

Введение

«Интеллект ChatGPT абсолютно бесчеловечен», — сказал Джеффри Хинтон (Geoffrey Hinton: ChatGPT's Intelligence Is Absolutely Inhuman). Принимая во внимание это не слишком приятное выражение «крестного отца» искусственного интеллекта, давайте взглянем на три сектора нашей жизни: общественные дела, военное дело и медицину, посмотрим, как на них влияет искусственный интеллект.

Искусственный интеллект (ИИ) идеален для правителя, который хочет тотальной власти над народом (чтобы никто даже не посмел усомниться в его власти).

Это идеально подходит для генерала, цель которого — «наплодить» как можно больше трупов противника. Беда только в том, что противник не дремлет и выдумывает всевозможные контрмеры. С другой стороны, системы вооружений становятся все более сложными. Им требуется все более мощное программное обеспечение, и оказывается, что не хватает человеческих ресурсов даже для разработки и эксплуатации этих систем вооружения, не говоря уже о киберугрозах.

Современная информационная среда ставит все новые и более сложные задачи перед национальной безопасностью, поскольку мир переходит от индустриальной эпохи к информационной. Достижения в области информационных технологий, беспроводной связи и социальных сетей увеличили скорость и диапазон информации, рассредоточили власть над информацией и изменили социокультурные нормы. Конкуренты и противники США используют эти достижения и связанное с ними влияние в информационной среде, чтобы свести на нет преимущества США в обычных вооруженных силах. Министерство обороны (Department of Defense, DOD) определяет информационную среду как совокупность индивидуумов, организаций и систем, которые собирают, обрабатывают, распространяют информацию или

действуют на ее основе. Эта среда имеет три измерения: (1) физическое, (2) информационное и (3) познавательное¹.

Физическое измерение относится к осязаемому, реальному миру. Это люди вокруг нас, начальники и управляющие машины, газеты, книги, вышки связи, компьютеры, серверы, ноутбуки, смартфоны и многие другие устройства.

Информационная среда связана со сбором, обработкой, хранением, распространением и защитой информации; где и как отдаются приказы и как контролируется их исполнение, в чем и как выражаются намерения командира.

Когнитивное (познавательное) измерение ориентировано на человека. Оно имеет множество аспектов: убеждения людей, нормы, слабости, мотивы, эмоции, опыт, мораль, образование, психическое здоровье, идентичность и идеологии.



Три измерения киберугроз. Источник: GAO-22-104714

Совместное использование различных технологий, электромагнитного спектра (EMS) и данных является неотъемлемой особенностью работы Министерства обороны в информационной среде. С этим связаны киберугрозы. Национальная стратегия США и Министерства обороны признают, что такие страны, как Китай, Россия, Иран и Северная Корея, представляют угрозу в информационной среде. Необходимо также учитывать негосударственные организации, агентов инсайдерской угрозы, иностранных террористов, транснациональные преступные организации, которые угрожают персоналу Министерства обороны как внутри страны, так и за рубежом.

¹ GAO-22-104714. Information Environment: Opportunities and Threats to DOD's National Security Mission. Sep 21, 2022.

Угрозы могут принимать различные формы: сбор информации или разведывательных данных, влияние на принятие решений, ухудшение возможностей электромагнитного спектра и кибератаки, все это отрицательно влияет как на персонал Министерства обороны, так и на возможности вооружения.



Киберугрозы в информационной среде. Источник: GAO-22-104714

С медициной тоже все очень сложно. Если подчиняться властям и объявлять эпидемию (как это было с ковидом в Латвии, когда эта эпидемия вызвала самую высокую смертность в Европе), то говорить тут нечего (даже рот открывать не рекомендуется). Однако если подойти к этому более разумно, то о роли искусственного интеллекта стоит говорить со всей серьезностью, стоит подумать, как помочь человеку жить долго, в полном разуме и здравии.

Международный спор идет об ограничении глубокого обучения (Deep Learning): его неконтрольное развитие потенциально может угрожать даже человечеству. В следующем тексте, особенно в третьей части (по медицине), будут примеры машинного обучения (Machine Learning), распознавания изображений (Computer Vision), дискриминантного анализа (как одного из методов Support Vector Machines) и нейронных сетей.

Будет ли искусственный интеллект действительно угрожать человечеству? Оправданы ли опасения? Тут стоило бы прислушаться к словам великого философа-просветителя Иммануила Канта (1724–1804). По его мнению, человек больше склонен ко злу, чем к добру. Человек часто пытается уклониться от следования основным принципам морали, то есть идет от святости и совершенства к аду и несовершенству. Кант говорит: «Закон,



Классификация методов искусственного интеллекта

живущий в нас, называется совестью. Совесть — это на самом деле наша способность действовать в соответствии с этим законом». Смогут ли конструкторы роботов дать им совесть и, тем более, способность ей следовать?



Сколько угроз вокруг: как их все предотвратить?
И после этого — останется ли время на добрые дела?

Предисловие. Некоторые мысли об искусственном интеллекте

Сначала я процитирую первое предложение книги. «Интеллект ChatGPT абсолютно бесчеловечен», — говорит Джеффри Хинтон (Geoffrey Hinton: ChatGPT's Intelligence Is Absolutely Inhuman)». Ниже приведены мои мысли об искусственном интеллекте «в свете этого не очень приятного выражения "крестного отца" искусственного интеллекта». Я доказывал автору, что слово «бесчеловечный» на самом деле имеет две разные семантики, первая — что-то очень плохое, а второе — что-то, что находится за пределами человеческих возможностей (например, летать по воздуху, как птица). Поскольку это высказывание вызвало большой резонанс в мировой прессе, за данным интервью Хилтона последовало более точное объяснение ситуации второй семантики этого утверждения². Вот выдержки из этого интервью:

Интеллект ChatGPT абсолютно нечеловеческий, поскольку в основе интеллекта Homo Sapiens лежит способность языкового общения, выработанная в ходе эволюции биологических существ, которая характеризуется передачей закодированных символов ограниченного количества данных на низких скоростях.

Интеллект ChatGPT основан на искусственно созданном интеллектуальном объекте с расширенными возможностями цифровой связи, характеризующимся передачей огромных объемов данных со скоростью, невообразимой для человека.

Я упомяну еще одну цитату Хилтона об искусственном интеллекте. «Во многих заголовках говорится, что, по моему мнению, [работу ChatGPT] следует прекратить сейчас, но я никогда этого не говорил», — утверждает Хинтон. «Прежде всего, я не думаю, что это возможно, и я думаю, что нам следует продолжать развивать эти работы, потому что они могут привести ко многим замечательным вещам. Но нам нужно приложить равные усилия, чтобы уменьшить или устранить возможные злые последствия.»

С автором книги я обсуждал угрозу интеллекта ChatGPT человечеству. Он подсказал оставить интригу для читателя и оставил лишь первую интерпретацию слова «бесчеловечный». Наверное, это правильно — в каждом литературном произведении

² <https://mpost.io/geoffrey-hinton-chatgpt-has-absolutely-inhuman-intelligence>

должна быть какая-то интрига, что-то должно быть подчеркнуто больше, что-то меньше. Я очень горжусь автором книги, моим чрезвычайно талантливым и разносторонним однокурсником по Латвийскому Университету (1954–1959) Манфредом Шнепс-Шнеппе. Так получилось, что наш дальнейший путь в жизни и особенно в науке для нас обоих был весьма схож. В этой книге Манфред неоднократно упоминает академика Колмогорова — одного из самых выдающихся математиков мира середины XX века, который, как мы теперь видим, сравнительно близко подошел к возможностям современного искусственного интеллекта. Наша научная деятельность во многом была связана с дальнейшим развитием идей Колмогорова, мне доводилось многократно встречаться с Колмогоровым лично и обсуждать различные научные проблемы.

Моя кандидатская диссертация (сейчас она называлась бы докторской) была посвящена проблеме универсальности наиболее общих математических автоматов (растущих автоматов), сформулированной Колмогоровым. Мне удалось доказать, что возможен конечный набор математических элементов, из которого можно «собрать» любой растущий автомат — в каком-то смысле любой математический «живой организм», способный трансформироваться и делать свои копии. Говоря в терминах начала статьи, можно сказать, что речь шла о математической «небиологической» теории живых организмов, которую можно определить с помощью простых математических элементов. Уже тогда (60 лет назад) целью было моделировать не только в самовоспроизведение этих математических организмов, но и в их эволюцию. Прискорбно констатировать, что подход к проблеме, по-видимому, опередил свое время как минимум на полвека. Даже сегодня мы до сих пор не можем (или не умеем) моделировать нетривиальную эволюцию даже на самых мощных компьютерах.

В заключение еще несколько слов о вреде искусственного интеллекта. Я считаю, что его опасность во много раз меньше, чем отравляющие боевые вещества и атомное оружие, которые до сих пор не применялись в больших масштабах.

Я принадлежу к тем оптимистам, которые верят, что искусственный интеллект в конечном счете изменит наш мир весьма позитивным образом.

Янис Барздиньш,
профессор Латвийского Университета,
академик Латвийской АН

Юрмала, 21 января 2024 года