

PHILOSOPHICAL SCIENCES

ARTIFICIAL CONSCIOUSNESS THE NEXT STAGE IN THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE

Karakachanov A.

PhD of Philosophy

Associate Professor

The University of Library Studies and Information Technologies (SULSIT)

Sofia, Bulgaria

ИСКУССТВЕННОЕ СОЗНАНИЕ СЛЕДУЮЩИЙ ЭТАП РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Каракачанов А.

Доктор философии

Доцент

Университет библиотекознания и информационных технологий (УниБИТ)

София, Болгария

<https://doi.org/10.5281/zenodo.10975293>

Abstract

The article substantiates the inevitability of the emergence of artificial consciousness as a result of the development of artificial intelligence. The achievements of neurophysiology and cognitive psychology are used to argue this thesis. The basic stages through which any artificial or natural self-governing system must pass in order to reach the stage of consciousness and self-awareness are developed.

Аннотация

В статье обосновывается неизбежность возникновения искусственного сознания в результате развития искусственного интеллекта. Для аргументации этого тезиса используются достижения нейрофизиологии и когнитивной психологии. Разработаны основные этапы, через которые должна пройти любая искусственная или естественная система самоуправления, чтобы достичь стадии сознания и самосознания.

Keywords: Consciousness, sign, signal, artificial intelligence

Ключевые слова: Сознание, знак, сигнал, искусственный интеллект.

Сформулированное таким образом название настоящей статьи требует добавить, что предполагаемая неизбежность возникновения искусственного сознания весьма вероятна, если усилия по развитию искусственного интеллекта продолжатся. Чтобы обосновать это утверждение, необходимо сначала определить, что мы подразумеваем под «искусственным интеллектом».

Идея мыслящих искусственных существ исходит из античности и имеет длительную историю [8, с. 17-34], но более четкую формулировку получила в XX веке, особенно после введения термина «искусственный интеллект» в 56 г. Джоном Маккарти.

В целом проблема интеллекта наиболее полно рассматривается в когнитивной психологии. Но в нем выделяются десятки (по некоторым данным даже более 70) различных определений. Однако можно несколько условно заявить, что «интеллект — это способность учиться на основе опыта, используя метакогнитивные процессы для облегчения обучения, а также способность адаптироваться к окружающей среде. Это может потребовать различной адаптации в разных социальных и культурных контекстах» как утверждает Р. Стернберг [9, с. 555]. Похожее определение предлагает Говард Гарднер [1, с. 35], особенность которого состоит в

том, что в обоих случаях подчеркивается специфика интеллекта как адаптации в определенной культурной среде.

Очевидно, что для кибернетической системы, в которой возможность принятия решений ограничена ее программой, и даже когда эта система достигает высоких результатов в пределах определенных для нее границ, она остается ограниченной по интеллекту по сравнению со многими биологическими организмами. Причина в том, что последние представляют собой самоорганизующиеся системы, эволюционно ограниченные лишь необходимостью сохранения гомеостаза (выживания организма) или создания потомства, т.е. выживание вида. Это дает им несравненно больший простор для эволюционного развития и, в конечном итоге, развития интеллекта. У человека эти два эволюционных ограничения в принципе преодолимы, главным образом, благодаря наличию сознания, которое делает его максимально свободной самоорганизующейся системой.

Но есть еще одно важное ограничение современных кибернетических систем, которое не позволяет им быть самоорганизующимися системами. Это отсутствие возможности онтогенетических изменений системы, которая у живых организмов

подкрепляется наследственностью признаков и соответствующим филогенетическим развитием. Образуя, огромным недостатком современных кибернетических систем является то, что они не могут самостоятельно изменять как свое программное, так и аппаратное обеспечение в ответ на вызовы внешней среды.

Тезис, который я отстаиваю, заключается в том, что если будут преодолены технологические препятствия на пути превращения кибернетических систем в максимально раскрепощенные самоорганизующиеся системы, способные «эволюционировать» программное и аппаратное обеспечение, постепенная эволюция этих искусственных самоорганизующихся систем от форм низшего интеллекта к формам высшего интеллекта неизбежно приведет к структуре тех процессов, которые лежат в основе сознания. Возникновение в этих условиях животной (сигнальной) или человеческой (знаковой) формы сознания и самосознания практически неизбежно. Чтобы обосновать этот тезис, я должен вкратце рассмотреть вопрос о том, что представляет собой сознание, а также о некоторых основных нейрофизиологических механизмах, на которых строится психика? Но прежде, чем сделать это, необходимо обозначить некоторые важные изменения в мировоззрении последних десятилетий, которые не могут не отразиться на философско-методологическом подходе к данному вопросу.

Во-первых, с развитием высоких технологий постепенно уходит в небытие один из глубоко укоренившихся в массовом сознании древнейших философских тезисов: о сознании (духе, душе) как совершенно противоположном материи или, в лучшем случае, как о чем-то «квази» материальном. Позиции, защищаемые в том числе философами, декларирующими материалистические, даже марксистские позиции. Последнее, как ни парадоксально звучит, коренится в одном из догматических положений марксизма об «идеальности» сознания. Классическим примером в этом смысле является Э. Ильенков, определяющий идеальное как схему «реальной предметной деятельности человека, согласующуюся с формой вне головы, вне мозга... его можно представить и рассматривать как особое, абсолютно самостоятельное от устройства мозга» [4, с. 136].

Болгарский философ Сава Петров практически вступает на путь частичной субстанциализации сознания утверждая - «идеальное, универсальное, каким оно представляется «внутреннему видению» субъекта, должно быть включено в материальное... Очевидно, необходимо соответствующим образом обобщить понятие материального, чтобы его абсолютная гносеологическая противоположность идеальному была релятивизирована, не отменяя его» [6, с. 74]. В конце концов, Сава Петров признает «квазиобъективного монстра», который как отмечает Л. Гурова «противоречит основному принципу материалистического монизма» [2, с. 99].

Философские споры бесконечны, а решения приходят из практики, которая отправляет в небы-

тие ту или иную глобальную теоретическую проблему. Поэтому процесс «секуляризации» вопроса о сознании от этой идеалистической традиции получил импульс, со стороны современных информационных дисциплин, которые вполне целенаправленно пытаются конструировать психические процессы. Причина не является ни философской, ни идеологической, а вытекает из сущности любой истинной науки, которая по самой своей природе отказывается иметь дело с вещами, которые невозможно сконструировать. Такова «душа» в идеалистической философской традиции.

Во-вторых, современная нейрофизиология уже предоставила немало данных, которые хоть и не объясняют полностью сложные психические процессы, но позволяют существенно приблизиться к сущности сознания. Исходя из них, очевидно, что структура процессов сознания строится на более элементарных функционалах - безусловном и условном рефлексе. Последние неизбежны как основной этап развития психики и являются прежде всего проявлением сигнального этапа ее развития, который мы наблюдаем преимущественно в элементарной сенсорной психике.

Следующими уровнями развития психики являются перцептивный (отражение целостных объектов), а затем «интеллектуальный этап» — чувственное восприятие объективных отношений между объектами как объективных «ситуаций». На последнем этапе зарождается «знаковая» организация высших психических функций, которые становятся ведущими у человека.

Опираясь на знание этих основных физиологических процессов, мы можем приступить к раскрытию сущности сознания. Литература по данному вопросу, особенно по дисциплине «Философия сознания» (Philosophy of Mind), огромна, а если включить в нее труды ряда авторов от древнейших времен до современности, в которых вопросы о душе, духе, мышлении (все термины обозначают эту неуловимую психическую реальность), вполне можно впасть в отчаяние. Но как бы претенциозно это ни звучало, последние десятилетия дали нам возможность гораздо более оптимистично взглянуть на эту извечную проблему.

Не вдаваясь в подробности, как я это сделал в других публикациях [5], сформулирую дефиницию сознания как постоянного циклического процесса сравнения и синтеза сигналов. В этом варианте сознание можно обнаружить и у высших видов животных в виде так называемого «животное сознание» - или как циклический процесс, протекающий без участия знакоорганизованных нервных структур (цикл, не затрагивающий долевые отделы мозга). Однако у человека, помимо «животного сознания», появляется специфическое человеческое сознание. Это тоже процесс циклического синтеза информации, но уже на уровне высших психических функций, т. е. включает в себя филогенетически самые молодые отделы мозга (кору головного мозга), структура которых организована изоморфно знаку. Появление нового качества — «человеческого сознания» обусловлено окончательным

наложением знакового принципа организации структур коры головного мозга и циклического процесса синтеза информации, но уже с ведущей ролью ассоциативных полей.

Такой подход к сознанию основан на исследованиях ряда авторов. Согласно Эдельману [10] чувство «Я» и восприятие пространственно-временной непрерывности окружающего мира возникает в результате постоянного ввода в проекционные поля мозга информации, хранящейся в памяти. При этом сигналы памяти вводятся заново так, как если бы они были внешними. Этот процесс чрезвычайно быстрый с циклической повторяемостью примерно каждые 100 мс. То есть каждый момент мы словно вспоминаем, каким был мир 100 мс назад, и сравниваем его с настоящим. Это обеспечивает непрерывность сознания. Похожую гипотезу разработали В. Сергин [7, с. 627-639] и Иваницкий [3]. Гипотезу о возврате возбуждения в первичную кору как о механизме возникновения зрительных ощущений (как акта осознания) высказывают также П. Стериг и С. Брандт. [13, с. 117]. По их мнению, в основе «осознания» тактильных ощущений лежит возврат в первичную кору сигналов из вторичных полей.

Как мы видим, поступающие стимулы (сигналы) сопоставляются со следами памяти, которые не организованы в какой-либо самостоятельный образ. Также они хранятся в разных структурах мозга и получают свою интеграцию только в процессе осознания. Другими словами, сознание «оживает», а вместе с ним и интроспективно регистрируемое ощущение «образа» только в ходе циклического процесса сравнения данных извне и данных, хранящихся в памяти.

Если перевести все это на философский язык, то можно сказать, что сознание и психика — это процесс отношения субъекта к объекту, но в первом случае это циклический процесс. Тогда как второй (психический акт) представляет собой единовременный материальный процесс, в ходе которого формируется отношение (т.е. это процесс отношения) субъекта к объекту. Предмет — это заранее фило- или онтогенетически закрепленные в памяти, сигнальные (у низших животных) или знаковые (у некоторых высших антропоидов и человека) организованные ответы (реакции) на раздражения. Объект — это раздражения, поступающие в проекционные зоны коры головного мозга.

В чистом виде сигнал, как и знаковая психическая реакция, представляют собой разовые действия по принципу открытой дуги — поступающий сигнал проходит через разные структуры мозга, не возвращаясь, однако, к месту из чего оно пришло. Но в реальной жизни организма это происходит неоднократно и с большой частотой. С эволюционным развитием и усложнением условнорефлекторной деятельности возникает необходимость постоянного сравнения заученных комплексов восприятий с полученными в данный момент внешними восприятиями. Вполне логично, что те из отдельных процессов, построенных по принципу рефлекторной дуги, превращаются в циклические,

т.е. постоянно повторяющимися. И это является основой процесса осознания. Важная деталь здесь заключается в том, что когда этот циклический процесс построен на сигнальной нейронной активности, мы имеем так называемую «животное сознание», а когда это происходит на базе нервных структур, организованных изоморфно структуре знака, перед нами предстает человеческое сознание. Последняя, высшая ступень сознания — это самосознание, при котором «моделируемая» внешняя реальность включает в себя основные качества личности, в мозгу которой происходят процессы сознания. Таким образом, в общей картине мира среди множества предметов появляется еще один «предмет», в качества которого входят не только данные о собственном теле, поступающие от органов чувств, но также запоминаемые и связанные с ними психические состояния. Но это возможно лишь при наличии психики организовано на базе знаков, начальные проявления которой мы находим у ограниченного числа высокоразвитых видов животных.

После описанных таким образом этапов психического развития, интеллекта и сознания, как и их взаимосвязей, становится ясно, что развитие интеллектуальных кибернетических систем (искусственного интеллекта) невозможно без прохождения этих этапов. Высшая фаза этого развития — возникновение сознания и самосознания!

В заключение я коснусь принципиального вопроса, который неизбежно будет поставлен на повестку дня в связи с искусственным сознанием. А вопрос в том: можем ли мы признать, что искусственные (кибернетические) устройства, обладающие сознанием, этически равны людям? Принять их как существ, имеющих не меньшее, чем люди, право на жизнь, свободу и неприкосновенность? Противники этой позиции могли бы возразить как минимум по трем направлениям: у них нет души, они не являются биологическими существами (находятся на другой субстратной основе) и не возникли естественным путем — их создал человек.

Во-первых, несколько слов насчет аргумента о существовании «души» (в смысле философского идеализма) как основы превосходства человека над всеми другими природными или искусственными системами, обладающими сознанием. Как я уже указывал, существование нематериальных (духовных) сущностей недоказуемо, тем более они не могут быть сконструированы и поэтому остаются вне научного анализа. Однако социальная практика требует операционалистских решений, имеющих применение в жизни. Если мы попытаемся и дальше внедрять понятие «душа» как основной объяснительный и моральный принцип в социальную практику, то нам придется отказаться от любых воздействий на мозговые структуры человека, не говоря уже о протезировании его частей. Это сведет общество на уровень мышления средневековой религиозной секты. Естественно, можно было бы попытаться выйти из ситуации, встав на позицию эпифеноменализма и заявив, что нет никакой проблемы в проведении процедур на структурах мозга, поскольку душа — это нечто, что парит над

ними. Но тогда возникает принципиальный вопрос – почему душа не может парить и над искусственными структурами?!

По поводу утверждения, что сознание можно построить только на основе белковых структур. Из приведенных фактов ясно, что все характеристики таких психических функций, как интеллект и сознание, могут быть сконструированы (если не сейчас, то в ближайшем будущем) искусственно. Вопрос в том, должен ли он быть основан на углероде или же он может быть основан на другом химическом элементе или каких-то технических элементах? Я не буду здесь вдаваться в вопрос о том, как теоретически можно построить достаточно сложные молекулы на основе любого химического элемента, кроме углерода. Важно то, что серьезной теоретической основы для так называемой «углеродный шовинизм» попросту говоря нет!

Что касается неравенства искусственных систем, в том числе искусственного интеллекта, по отношению к их создателю, то мы можем определить эту «стадию» как «шовинизм по происхождению». Другими словами, это попытка найти принципиальное различие между естественными и искусственно созданными системами, даже если они имеют идентичные функциональные характеристики. Ведь экстракорпоральное оплодотворение – это тоже искусственное создание живой системы с вышеперечисленными характеристиками, называемой зиготой, но мы не объявляем созданное таким образом человека неполноценным. Нет принципиальной разницы, будет ли искусственное строительство происходить на клеточном уровне, на уровне сложных молекул или даже на атомном уровне!

Выводы, которые мы можем сделать, заключаются в том, что, если человечество продолжит усилия по созданию искусственного интеллекта, в результате оно, по всей вероятности, получит искусственное сознание. Тогда оно встанет перед дилеммой: признать ли его этически равным человеку со всеми вытекающими, в том числе юридическими, последствиями или попытаться подчинить

его себе со всеми возможными негативными последствиями.

Список литературы:

1. Гарднър Х., Нова теория за интелигентността. Множествените интелигентности на 21 век. Издателство Сиела, С. 2004, с. 35.
2. Гурова, Л. Понятия за обективно и субективно: 35 години по-късно. Философски алтернативи 2000, 3-4, с. 99.
3. Иваницкий А.М. Мозговая основа субъективных переживаний: гипотеза информационного синтеза // Журн. высш. нервн. деят. 1996. Т. 46. N 2.
4. Ильенков, Э. Проблема идеального. Вопросы философии, 1979, № 6., с. 136.
5. Каракачанов, А. Интеграционна теория за съзнанието. В: списание NotaBene, бр. 28, 2014, NotaBene-bg.org ISSN 1313-7859
6. Петров С. Доминанти във философията и науката. Университетско издателство “Св. Кл. Охридски” 1993, с. 74.
7. Сергин В.Я. Сознание как система внутреннего видения // Журн. высш. нерв. деят. 1994. Т.44. N 4-5. С. 627-639.
8. Тодорова, М. Изкуственният интелект. Кратка история на развитие и етични аспекти на темата. Издателство „Изток-Запад“, 2020
9. Стърнбърг, Р. Когнитивна психология. Издателство „Изток-Запад“, 2012, стр. 555
10. Edelman G.M. The remembered present. A biological theory of consciousness. New York: Basics Books, 1989. 346 pp.
11. Nigel J.T. Mental Imagery, Philosophical Issues About, Encyclopedia of Cognitive Science [Volume 2, pp. 1147-1153 - Editor in Chief, Lynn Nadel. London: Nature Publishing/Macmillan, 2003]
12. Ralph D. Consciousness and the brain <http://home.earthlink.net/~dravita/>
13. Stoerig P., Brandt S. The visual system and levels of perception: properties of neuromental organization // Theoretical Medicine. 1993. V.14. No. 2. P. 117