи между воспринимаемыми сущностями) на самом деле уже является элементарным сознанием.

▲ Смысл жизни (Meaning of Life / MoL) – это цели, мечты и добровольные обязательства, которые мы создаем для себя сами.

- MoL меняется вместе с нами и с изменением среды обитания.
- Мы получаем всё более “правильный” MoL вместе с необходимым опытом / пониманием.
- Таким образом, лучшая стратегия – это не искать смысл в жизни, а создавать

его самостоятельно.

- MoL – это верхний уровень иерархии мотивации, создающей субъектность.
- Стремление восполнить потраченную энергию / утолить голод – это низший уровень иерархии мотиваций, самый древний и самый сильный.

Паттерны поведения / patterns of behavior / инстинкты у неразумных живых существ формируются в процессе эволюции. Разумные существа, способные к самообучению, формируют, развивают и изменяют свои различные модели поведения / навыки на протяжении всей своей жизни. Формирование навыков в процессе самообучения осуществляется с помощью "медленной" системы мышления. Сформированные навыки сохраняются в "быстрой" системе мышления. Стандартный механизм закрепления паттернов поведения в качестве быстро воспроизводимых навыков – это, по сути, просто повторение. Повторяющиеся осмысленные действия превращаются в автоматически воспроизводимые навыки после определенного количества повторений.

Такие модели поведения описываются терминами “эвристики” / “mental shortcuts” в системе “быстрого” мышления Канемана 1. Понятие "поведенческие паттерны" / “behavioral patterns” является более общим, поскольку оно охватывает не только разумных живых существ, способных к "медленному" мышлению по системе Канемана 2, но и неразумных живых существ, чьи поведенческие паттерны / инстинкты сформированы эволюцией.

- "Логическое" / "медленное" и "статистическое" / "быстрое" мышление полностью интегрированы и дополняют друг друга.

"Интуиция" и "здравый смысл" в статистически значимых случаях очень помогают сделать правильный выбор без долгих раздумий. Но в нечастых, но важных случаях они ошибаются. Логическое рассуждение находит правильное решение. Статистический вывод находит наиболее вероятное решение.

- Используя статистику вместо логики, мы всегда будем получать усредненные ответы вместо правильных. В простых случаях они будут совпадать почти всегда, в сложных – почти никогда.
- Байесовское понимание LLM является аналогом системы “быстрого” мышления Канемана 1. Байесовское понимание, отвечающее за “здравый смысл” / интуицию, содержит статистически накопленные результаты логических рассуждений для того, чтобы быстро воспроизвести их в виде готовых решений. Текстовые подсказки / промпты, вводимые пользователями LLM, являются триггерами для воспроизведения наиболее контекстуально схожих "поведенческих паттернов" / готовых текстов из большого набора обучающих данных.
- Логический анализ и понимание – это аналог системы “медленного” мышления Канемана 2.

Таким образом, присущее LLM байесовское понимание может обеспечить часть “реального” AGI, ответственную за усредненный “здравый смысл”, но только “логическое” понимание гарантирует логические рассуждения и отсутствие противоречий / “галлюцинаций” в текстах, сгенерированных искусственным интеллектом.

- Простая реализация системы мышления Канемана 1 / 2 возможна на графе знаний AGI:
 - Все сохраненные решения / пути на графе представляют “быстрое” мышление системы Канемана 1.
 - Все неизвестные пути, которые должны быть найдены / вычислены, являются предметом “медленного” мышления системы Канемана 2.
 - Пути, найденные системой “медленного” мышления Канемана 2, могут быть сохранены в системе “быстрого” мышления Канемана 1 с помощью некоторой служебной функции. В простейшем случае критерием сохранения "медленного" результата мышления в качестве "быстрого" решения может быть многократное использование. Результаты часто используемых "медленных" вычисленных решений сохраняются для дальнейшего "быстрого" воспроизведения без новых повторных вычислений.

8. Понимание и Предсказание

- Предсказание – это лишь подмножество понимания.
 - Точно так же, как отношение порядка является подмножеством множества всех отношений.
 - “Предсказание” и “понимание” – это довольно разные вещи. Хотя “понимание” включает в себя “предсказание”, “предсказание” частично возможно и без “понимания”.
- Нет смысла “предсказывать” то, что вы понимаете, знаете или сами создаете. Нет необходимости “предсказывать” следующее слово в создаваемом вами предложении, в отличие от того предложения, которое вы лишь пытаетесь угадать.
 - На 100% гарантировано предсказание событий, причины которых вам понятны или причиной которых являетесь вы сами.
- Предсказание основано на вероятности. Понимание основано на достоверности.
 - LLMs (Большие Языковые Модели) оперируют вероятностями. Реальный AGI (Общий Искусственный Интеллект) оперирует неопределенностью, через механизм понимания достигая достоверности.

9. Понимание и Интеллект

- Интеллект основан на понимании.
 - Понимание формирует смыслы. Смыслы формируют знание. Интеллект как оператор смыслов процессирует эту деятельность. Таким образом, интеллект использует понимание для оперирования смыслами создавая знания.
 - “Реальный” интеллект использует механизм рассуждения / понимания для создания текстов / событий / знаний, а не просто “предсказывает” их.
- AGI – “Общий” не означает “Сильный”:
 - “Сильный” интеллект находит наилучшее решение проблемы в пространстве известных альтернатив.
 - “Общий” интеллект переформулирует неразрешимую проблему, имеющую пустое пространство альтернатив, чтобы найти по крайней мере какое-то подходящее / не обязательно лучшее решение.

Обработка обратной связи является ключевым фактором для реального интеллекта. Обратная связь присутствует во всех саморегулирующихся системах. Поддержание жизни как системы гомеостаза требует обратной связи. В простейших механических системах саморегулирования обратная связь представлена жестким механическим регулятором, который позволяет системе поддерживать баланс процесса функционирования. Например, в старых маятниковых часах обратная связь реализована в виде анкерного механизма, который поддерживает стабильное равновесие системы за счет расходования потенциальной энергии пружины или груза. Для сложных интеллектуальных систем / AGI / HLAИ обратная связь является важным средством самообучения системы.

Норберт Винер считал, что все разумное поведение является результатом механизмов обратной связи.

“I repeat, feedback is a method of controlling a system by reinserting into it the results of its past performance. If these results are merely used as numerical data for the criticism of the system and its regulation, we have the simple feedback of the control engineers. If, however, the information which proceeds backward from the performance is able to change the general method and pattern of performance, we have a process which may well be called learning.” / “Я повторяю, обратная связь – это метод управления системой путем повторного внедрения в нее результатов ее прошлой работы. Если эти результаты просто используются в качестве числовых данных для критики системы и ее регулирования, то мы

получаем простую обратную связь от инженеров по управлению. Однако, если информация, поступающая в обратном порядке от исполнения, способна изменить общий метод и структуру исполнения, мы имеем дело с процессом, который вполне можно назвать обучением". (Wiener, N., 1950).

Для интеллектуальных систем процесс самообучения, основанный на обратной связи, требует механизма понимания, описанного выше. Обратная связь для интеллектуальных систем может быть представлена в виде текста, который система должна быть в состоянии понять / интерпретировать как знания. Итак, для интеллектуальных систем обратная связь – это средство самообучения, а понимание – это механизм интерпретации внешних данных, предоставляемых посредством обратной связи, как знаний.

10. Сознание

- ▲ Сознание есть восприятие с пониманием.
 - То есть сознание – это осмысленное восприятие.
 - Сознание начинается с квалиа (физиологических данных / сенсорного опыта, хранящихся в мозге в невербальной форме), но формируется пониманием как некоторые отношения между элементами квалиа.
 - Необходимо проводить различие между сознанием и самосознанием. Самосознание – это восприятие своего собственного сознания.
 - Сознание и понимание неразделимы.
 - Восприятие без понимания является бессознательным.
 - Понимание определяет восприятие как сознательное.
- Бессознательное восприятие формирует квалиа / сенсорный опыт / нейробиологические данные в невербальной форме.
- Сознательное восприятие формирует знание / субъективную онтологию.
 - Сознание на самом деле в основном только тем и занимается, что создает персональные / субъективные онтологии в реальном времени.
- Сознание может возникать в компьютерах, которые обрабатывают информацию определенным образом, а именно формируют знания, используя понимание из “сырых” данных, воспринимаемых в реальном времени.
- Сознание не является бинарным. Процесс обретения сознания является последовательным, постепенным.
 - Чем более сложные / абстрактные онтологии / отношения создаются, тем более высокий уровень сознания достигается.
 - Количественное измерение уровня сознания возможно представить как мощность графа знаний, содержащего созданные сознанием субъективные онтологии.

Элементарное сознание возникает, когда осознаются элементарные отношения между воспринимаемыми сущностями реального мира. Таким элементарным отношением может быть причинно-следственная связь, которая является формой отношения порядка. Брошенный камень падает на землю. Это может быть опасно. Осознание такой взаимосвязи уже является признаком элементарного "заземленного" сознания.

- Причина – это первый элемент упорядоченной пары сущностей, связанных отношением порядка. Следствие является вторым элементом такой пары.

Стремление понять элементарный порядок / причинность значительно увеличивает выживаемость живых существ, и это является наиболее важным фактором для возникновения первичного "заземленного" сознания.

▲ Хаос – это отсутствие порядка.

◦ Или непонимание / невосприятие порядка.

Всё является хаосом до тех пор, пока не будет воспринят или установлен субъективно понятый порядок. Последовательность знаков числа π – это "хаос", вполне однородный генератор случайных чисел для тех, кто не знает формулу / порядок вычисления числа π . Музыка Баха – это тоже "хаос", то есть просто бессмысленный шум для тех, кто ее не понимает. Но для некоторых это высший порядок. "Бах – доказательство существования Бога", перефразировано из (Cioran, E., 1995)

◦ Порядок – это форма смысла / отношения, а именно отношение порядка.

Хорошо известно очевидное утверждение о том, что мозг был сформирован эволюцией на основе "сенсорного заземления" / "sensory grounding" / обратной связи от органов чувств. Но эволюция искусственного "разума" вовсе не обязательно должна повторять эволюцию живого мозга.

Очевидно, что сознание живых существ неотделимо от "заземленного" восприятия, сформированного эволюцией. Однако "абстрактное" восприятие, которым обладает человек, характеризует более высокий уровень сознания, способный воспринимать виртуальные сущности, отличные от чувственно воспринимаемой "реальности". Нет ни необходимости, ни возможности "копировать" или воссоздавать такую целостную сущность, как человеческое сознание. Благодаря небинарной природе сознания, достаточно просто начать процесс обретения элементарного сознания как формирования простейших отношений между элементами восприятия. Таким образом, сознание человеческого уровня будет сформировано как естественная стадия такого развития. Такой подход вполне конструктивен, "программируем".

Любой пробел в объяснениях обычно возникает при отсутствии строгих определений обсуждаемых понятий и предметной области. Неправильное определение / масштабирование предметной области искажает проблему, препятствуя формированию подходов к ее решению. Сознание для нейрофизиологов и иже с ними – это не просто "трудная", а неразрешимая проблема именно из-за неправильного масштабирования предметной области. Сознание – это не вопрос нейрофизиологии. В нейронах нет сознания. Послойное изучение компьютерного процессора химиками с целью понимания системы его команд является примером такого же неправильного определения предметной области.

"Нейрофизиологический" подход смог сформулировать "трудную проблему сознания" (Chalmers, D., 1995), но не решить ее. Я называю это "нейрофизиологическим тупиком". Бессмысленно изучать мозг и нейроны в поисках сознания. Необходимо изучать ум и логику рассуждений. Почувствуйте разницу.

"Трудная проблема сознания", сформулированная в терминах, определенных выше, становится не такой уж трудной:

- сознание – это постоянный процесс осознания, а именно осмысленное восприятие воспринимающим субъектом реальных или виртуальных сущностей в реальном времени, т.е. формирование отношений / смыслов между воспринимаемыми данными;
- более низкий уровень сознания реализуется посредством "заземленного" восприятия, которое обрабатывает / формирует смыслы в воспринимаемых квалиа в реальном времени. Именно так воспринимается лист бумаги с написанным на нем числом π ;
- более высокий уровень сознания реализуется посредством "абстрактного" восприятия, которое выполняет постобработку / формирование смыслов в воспринимаемых квалиа, хранящихся в памяти. Именно так воспринимается само число π ;

- Ощущения еще не сознательны. Но чувства, которые возникают в результате понимания / осознания ощущений, уже являются осознанными. Непосредственное восприятие зеленого или кислого еще не является осознанным, это всего лишь данные, полученные от физиологических рецепторов. Понимание / формирование взаимосвязей между такими данными превращает эти ощущения в сознательное восприятие. Понимание в режиме реального времени сравнивает / отличает кислое от некислого, зеленое от незеленого. И это уже первичное / элементарное сознание.

Отношения / смыслы, постоянно формируемые сознанием, формируют знание / субъективную онтологию субъекта.

- Субъективная / личностная онтология на самом деле является результатом осознания взаимосвязей в воспринимаемом мире, восприятия его структуры, метаданных о содержании мира. Такой процесс осознания / распознавания является проявлением сознания. Без сознания невозможно воспринять онтологию / структуру окружающего мира или абстрактных сущностей. Просто лист бумаги, на котором описана эта онтология, ее внешний вид, запах и цвет. Это простое бессознательное восприятие.

- Таким образом, осознание отношений между различными реальными или виртуальными сущностями создает онтологические / семантические структуры, т.е. знание / субъективную онтологию.

- Сознание в основном только и занимается созданием своей собственной субъективной онтологии либо в режиме реального времени, интерпретируя восприятие, либо обрабатывая ранее полученные данные.

11. Сознание Субъективно

- В виртуальном мире нелегко отделить "оригинал" от "симуляции".

- “Если это выглядит как интеллект, плавает как интеллект и крикает как интеллект, то, вероятно, это и есть интеллект” – известная поговорка, перефразированная мной.

- Сознание субъективно.

- Сознание другого субъекта существует для нас только в том случае, если мы воспринимаем этот субъект (этого субъекта) как сознательный. Сознание другого субъекта не существует для нас, если мы не воспринимаем этот субъект как сознательный — независимо от любых других способов определения наличия сознания у такого субъекта. Мы можем верить утверждениям о наличии сознания у какого-то субъекта, но не воспринимать его как сознательный. Только наше субъективное восприятие является для нас доказательством существования сознания у других людей.

Если мы не воспринимаем поведение какого-либо объекта как сознательное, или если объект не проявляет никакого воспринимаемого нами поведения, то у нас нет способа определить наличие сознания у такого объекта. Возможно, такой объект не обладает сознанием. Также возможно, что наши каналы восприятия недостаточно мультимодальны, чтобы воспринимать поведение объекта как сознательное.

Число π и музыка Баха существуют только для тех, кто их воспринимает или, по крайней мере, знает формулу π / нотную запись. Восприятие чего-либо как сознательного и осмысленного субъективно.

- Два субъекта взаимно признают друг друга обладающими сознанием тогда и только тогда, когда:

- их каналы восприятия согласованы / совместимы;
 - их поведение / взаимодействие / коммуникация взаимно воспринимаются как осмысленные;
 - каждый из них считает себя сознательным.

Если один из этих субъектов не считает себя сознательным, то такой субъект не может считать другого субъекта как обладающего сознанием. Т.е., чтобы воспринимать чужое сознание, необходимо также обладать сознанием. Однако возможно обратное. Субъект, обладающий сознанием, может считать какой-либо объект как обладающий сознанием, даже если тот не считает себя / не является таковым, более того, например, управляется каким-то жестким алгоритмом (воспринимаемым субъектом, обладающим сознанием как осмысленный).

Камни могут обладать сознанием, но этот факт останется неизвестным, потому что сознание субъективно и проявляется через поведение / взаимодействие / коммуникацию, которых у камней нет. Мы считаем, что камни не обладают сознанием только потому, что у нас нет возможности проверить это. Возможно, камни просто не хотят с нами общаться.

- Не имеет значения, как “китайская комната” (Searle, J., 1980) дает ответы. Важно только то, готовы ли мы квалифицировать эти ответы как осознанные.

- “Китайская комната” (или любой другой объект) не обладает сознанием, если она не говорит по-китайски таким образом, что наше субъективное восприятие квалифицирует ее как обладающую сознанием.

- “Китайская комната” (или любой другой объект) обладает сознанием, если она говорит по-китайски таким образом, что наше субъективное восприятие квалифицирует ее как обладающую сознанием.

- Нет другого способа определить, что какой-то объект обладает сознанием, кроме нашего субъективного восприятия. Если вы не говорите по-китайски, вы не сможете квалифицировать своего собеседника как обладающего сознанием, несмотря на все его попытки объяснить вам это по-китайски.

Конечно, у нас с вами есть сознание, независимо от чьего-либо восприятия. Но это верно только для меня самого и для вас самих, но не для других. И это будет справедливо для других только тогда, когда они смогут воспринимать наше поведение как сознательное.

Всё это потому, что сознание – это восприятие с пониманием, и сознание субъективно.

Очень важно понимать, что утверждение о субъективности сознания не отрицает возможности его объективной оценки. Напротив, краткое определение сознания как восприятия с пониманием, сформулированное в этой статье, позволяет нам правильно сфокусировать разработку объективных тестов, позволяющих выявить наличие сознания у тестируемых субъектов.

Такие тесты должны диагностировать одновременное наличие определенного уровня восприятия и определенного уровня понимания, проявляемого испытуемыми в режиме реального времени.

Очевидная субъективность сознания при таком тестировании проявится в том, что реальные люди, которые, несомненно, обладают сознанием, не смогут пройти некоторые из этих тестов, но какое-то жестко запрограммированное программное обеспечение, не обладающее сознанием, успешно пройдет их. Все это уже наблюдается – многие реализации LLM успешно проходят тест Тьюринга, что не означает, что они обладают сознанием и интеллектом человеческого уровня. Проблема “объективизации” сознания путем разработки “объективных” тестов принципиально неразрешима из-за субъективности сознания, утверждаемой в этой работе, но для частных случаев и конкретных предметных областей разработка таких объективных критериев наличия сознания вполне возможна. Это очевидный дуализм субъективности и объективности сознания, зависящий от рассматриваемой области восприятия. Если испытуемый воспринимает и понимает число π , музыку Баха или китайский язык, это, несомненно, осознанно. Но если субъект их не воспринимает, это не значит, что у субъекта нет сознания. Возможно, область восприятия с пониманием субъекта находится в другом домене. Потому что сознание субъективно.

12. Сознание и Чувства

▲ Чувства есть ментальные интерпретации ощущений / эмоций.

Такие интерпретации требуют понимания / осознанности.

- Таким образом, Чувства \equiv {Ощущения U Понимание}.
- Чувства не обязательны для сознания, но чувства всегда осознанны.
- Чувства / ощущения / рефлексy / навыки / инстинкты / поведенческие паттерны являются триггерами мотивации.

13. Субъектность и Мотивация

▲ Субъектность есть способность быть субъектом.

- HLAI / AGI может обладать субъектностью, также называемой агентностью / agency .
- Агентность, способность быть агентом это не что иное, как субъектность.

▲ Субъект есть наблюдатель, способный к восприятию.

◦ Главной характеристикой субъектности является способность воспринимать. Восприятие – это основа субъектности. Но способность влиять на окружающий мир и вести себя непредсказуемо и целенаправленно имеет решающее значение для определения сущности как субъекта для любых других субъектов.

• Субъект обладает субъектностью, если его поведение мотивировано / целенаправленно и непредсказуемо.

◦ Субъектность субъективна. Пусть это не выглядит тавтологией, потому что это не тавтология. Поскольку очевидно, что восприятие целенаправленного и мотивированного поведения достаточно субъективно.

▲ Объект есть нечто, воспринимаемое как целое.

▲ Система есть объект с разнообразием.

◦ “Объект” означает, что система может быть каким-то образом отделена от окружающего мира в восприятии.

◦ “Разнообразие” означает, что система имеет воспринимаемую структуру — атрибуты, поведение, элементы и т.д.

◦ Таким образом, субъект также может быть понимаем как объект или система, которые ведут себя непредсказуемо и целенаправленно.

▲ Одушевленность – это сознательная субъектность.

◦ Таким образом, субъекты, обладающие самосознанием, одушевлены.

◦ Сущности, обладающие одушевленностью, могут считаться живыми, независимо от их природы.

• Сознание включает в себя субъектность. Но субъектность не обязательно означает наличие сознания.

• Субъектность – основа автономии искусственного интеллекта.

◦ Автономная / независимая система искусственного интеллекта должна обладать субъектностью.

▲ Мотивация – это осознанная или вынужденная потребность в действии.

◦ Объект с мотивацией – это субъект.

◦ Субъектность определяется мотивацией.

Электронный прицел без мотивации не опасен. Шариковая ручка также очень опасна, если есть мотивация. Калькулятор опасен, если он используется в опасных целях с опасной мотивацией. Добавление некоторой мотивации к любому приложению для обработки текстов, переводчику или поисковой системе превращает эти приложения в "интеллектуальных агентов" и придает им субъектность. Искусственный интеллект без мотивации – это просто инструмент, не имеющий

субъектности, сам по себе не опасный. Мотивация – это ключевое слово. Негативная мотивация – это то, чего нам следует опасаться и что должно быть предметом настройки / “выравнивания” искусственного интеллекта / “AI alignment”.

- В 1960 году Норберт Винер сформулировал проблему “выравнивания” ИИ как несколько нестрогую формулировку. С тех пор никакого формального определения этой проблемы не появилось.

“If we use, to achieve our purposes, a mechanical agency with whose operation we cannot interfere effectively ... we had better be quite sure that the purpose put into the machine is the purpose which we really desire.” / “Если мы используем для достижения наших целей механическую агентность, в работу которого мы не можем эффективно вмешиваться ... нам лучше быть совершенно уверенными в том, что цель, заложенная в машину – это та цель, к которой мы действительно стремимся”. (Wiener, N., 1960)

- “Выравнивание” ИИ, понимаемое как набор начальных ограничений, принципиально невозможно.
 - “Реальный” ИИ, формирующий свои рассуждения и цели независимо, не может быть “выравнен”.
 - Частичное решение проблемы существует только в виде непрерывного / последовательного процесса улучшения ограничений.
 - Выравнивать “аппроксимацию” несложно. Но выравнивать “рассуждения” / “выводимость” не так-то просто. Почувствуйте разницу.

Мы не можем достаточно надежно “выравнивать” рассуждения ИИ, они могут быть лишь несколько ограничены. Это фундаментально сложная проблема. Поскольку “настоящее” рассуждение может быть только “автономным” / независимым. Но в любом случае мы должны стараться регулировать мотивацию искусственного интеллекта. Искусственный интеллект без мотивации не обладает субъектностью. Однако он может быть использован субъектами / людьми с негативной мотивацией. Для “выравнивания” ИИ важна только мотивация – встроенная / в приложении, или внешняя / управляемая человеком.

- Субъектом мотивации в ИИ является либо разработчик ИИ, либо сама сущность / экземпляр ИИ.
 - Мотивация является основой для определения ответственности.
 - Если мотивация в автоматически сгенерированном контенте / действиях искусственного интеллекта доказана, то субъект мотивации несет ответственность за контент / действия.
 - Права появляются вместе с обязанностями и только после обретения субъектности. Субъект искусственного интеллекта может приобрести какие-либо публичные права только после признания его субъектности.

14. Субъектность и Свобода Воли

▲ Свобода воли для искусственного интеллекта – это свобода алгоритма выбирать решение при наличии альтернатив, используя либо встроенную мотивацию, либо метод случайного выбора при ее отсутствии.

- Свобода в неопределенности.

Свобода подразумевает наличие неопределенности, наличие альтернатив. Там, где все предопределено и нет неопределенности, нет и свободы. Только наличие неопределенности обеспечивает возможность выбора. Если нет неопределенности, то нет и выбора. Наличие двух определенных альтернатив уже является неопределенностью. И только эта неопределенность позволяет сделать выбор из таких альтернатив.

Маленький ребенок верит, что простейшая игровая программа в крестики–нолики 3x3 обладает субъектностью, то есть непредсказуемым целенаправленным поведением, но для тех, кто знает этот простой алгоритм, это не так. Тот, кто манипулирует другими людьми и знает наперед каждый их шаг, считает их объектами, а не субъектами. Но если эти люди не знают об этом, то они считают себя субъектами со свободной волей.

- Субъектность определяется способностью к мотивированному поведению.
- Свобода воли определяется способностью делать мотивированный выбор в условиях неопределенности.

15. Основное различие между LLM и “реальным” AI / AGI / HLAИ

*"Большинство всегда неправо; меньшинство иногда бывает правым".
– перефразировано из (Ibsen, 1882)*

- Используя статистику вместо логики, мы всегда будем получать усредненные ответы вместо правильных. В простых случаях они будут совпадать почти всегда, в сложных – почти никогда.
 - Это именно то, что делают LLM / Большие Языковые Модели.
- LLM оперируют вероятностями. AGI работает с неопределенностями.
 - Способность оперировать неопределенностью вместо вероятности позволяет “реальному” AGI решать когнитивные задачи.

LLM компилирует контексты из набора данных, приближая их к подсказке / промпту, всегда давая какой–то ответ, возможно, неверный. AGI формирует осмысленные контексты, используя рассуждения, не давая никакого ответа, если данных недостаточно и неопределенность не устранена. “Реальный” AGI обладает субъектностью и создает осмысленные тексты. Имитационный / статистический искусственный интеллект предсказывает наиболее вероятные тексты.

Несмотря на то, что “Байесовское” и “Булево” понимание дополняют друг друга, это разные ветви эволюции искусственного интеллекта, которые будут развиваться независимо друг от друга. LLM не смогут эволюционировать до уровня AGI “естественным” образом, простым масштабированием, они могут быть только интегрированы. Вопрос бесшовной интеграции между LLM и AGI чрезвычайно важен и фундаментален. Это интеграция между аппроксимацией и обобщением. Между вероятностью и рассуждением. LLM предоставит ассоциативные большие данные, а AGI обеспечит их логическую обработку.

- Могут ли LLM обладать сознанием? Определенно, нет. Потому что любая LLM – это процессор статистических аппроксимаций.
- Будет ли AGI обладать сознанием? Определенно, да. Потому, что сознание – это осмысленное восприятие, а AGI / “реальный” интеллект – это оператор смыслов.

16. Дополненный Интеллект / Augmented Intelligence

*«If you gaze long enough into an abyss, the abyss will gaze back into you». /
«Если долго всматриваться в бездну — бездна начнет всматриваться в тебя».
(Nietzsche, 1886)*

- Из теории множеств: любая сущность является дополнением к своему дополнению.
- Изобретая удобства цивилизации, человек становится зависимым от них и больше не может (или не хочет) выживать голым в джунглях.
 - Использование автопилота лишает навыков самостоятельного управления автомобилем.
 - Использование грамматических подсказок на телефоне позволяет не помнить грамматику.
 - Использование калькулятора лишает навыков арифметических вычислений.

Используя искусственный интеллект, перестаёшь тренировать свой собственный ум и становишься зависимым от искусственного интеллекта. Дополняя собственный мозг искусственным, сам становишься дополнением. Это не всегда плохо и не всегда хорошо, это просто констатация факта.

Дополненный Интеллект уже давно существует. Дополненный интеллект появился давно вместе с первым калькулятором (и даже абак можно считать таковым). Любое устройство, технология, предназначенные для облегчения / дополнения умственной деятельности, уже создаёт такой дополненный интеллект.

Дополненный Интеллект через некоторое время может превратиться в «Замещающий Интеллект» / «Substitutional Intelligence». Любые неиспользуемые умения и навыки, как известно, теряются, забываются, а неиспользуемые органы со временем атрофируются, деградируют, способность их применять исчезает. Это относится и к мышлению. Примером является все тот же калькулятор. Многие ли сегодня могут без него перемножить 2 числа? Очень немногие. А раньше могли почти все. Возможность скачать из Интернета и скопипастить курсовую работу приводит к деградации аналитического мышления у современных студентов. Соответствующие мыслительные способности деградируют или не развиваются. Естественный интеллект заменяется искусственным. Это реальная проблема. Для нескольких процентов людей новые возможности, связанные с ИИ, позволяют развить навыки мышления, освободив разум от рутинной работы по поиску информации и изобретения того, что давно существует. Для подавляющего большинства людей ИИ позволит просто перестать думать там, где раньше они были вынуждены это делать. Они замещают свою способность думать “суперкалькулятором”, который думает за них.

- Экземпляр AGI превращается из Дополненного Интеллекта в Замещающий Интеллект, если:
 - Естественный Интеллект (ЕИ) уже не может обойтись без данного экземпляра AGI, который выполняет жизненно важные функции
 - Экземпляр AGI обладает свободой воли, принимает собственные решения и его нельзя или нецелесообразно отключать.

Соответственно, в таких случаях Естественный Интеллект (ЕИ) очевидным образом сам превращается в Дополненный Интеллект.

- Итак, любой симбиоз (всякое взаимодействие) Естественного Интеллекта (ЕИ) и Искусственного Интеллекта (ИИ) (любого уровня сложности, начиная с калькулятора) создает интеллектуальную систему, в которой ЕИ и ИИ являются взаимодополняющими. Взаимная ценность и роли интеллектуальных компонентов такой системы зависит как от интегрированной целевой функции системы в целом, так и от частных целевых функции компонентов системы, обладающих субъектностью. Калькулятор субъектностью не обладает и его частной целевой функцией можно пренебречь. Но автопилот грузового автомобиля, принимающий решение о предотвращении аварии на дороге, может обладать субъектностью и его решение может превалировать над решением водителя-человека, не обладающего достаточно быстрой реакцией.
- Интеллект, трактуемый как способность понимать, интегрированный с каналами восприятия, образует сознание. Таким образом, дополняя каналы восприятия / органы чувств и инструменты обработки данных / понимания, мы дополняем / расширяем как интеллект, так и сознание.

17. Выводы

Таким образом, ответ на вопрос "Может ли искусственный интеллект обладать сознанием?" однозначно утвердительный. Искусственное сознание в смысле определений, сформулированных в данной публикации, вполне реализуемо в программном коде.

Из-за небинарной природы сознания реализация способности ИИ обладать сознанием может быть постепенной и начинаться с адаптивной способности программного кода определять простые отношения / причинно-следственные связи между воспринимаемыми в реальном времени сущностями окружающего мира / элементами кватерниона.

Ссылки

- Kahneman, D., 2011, Macmillan. "Thinking, Fast and Slow"
- Searle, J., 1980, Behavioral and Brain Sciences, Cambridge University Press. "Minds, Brains, and Programs"
- Chalmers, D., 1995, Journal of Consciousness Studies. "Facing up to the problem of consciousness"
- Friston, K.; Kilner, J.; Harrison, L., 2006. Journal of Physiology-Paris. "A free energy principle for the brain"
- Wiener, N., 1950, The Riverside Press, Cambridge, Massachusetts. "The Human Use of Human Beings: Cybernetics and Society"
- Wiener, N., 1960, Science. 131 (3410): 1355–1358. "Some Moral and Technical Consequences of Automation"
- Cioran, E., 1995, Chicago: University of Chicago Press. "Tears and Saints"
- Peirce, C., Unpublished notes, circa 1898. "Qualia-Consciousness"
- Peirce, C., Welch, Bigelow, 1870. "Description of a Notation for the Logic of Relatives"
- Nietzsche, F., 1886, "Beyond Good and Evil (Jenseits von Gut und Böse): Prelude to a Philosophy of the Future", Chapter IV. Apophthegms and Interludes, §146
- Ibsen, H., 1882, Copenhagen. "An Enemy of the People"
- Senkevich, V., 2022. "Existence and perception as the basis of AGI (Artificial General Intelligence)". arXiv:2202.03155 [cs.AI]