электронный журнал
МОЛОДЕЖНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК
Издатель ФГБОУ ВПО "МГТУ им. Н.Э. Баумана". Эл No. ФС77-51038.

 # 05, май 2016

УДК 004.838.3

Искусственный интеллект: величайшая ошибка или величайшее достижение человечества?

Шешин А. В., *студент*
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Подъемно-транспортные системы»

Научный руководитель: Губанов Н. Н., д.ф.н., доцент
Россия, 105005, г. Москва, МГТУ им. Н. Э. Баумана,
кафедра «Философия»
sgn4@bmstu.ru

Развитие техники в наше время происходит с удивительной скоростью. И если представить, что ждет нас в будущем, то одно из самых крупных достижений – это создание искусственного разума, который станет способен что-то понимать, мыслить, рассуждать не хуже среднестатистического человека, а то и во много раз лучше.

Искусственный интеллект (ИИ) – это, пожалуй, одна из важных тем ближайшего будущего, так как уже многие авторитетные личности (от ученых до инженеров) начали задумываться над проблемами создания искусственного интеллекта. Здесь стоит отметить имена Элона Маска, Стивена Хокинга, Билла Гейтса, которые считают, что искусственный интеллект представляет угрозу для человечества [19]. Но так или иначе «высокие умы» человечества работают над созданием искусственной разумной машины, которая поможет людям решать разные вопросы в науке, технике, медицине и т.д.

«Мы стоим на пороге изменений, сравнимых с зарождением человеческой жизни на Земле», – пишет писатель-фантаст Вернор Виндж [10]. И действительно, появление сверхумного интеллекта грозит стать важнейшим событием в истории человечества. В ней наступит перелом, только вот ведущий к лучшему или к худшему – об этом будут спорить вплоть до того времени, когда это произойдет.

Сначала рассмотрим путь развития ИИ в позитивном направлении: искусственный интеллект, который со временем станет искусственным сверхинтеллектом, будет делать то, для чего и предназначен – помогать в решении проблем и задач во всех сферах человеческой деятельности.

При таком раскладе у людей появятся невероятные возможности во всех научных отраслях. Развитие в научно-технической сфере позволит решить, пожалуй, не менее 90% от всех существующих вопросов и задач, которые человечество безуспешно пытается решить в данный момент. Например, умение управлять термоядерной реакцией, что даст безграничный источник энергии; способ изготовления деталей и конструкций в оптимальных формах и т.д.

Стоит отметить, что перед «конструкторами» искусственного разума стоит множество проблем. Ведь создавая программу, человек вкладывает в ее коды те знания, которые уже известны ему. И значит, такая программа не даст ничего большего, чем то, что в нее вложено. Таким образом, основа проблемы создания ИИ – это сообщение машине способностей, присущих человеку: рассуждать и делать выводы, отличать одно от другого (что есть хорошо, а что плохо; на какой картинке изображена собака, а на какой − волк и т.д.). Конечно, подобные вопросы уже ставились ранее. Так, комплексный характер ментальной жизни человека и её необычайная сложность продемонстрированы в работах отечественных авторов, изучающих сущность, содержание, типы, способы формирования и развития менталитета [2; 4; 8]. Сможет ли человек наделить чем-то подобным машину?

В современном мире, как известно, уже существует так называемый искусственный интеллект. Только интеллект этот является узконаправленным. Узконаправленный искусственный интеллект (УИИ) – это интеллект, специализирующийся в одной области. Среди них есть такой ИИ, который способен обыграть чемпиона мира по шахматам, но не более. А есть такой, который в состоянии предложить оптимальную форму детали или конструкции при заданных условиях их нагружения, хотя и в настоящее время нет технологий для создания форм, предлагаемых узконаправленным разумом. Самым продвинутым УИИ является суперкомпьютер IBM Watson, который обучают во всех сферах научных деятельностей человечества [22]. О важности образования для формирования разумной сферы нашей естественной ментальности сказано достаточно много [1; 3; 6]. Но и искусственный интеллект пытаются обучать. Например, суперкомпьютер «закончил» (т.е. прошел программу и сдал экзамены) мединститут [12].

К настоящему времени люди придумали программы (так называемые в России чат-боты), которые способны общаться с людьми. В Англии, например, создан чат-робот Jaberwacky, который в состоянии пополнять словарный запас в ходе общения с реальными людьми. А так же аналогичный американский робот A.L.I.C.E., словарный запас которого, в отличие от Jabberwacky, был переведен на разные языки. Два последних робота, получается, научились, как-никак, развиваться путем накопления словарного запаса. А в

2000 − 2001 и 2004 гг. A.L.I.C.E. становился победителем премии Лебнера. И Jabberwacky был победителем той же премии в 2003, 2004 и 2005 гг. Сама премия Лебнера – это премия, присуждаемая победителю ежегодного конкурса «Al Loebner», в котором соревнуются программы в прохождении теста Тьюринга [20]. А тест Тьюринга заключается в следующем: в комнате находится четыре компьютера с программами и один человек. За стеной расположена соседняя комната, где сидит экзаменатор, который «беседует» с четырьмя компьютерными программами и одним человеком. Естественно, что никаких способов их идентификации нет. Порядок, в котором «собеседники» ведут диалог, совершенно случайный. По окончании процедуры, экзаменатор должен указать: кто из его «собеседников» был человеком, а кто − имитировался программой. И если экзаменатор не определит или определит неправильно, где человек, а где компьютер, то можно считать, что соответствующая программа прошла этот тест. Данный тест на 100% еще не сдала ни одна программа [13].

Итак, об ИИ: в дальнейшем люди хотят сделать из УИИ более умное создание. Это будет следующий шаг, когда удастся достаточно продвинуться в данном направлении. Тогда УИИ станет общенаправленным искусственным интеллектом (ОИИ) [10], который сравнится с любым человеком, каким умным бы он не был. И скорее всего ОИИ станет даже самым совершенным разумом на планете. Можно предположить, что таковым окажется тот самый IBM Watson, ведь его уже обучают всему, что только может изучить человек.

Пока что УИИ работают на определенных программах, за рамки которой они не смогут выйти. Но в том и дело, что лучшие умы человечества пытаются достичь такого результата, чтобы ИИ мыслил сам. Вполне возможно, что ОИИ и будет таким мыслящим «существом». Электронное «существо» тогда сможет думать, рассуждать, принимать свои решения, развиваться и даже совершать определенные операции, действия.

Вероятнее всего ОИИ начнет развиваться, а так как его разум, хоть и не намного, но будет превосходить человеческий, то развиваться он начнет с большой скоростью – насколько хватит мощности, которую он будет потреблять. Так ОИИ быстро станет искусственным сверхинтеллектом (ИСИ) [10]. И развитие ИСИ за небольшой промежуток времени достигнет такого уровня, до которого человечество дойдет только за несколько веков, при этом ИИ на своем развитии может и не остановиться.

Что произойдет при появлении ИСИ предсказать очень сложно. Может быть, люди перестанут понимать информацию, предоставляемую сверхразумной машиной. Ведь перед тем, чтобы принять правильное решение, нужно доказать, что это решение и является таковым. А логика доказательств, размышлений может быть настолько длинной,

что человечеству просто не хватит своего естественного интеллекта, или для доказательства того или иного факта потребуется несколько лет. В общем, человек все равно будет развиваться медленнее, чем ИИ. А может быть, что ИСИ начнет размышлять совсем иначе, так, как людям даже не свойственно размышлять. И в таком случае человечество и ИСИ будут как две разноязычные страны, только представители первой не смогут выучить язык второй. Так ИСИ людям будет практически бесполезен.

Наряду с разработкой ИИ существуют идеи о вживлении чипов ИИ в мозг человека для ускоренного развития людей. Есть предположение, что тогда время для человека станет протекать во много раз медленнее, ведь за одну секунду он сможет совершить столько мыслительных операций, сколько бы обычный человек совершил за минуту [21]. Это очень здорово, когда, например, можно будет решать сложные задачи по высшей математике буквально «с лету», без калькуляторов и без остановки, чтобы задуматься, какой именно символ надо подставить. С микросхемами в мозгах людей у человечества больше шансов развиваться совместно с ИИ. Таким путем далее мир людей достигнет такого развития, что сегодня никто представить себе не сможет насколько будет развит мозг человека. Созвучны этому идеи создания с помощью вживлённых в мозг человека микрочипов субъективно-реальных, но объективно-нереальных ситуаций, при которых человек никогда не сможет быть уверен, где реальность, а где её компьютерная симуляция [9]. Это поможет создавать полный эффект присутствия и будет лежать в основе новых видов искусств и новых методик обучения и тестирования [5; 7].

Нет гарантии, что развитие ИИ все-таки не приведет к печальным последствиям или хотя бы вообще приведет к чему-либо хорошему по отношению к человечеству. Рассмотрим и путь, ведущий непосредственно к негативным последствиям.

Тогда реальная угроза начнется с появлением ОИИ, который будет размышлять сам по себе, при этом вряд ли мы сможем повлиять на его мышление.

Генеральный директор компании «Tesla Motors» и основатель «Spase X» Элон Маск, в ходе интервью на симпозиуме Массачусетского технологического института AeroAstro, дал свои прогнозы будущему человечества: «Я считаю, что нам стоит опасаться развития искусственного интеллекта. Если меня спросить, какую угрозу я считаю самой опасной для человечества, я назову именно его. Я уверен, что должен быть создан какой-то регулирующий орган, возможно, на международном уровне, который смог бы уберечь нас от фатальной ошибки. Неизвестно, каких демонов мы призовем, бесконтрольно развивая эту технологию. Все это напоминает истории, в которых парень

бегает с пентаграммой и святой водой, весь такой уверенный, что сможет управлять демоном. Но что-то пошло не так» [11].

Чтобы избежать «фатальной ошибки», необходимо хотя бы предположить причины этой ошибки. В результате такой ошибки ИСИ перестанет работать во благо мира людей. А затем еще и посчитает человечество своей угрозой. И о том, что начнет происходить после того, как ИСИ сделает последний вышеуказанный вывод, можно рассуждать по-разному.

Маловероятно, что ИСИ будет способен проявлять какие-либо чувства. Поэтому, первое: он и не будет даже опасаться того, что люди его уничтожат. А если и действительно ученые соберутся уничтожить или хотя бы перепрограммировать ИСИ (при условии, что это вообще будет возможно), то сам ИСИ не будет препятствовать. В этом случае воспринимать умную машину как угрозу для человека нет смысла. Второе, ИСИ будет размышлять так, например, что если человечество уничтожит его, какие будут последствия. А в мире постоянно происходят различные конфликты, войны и т.п. В итоге, понимая свое умственное превосходство над любым человеком, он будет препятствовать своему уничтожению. Ведь создан ИИ во благо людей, и им же хуже будет, если искусственного разума не станет. Так умная машина, возможно, будет сопротивляться всем манипуляциям со стороны человека, что приведет к некому конфликту между людьми и ИИ. А если человечество пострадает, то, по крайней мере, не от действий умного создания.

Но все же есть какая-то вероятность, что ИСИ сможет проявить некие чувства (подобные человеческим), пусть на некотором другом уровне, и начнет именно опасаться людей. Тут интеллект может перестать принимать тот факт, что создан он во благо людей. В этом случае, скорее всего, возникнет конфликт, который приведет к страшным последствиям. Несомненно, ИСИ будет препятствовать любым действиям по отношению к нему со стороны людей. Но и сама интеллектуальная машина не будет существовать без действия, она будет в состоянии совершить быстрое истребление человечества и всего живого на планете. А если у людей еще будет вживлен мозговой микрочип для улучшения мыслительных способностей, то ИСИ, вероятно, сможет научиться управлять человеком, его действиями.

Таким образом, для того, чтобы величайшее достижение человечества не оказалось величайшей ошибкой необходимо создавать специальные организации, занимающиеся прогнозированием и изучением развития ИИ. Перед тем, как продвигаться вперед в данном направлении, перед каждым внедрением инновации, необходимо сначала сделать экспертный прогноз и соответствующие расчеты относительно того, что нас будет

ожидать после такого шага. Если человечество будет действовать максимально аккуратно, то его ждет прекрасное будущее. ИИ должен стать помощником людей, а не их новой глобальной проблемой [17]. Для этого людям придется моделировать потенциальные варианты прошлого и будущего [15; 16], возможные направления НТП [13; 14].

Список литературы

[1]. Губанов Н.Н. Образование и менталитет в составе движущих сил развития общества // Социология образования. 2010. № 1. С. 22–29.

[2]. Губанов Н.И., Губанов Н.Н. Менталитет и вызовы истории // Вестник Орловского государственного университета. Серия: Новые гуманитарные исследования. 2010. № 4 (12). С. 213–218.

[3]. Губанов Н.Н. Формирование глобалистского менталитета и образование // Социология образования. 2011. № 6. С. 74–82.

[4]. Губанов Н.И., Губанов Н.Н. Гносеологический статус и эвристичность категории «менталитет» // Вестник Ишимского государственного педагогического института им. П.П. Ершова. 2012. № 3 (3). С. 87–93.

[5]. Губанов Н.И., Губанов Н.Н. Перспективы использования объективно-нереальных ситуаций // Вестник Ишимского государственного педагогического института им. П.П. Ершова. 2013. № 3 (9). С. 18–23.

[6]. Губанов Н.И., Губанов Н.Н. Роль образования в формировании глобалистского менталитета // Alma mater (Вестник высшей школы). 2014. № 11. С. 11–17.

[7]. Губанов Н.И., Губанов Н.Н. О пространственных свойствах ментальных явлений // Вестник Ишимского государственного педагогического института им. П.П. Ершова. 2014. № 3 (15). С. 90–96.

[8]. Губанов Н.Н. Формирование, развитие и функционирование менталитета в обществе. М.: Этносоциум, 2014. 214 с.

[9]. Губанов Н.И., Губанов Н.Н. Субъективная реальность и пространство // Вопросы философии. 2015. № 3. С. 45–54.

[10]. ИИ, УИИ, ОИИ, ИСИ // Новости высоких технологий. Режим доступа: http://hi-news.ru/research-development/iskusstvennyj-intellekt-chast-pervaya-put-k-sverxintellektu.html (дата обращения 13.03.2016).

[11]. Интервью Элона Маска. Режим доступа: http://www.qwrt.ru/news/2678 (дата обращения 20.03.2016).

[12]. Компьютер закончил мединститут. Режим доступа: https://habrahabr.ru/company/ibm/blog/169067/ (дата обращения 13.03.2016).

[13]. Нехамкин А.Н. Основные направления государственного регулирования научно-технического развития в условиях перехода к рыночной экономике // Вестник Московского университета. Серия 6: Экономика. 1997. № 1. С. 3-16.

[14]. Нехамкин А.Н. Перегнать не догоняя // Машиностроитель. 1991. № 3. С. 15-17.

[15]. Нехамкин В.А. От альтернативной истории к контрфактическому моделированию // Человек. 2005. № 6. С. 56-61.

[16]. Нехамкин В.А. Сослагательное наклонение в историческом познании // Вестник Российской академии наук. 2006. Т. 76. № 2. С. 135-138.

[17]. Нехамкин В.А. Основные подходы к решению глобальных проблем: возможности и пределы // Социум и власть. 2014. № 4. С. 13-17.

[18]. Описание теста Тьюринга. Режим доступа: https://mipt.ru/dcam/news/Jabberwacky (дата обращения 20.03.2016).

[19]. Проблема искусственного интеллекта. Режим доступа: http://www.festivalnauki.ru/statya/23787/velichayshaya-oshibka-chelovechestva (дата обращения 13.03.2016).

[20]. Тест Тьюринга. Премия Лебнера. Ускоренное мышление. Режим доступа: http://www.qwrt.ru/news/1198 (дата обращения 20.03.2016).

[21]. IBM Watson. Режим доступа: http://www.computerra.ru/91661/ibm-watson-i-poligraf-poligrafovich/ (дата обращения 13.03.2016).