Искусственный интеллект, сознание и психиатрия

Giulio Tononi, Charles Raison. Artificial intelligence, consciousness and psychiatry. *World Psychiatry*. 2024;23(3):309-310

В 1966 году исследователь из Массачусетского технологического института представил ELIZA – компьютерную программу, которая симулировала личностно-ориентированную (клиент-центрированную) психотерапию К. Роджерса. Она перефразировала слова пациента в вопросы, следуя заранее подготовленным простым, но эффективным скриптам. Это была одна из первых успешных программ на основе искусственного интеллекта (ИИ). К разочарованию его создателя, некоторые люди принимали ELIZA за настоящего психотерапевта, возможно, из-за нашей склонности приписывать объектам сознание, если мы обнаруживаем у них интеллект, а в особенности – речь.

Неуклюжая попытка создания ИИ в виде ELIZA в наши дни преобразилась в невероятно красноречивого голема. ChatGPT легко переговорит, перепишет и превзойдет самого Зигмунда Фрейда. Поскольку большие языковые модели (large language models, LLM) обладают сверхчеловеческим лексиконом, знаниями, памятью и скоростью, искусственные мозги теперь могут превзойти естественные в большинстве задач.

ELIZA была названа в честь цветочницы Элизы Дулиттл, героини пьесы Дж. Б. Шоу «Пигмалион», поскольку ее речь также улучшалась по мере научения. Оригинальный миф о Пигмалионе – скульпторе, вырезавшем из слоновой кости идеальную женщину Галатею и надеявшемся ее оживить, – еще более уместен в данном контексте: предвещает ли создание ИИ искусственное сознание, возможно, даже сверхчеловеческое? В настоящее время начинают формироваться два лагеря, дающих радикально разные ответы на этот вопрос.

Согласно вычислительной/функционалистской позиции, доминирующей в когнитивной нейронауке, ответ – «да»[11]. Когнитивная нейронаука предполагает, что мы, в конечном счете, являемся машинами со сложным программным обеспечением (которое может давать сбои и быть перепрограммировано). Нейронные алгоритмы распознают объекты и окружение, направляют внимание, удерживают элементы в рабочей и долговременной памяти. Сложные нейронные вычисления управляют когнитивным контролем, принятием решений, эмоциональными реакциями, социальным поведением и, разумеется, речью. С этой точки зрения, сознание должно быть просто еще одной функцией, возможно, обеспечивающей прямую трансляцию всей информации[2] или метакогнитивную оценку сенсорных сигналов[3]. В таком случае, когда компьютеры смогут воспроизводить те же функции, что и наш мозг, пусть и реализованные иным способом (функционалистская «множественная реализуемость»), они будут обладать сознанием так же, как и мы.

Несмотря на то что поведение LLM во многом напоминает осознанное, нет способа определить, есть ли у них сознание, и если да, то каким образом и в какой степени оно проявляется[1]. Также неясно, как определить, чувствуют ли они что-либо (предполагаем, что просто спросить не получится…).

Когнитивная нейронаука обычно придерживается подхода «внешней перспективы», предложенного Галилеем, который стал основой для многих научных открытий. С этой точки зрения, сознание представляется либо как «иллюзия пользователя»[4], либо как загадочное «эмерджентное» свойство (согласно теории систем, эмерджентность – наличие у системы свойств, не присущих ее компонентам по отдельности. – *Прим. перев.*). Однако, как заметил еще Лейбниц, такой подход не учитывает наш опыт – то, что мы видим, слышим, думаем и чувствуем. Этот скрытый дуализм с самого начала ставит вопросы не только перед нейронаукой, но и перед психиатрией: лечим ли мы мозг, психику или и то и другое? И если оба, то как они связаны? Является ли душа лишь эфемерным пассажиром мозга?

Теория интегрированной информации (Integrated information theory, IIT) представляет собой радикально иной подход[5], и мы придерживаемся именно его. IIT придерживается *внутренней перспективы,* опираясь не на то, что такое мозг и что он делает, а на то, что такое сознание. В конце концов, для каждого из нас опыт является неоспоримой реальностью, а мир – это лишь результат осмысления этого опыта, о чем врачи-психиатры должны быть хорошо осведомлены.

IIT сначала характеризует сущностные свойства сознания – те, которые неопровержимо верны для любого мыслимого опыта, – а затем ставит вопрос о том, что требуется для их объяснения в физических терминах. Крайне важно, что это приводит к отождествлению опыта во всем его богатстве со *структурой* (а не с процессом, вычислением или функцией) – структурой, которая определяет способность (нейронного) субстрата в его текущем состоянии выстраивать причинно-следственные связи. Фактически IIT предоставляет вычислительный аппарат для определения, по крайней мере принципиального, является ли субстрат сознательным, каким образом и в какой степени.

Данная теория может объяснить, почему определенные участки мозга связаны с сознанием, а другие, например, мозжечок или некоторые участки префронтальной коры, – нет. Она может объяснить, почему – из-за разрыва причинно-следственных связей – сознание теряется во время сна без сновидений, при анестезии и генерализованных припадках[6]. Кроме того, данная теория стала учитывать качественные аспекты опыта – ощущение пространства как протяженного, а времени как текущего[7]. Это приводит ко многим предсказаниям, которые можно проверить и которые на первый взгляд контринтуитивны, например: практически «молчащая» кора может поддерживать яркое переживание «чистого присутствия» (медитативного состояния без мыслей, эмоций и чувства «я». – *Прим. перев.*). Наконец, IIT привела к разработке совместного применения транскраниальной магнитной стимуляции и электроэнцефалографии, который в настоящее время является наиболее специфичным и чувствительным для оценки наличия сознания у пациентов, не реагирующих на внешние стимулы[8].

Если IIT верна, то, в отличие от позиции доминирующей вычислительной/функционалистской теории, ИИ лишен (и будет лишен) какой-либо искры сознания: он может говорить и вести себя так же или даже лучше, чем мы (он будет «функционально эквивалентен»), но он не будет «феноменально эквивалентным» (он вообще ничего не будет чувствовать)[5]. Перефразируя Т. Нагеля, невозможно понять, «что значит быть» компьютером, каким бы умным он ни был. Точно так же, как и мозжечок, компьютер не обладает структурой, необходимой для наличия сознания. Несмотря на то что компьютер может безупречно выполнять все «когнитивные» функции, которые нам необходимы, включая те, которые мы привыкли считать исключительно человеческими, эти процессы будут происходить «в тени». Они будут осуществляться так же неосознанно, как процессы в нашем мозге, которые плавно соединяют фонемы в слова, а слова – в предложения, чтобы выразить мимолетную мысль.

Если IIT верна, то приписывание сознания ИИ действительно является «экзистенциальной» ошибкой, потому что сознание – это существование, а не действие, а ИИ – это, наоборот, действия, а не существование. Под давлением естественного отбора биологические ограничения могут способствовать совместной эволюции интеллекта и сознания (т.к. высокоинтегрированные субстраты оказываются в более выигрышном положении)[9]. Однако в более широком контексте интеллект и сознание могут быть разделены. Может существовать опыт без функциональных способностей, которые мы связываем с интеллектом. Например, па-

циенты, почти не реагирующие на внешние раздражители, могут быть неспособны что-либо сделать или сказать, но могут обладать богатым миром субъективных переживаний[8]. Так же как и могучий интеллект может существовать без сознания: красноречивый ИИ может вести увлекательную беседу и поражать нас своим интеллектом, при этом не существует ничего, кроме потока предложений, которые он выдает – как писал Ф. Ларкин «No sight, no sound / No touch or taste or smell, nothing to think with / Nothing to love or link with» («бесчувственность сама – нет вкуса, запаха, нет тела, нет ума...» – *в переводе Л. Эпштейна*).

Искусственный интеллект представляет собой уникальную и насущную проблему не только для психического здоровья, но и для нашего существования как вида и нашего места в природе. Если популярные в настоящее время вычислительные/функционалистские подходы верны и мы – сильно ограниченные и часто дефектные биологические машины, то мы скоро будем вытеснены машинами из кремния, которые будут не только лучше и быстрее, но и будут обладать более насыщенным и богатым внутренним миром. Если же верна IIT, то каждое ощущение является экстраординарным и невероятным по своей сути феноменом, который требует особого нейронного субстрата, который невозможно воспроизвести, просто имитируя его функции.

Giulio Tononi, Charles Raison

Department of Psychiatry, University of Wisconsin, Madison, WI, USA

Перевод: Ланикин Илья Константинович, г. Москва
Редактура: Шишковская Татьяна Игоревна, г. Москва

Библиография

1. Butlin P, Long R, Elmoznino E et al. arXiv 2023;2308.08708.
2. Dehaene S, Changeux J-P. Neuron 2011;70:200-27.
3. Brown R, Lau H, LeDoux JE. Trends Cogn Sci 2019;23:754-68.
4. Norretranders T. The user illusion: cutting consciousness down to size. New York: Viking Penguin, 1991.
5. Albantakis L, Barbosa L, Findlay G et al. PLoS Comput Biol 2023;19:e1011465.
6. Tononi G, Boly M, Massimini M et al. Nat Rev Neurosci 2016;17: 450-61.
7. Haun A, Tononi G. Entropy 2019;21:1160.
8. Casarotto S, Comanducci A, Rosanova M et al. Ann Neurol 2016;80:718-29.
9. Albantakis L, Hintze A, Koch C et al. PLoS Comput Biol 2014;10:e1003966.

DOI: 10.1002/wps.21222

Выход в свет диагностического руководства МКБ-11 «Клинические описания и диагностические требования»

Dévora Kestel, Geoffrey M. Reed. Global launch of the ICD-11 Clinical Descriptions and Diagnostic Requirements (CDDR). *World Psychiatry*. 2024;23(3):310-311

МКБ-11, ставшая первым за три десятилетия глобальным пересмотром Международной классификации болезней (МКБ), была утверждена на 72-й Всемирной ассамблее здравоохранения в мае 2019 года и вступила в силу в январе 2022 года в качестве основы для статистического отчета о состоянии здоровья людей в странах – членах Всемирной организации здравоохранения (ВОЗ). В настоящее время страны по всему миру находятся на различных этапах внедрения МКБ-11 в свои клинические и информационные системы здравоохранения, и этот процесс будет продолжаться в течение следующих нескольких лет.

ВОЗ сделала важный шаг на пути к внедрению МКБ-11 в систему психиатрической помощи, опубликовав «Клинические описания и диагностические требования (Clinical Descriptions and Diagnostic Requirements, CDDR) для психических и поведенческих расстройств и нарушений нейропсихического развития МКБ-11[1].

CDDR представляет собой полноценное диагностическое руководство, предназначенное для помощи специалистам в области психического здоровья и другим специалистам здравоохранения по всему миру в целях более точной диагностики психических расстройств в клинической практике. Оно содержит последовательную и клинически полезную информацию для всех диагностических категорий главы МКБ-11, посвященной психическим, поведенческим расстройствам и нарушениям нейропсихического развития. Данное руководство включает в себя описание ключевых признаков каждого расстройства (основные признаки), их границы с нормой (пороговые критерии) и с другими состояниями (дифференциальная диагностика), а также особенности, связанные с течением расстройства, возрастной, гендерной и культуральной спецификой. По сравнению с диагностическим руководством для МКБ-10, в CDDR содержится значительно расширенная и дополненная информация, интегрированы и систематизированы данные о возрастных, гендерных и культуральных особенностях, а также более тщательно представлены указания для дифференциальной диагностики[2]. Кроме того, новое руководство основано на современных исследованиях и данных лучшей клинической практики.

Центральной целью Департамента ВОЗ по вопросам психического здоровья и злоупотребления ПАВ при разработке CDDR было создание инструмента, способствующего снижению глобального бремени психических расстройств. Исходя из этой цели, при разработке CDDR особое внимание уделялось клинической полезности и применимости по всему миру наряду с научной достоверностью[3]. Точная диагностика часто является первым критически важным шагом на пути к получению адекватного лечения и помощи. Руководство, более близкое к клинической практике и более универсальное для применения в разных странах, имеет больше шансов на систематическое внедрение, что, в свою очередь, способствует более раннему выявлению лиц, нуждающихся в помощи, и подбору эффективного лечения. Это также улучшает качество данных о состоянии здоровья, собранных в условиях клинической практики, которые используются для управления и распределения ресурсов на уровне медицинских учреждений, системы здравоохранения, отдельного государства и всего мира.

Для разработки CDDR Департамент сформировал шестнадцать экспертных рабочих групп в различных областях, обеспечив тем самым мультидисциплинарный подход, который включал представителей всех регионов ВОЗ, в том числе значительное количество специалистов из стран с низким и средним уровнем дохода. В процессе разработки предложений для МКБ-11 эти рабочие группы проводили тщательные обзоры научных данных, включая исследования, проведенные в рамках создания DSM-5.

Кроме того, предложенные диагностические требования для МКБ-11 прошли обширные испытания в рамках программы полевых исследований. Для содействия участию клиницистов в разработке CDDR была создана Всемирная сеть клинической практики (Global Clinical Practice Network, GCPN), в которую в настоящее время входят более