

СОДЕРЖАНИЕ

БЛАГОДАРНОСТИ	21
ПРЕДИСЛОВИЕ	22
САМОСОЗНАНИЕ ЧЕЛОВЕЧЕСКОГО УРОВНЯ У ИСКУССТВЕННОГО ИНТЕЛЛЕКТА	
МЮРРЕЙ ШАНАХАН	24
МЫСЛИТЬ — НЕ ЗНАЧИТ ПОДЧИНЯТЬ	
СТИВЕН ПИНКЕР	27
У ОРГАНИЧЕСКОГО ИНТЕЛЛЕКТА НЕТ БУДУЩЕГО В ДОЛГОСРОЧНОЙ ПЕРСПЕКТИВЕ	
МАРТИН РИС	31
ПОВОРОТНЫЙ МОМЕНТ ДЛЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА	
СТИВ ОМОХУНДРО	34
ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ — ЭТО Я	
ДИМИТАР САССЕЛОВ	37
НЕ МОЖЕШЬ ПОБЕДИТЬ — ПРИСОЕДИНЯЙСЯ	
ФРЭНК ТИПЛЕР	39
РАЗУМНЫЕ МАШИНЫ НА ЗЕМЛЕ И ЗА ЕЕ ПРЕДЕЛАМИ	
МАРИО ЛИВИО	41
ЛИЧНО Я ПРИВЕТСТВУЮ НАШИХ МЕХАНИЧЕСКИХ ПОВЕЛИТЕЛЕЙ	
ЭНТОНИ ГАРРЕТ ЛИСИ	43
НАШИ ГОСПОДА, РАБЫ ИЛИ ПАРТНЕРЫ?	
ДЖОН МАРКОФФ	46
СПРОЕКТИРОВАННЫЙ ИНТЕЛЛЕКТ	
ПОЛ ДЭВИС	49
СВЕРХРАЗУМНЫЙ ОТШЕЛЬНИК	
КЕВИН ХЭНД	51

ЭТО БУДЕТ ДИКАЯ ГОНКА	
ДЖОН МАЗЕР	54
ЭТОЙ ШТУКОЙ КТО-НИБУДЬ УПРАВЛЯЕТ?	
ДЭВИД КРИСТИАН	57
СВИДЕТЕЛЬ ВСЕЛЕННОЙ	
ТИМО ХАННЕЙ	60
ДАВАЙТЕ ПОДГОТОВИМСЯ!	
МАКС ТЕГМАРК	63
ВОПРОСЫ ДЛЯ РАСШИРЕННОГО ТЕСТА ТЬЮРИНГА	
ТОМАЗО ПОДЖО	67
ЭПОХАЛЬНОЕ ДЛЯ ЧЕЛОВЕКА СОБЫТИЕ	
ПАМЕЛА МАККОРДАК	71
ДОБРО ПОЖАЛОВАТЬ В ВАШЕ НОВОЕ СВЕРХЧЕЛОВЕЧЕСКОЕ «Я»	
МАРСЕЛО ГЛЕЙЗЕР	74
ВСЕ МЫ МЫСЛЯЩИЕ МАШИНЫ	
ШОН КЭРРОЛЛ	76
КРИЗИС КОНТРОЛЯ	
НИКОЛАС КАРП	79
МЫ ИХ ПОСТРОИЛИ, НО МЫ ИХ НЕ ПОНИМАЕМ	
ДЖОН КЛЕЙНБЕРГ, СЕНДИЛ МУЛЛАЙНАТАН	82
НАМ НАДО ДЕЛАТЬ ДОМАШНИЕ ЗАДАНИЯ	
ЯАН ТАЛЛИНН	86
КАКОЕ ВАМ ДЕЛО ДО ТОГО, ЧТО ДУМАЮТ ДРУГИЕ МАШИНЫ?	
ДЖОРДЖ ЧЁРЧ	88
МАШИНЫ НЕ МОГУТ ДУМАТЬ	
АРНОЛЬД ТРЕХУБ	92
У НИХ НЕТ НИ «Я», НИ СПОСОБНОСТИ К ЗЛОМУ УМЫСЛУ	
РОЙ БАУМЕЙСТЕР	93
ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ЧЕЛОВЕЧЕСКОГО ИНТЕЛЛЕКТА	
КИТ ДЕВЛИН	95

МАШИНА — МАТЕРИАЛЬНАЯ ВЕЩЬ	
ЭМАНЬЮЭЛ ДЕРМАН.....	98
Я МОГУ ОШИБАТЬСЯ	
ФРИМЕН ДАЙСОН.....	100
ПОЧЕМУ НЕЛЬЗЯ ВЫЧИСЛИТЬ БЫТИЕ И СЧАСТЬЕ	
ДЭВИД ГЕЛЕРНТЕР.....	101
НИ ОДНА МАШИНА НЕ ДУМАЕТ НАД ВЕЧНЫМИ ВОПРОСАМИ	
ЛЕО ЧАЛУПА.....	104
СИНГУЛЯРНОСТЬ — ГОРОДСКАЯ ЛЕГЕНДА?	
ДЭНИЕЛ ДЕННЕТ.....	106
НАНОИНТЕНЦИОНАЛЬНОСТЬ	
УИЛЬЯМ ТЕКУМСЕ ФИТЧ.....	110
ИГРЫ (И ПРОРОЧЕСТВА) РАЗУМА	
ИРЭН ПЕППЕРБЕРГ.....	114
КОЛОСС — ЭТО БДВ	
НИКОЛАС ХАМФРИ.....	116
ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ С САМОСОЗНАНИЕМ? НЕ В ЭТОМ ТЫСЯЧЕЛЕТИИ!	
РОЛЬФ ДОБЕЛЛИ.....	119
МАШИНЫ НЕ ДУМАЮТ, НО И ЛЮДИ — ТОЖЕ	
СЕЗАР ИДАЛЬГО.....	123
МЫ ЗАПУТАЕМСЯ В ВОПРОСАХ	
ДЖЕЙМС О’ДОННЕЛЛ.....	127
НЕ НАДО ПУТАТЬ ПРОИЗВОДИТЕЛЬНОСТЬ И КОМПЕТЕНЦИИ	
РОДНИ БРУКС.....	129
ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ СДЕЛАЕТ ВАС УМНЕЕ	
ТЕРРЕНС СЕЙНОВСКИ.....	133
ПОВЕРХНОСТНОЕ ОБУЧЕНИЕ	
СЕТ ЛЛОЙД.....	136
ЕСТЕСТВЕННЫЕ ТВОРЕНИЯ ЕСТЕСТВЕННОГО МИРА	
КАРЛО РОВЕЛЛИ.....	139

ТРИ ЗАМЕЧАНИЯ ОБ ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ ФРЭНК ВИЛЬЧЕК.....	141
КОГДА Я ГОВОРЮ «БРУНО ЛАТУР» Я ИМЕЮ В ВИДУ ВО ВСЕ НЕ «БАРАН ЛИ ТУР?» ДЖОН НОТОН	144
МЫ ПОКА ЧТО СТОИМ В САМОМ НАЧАЛЕ ПУТИ НИК БОСТРОМ.....	146
ЭВОЛЮЦИОНИРУЮЩИЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ ДОНАЛЬД ХОФФМАН.....	148
МЫСЛЯЩИЕ МАШИНЫ ЕСТЬ В КИНО РОДЖЕР ШЕНК.....	152
ПЕРЕСАДКА ГОЛОВЫ? ХУАН ЭНРИКЕС	156
ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ/ИСКУССТВЕННАЯ ЖИЗНЬ ЭСТЕР ДАЙСОН.....	159
МОЗГ И ДРУГИЕ МЫСЛЯЩИЕ МАШИНЫ ТОМ ГРИФФИТС.....	161
ОНИ ПРИНЕСУТ БОЛЬШЕ ПОЛЬЗЫ, ЧЕМ ВРЕДА МАРК ПЭЙГЛ.....	165
ДЕРЖАТЬ ИХ НА КОРОТКОМ ПОВОДКЕ РОБЕРТ ПРОВАЙН.....	168
СЛЕДУЮЩИЙ РЕПЛИКАТОР СЬЮЗАН БЛЭКМОР	170
ЧТО ЕСЛИ МЫ — МИКРОБИОМ КРЕМНИЕВОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА? ТИМ О'РАЙЛИ.....	173
ТЫ — ТО, ЧТО ТЫ ЕШЬ ЭНДИ КЛАРК.....	176
СИСТЕМА ПРАВА И ГОСУДАРСТВЕННОГО УПРАВЛЕНИЯ У ИСКУССТВЕННОГО ИНТЕЛЛЕКТА МОШЕ ХОФФМАН.....	180

РОБОТ С ТАЙНЫМИ ПЛАНАМИ	
БРАЙАН НАТСОН.....	183
УМЕЕТ ЛИ ПОДВОДНАЯ ЛОДКА ПЛАВАТЬ?	
УИЛЬЯМ ПАУНДСТОУН	186
НЕ БОЙСЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА	
ГРЕГОРИ БЕНФОРД.....	188
ЧЕГО-ЧЕГО Я БОЮСЬ?	
ЛОУРЕНС КРАУСС.....	192
ПРОЕКТИРОВОЧНЫЕ МАШИНЫ ДЛЯ РЕШЕНИЯ ПРОБЛЕМЫ СЛОЖНОСТИ МИРА	
ПИТЕР НОРВИГ	195
МАШИНЫ, КОТОРЫЕ РАССКАЗЫВАЮТ ИСТОРИИ	
ДЖОНАТАН ГОТТШОЛЛ.....	199
ПРОТОПИЯ, А НЕ УТОПИЯ ИЛИ АНТИУТОПИЯ	
МАЙКЛ ШЕРМЕР.....	201
ОГРАНИЧЕНИЯ БИОЛОГИЧЕСКОГО РАЗУМА	
КРИС ДИБОНА	204
ВСЯКОЕ ОБЩЕСТВО ИМЕЕТ ТАКОЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ, КОТОРОГО ЗАСЛУЖИВАЕТ	
ЙОША БАХ.....	207
ЗВЕРИ ОСТРОВА ИИ	
КВЕНТИН ХАРДИ.....	210
МЫ СТАНЕМ ОДНИМ ЦЕЛЫМ	
КЛИФФОРД ПИКОВЕР	213
РЕЗУЛЬТАТЫ ВНЕЗЕМНЫХ НАБЛЮДЕНИЙ ЗА ЧЕЛОВЕЧЕСКИМ ВЫСОКОМЕРИЕМ	
ЭРНСТ ПЁППЕЛЬ.....	216
КТО ОПЛАЧИВАЕТ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ, ТОТ И ЗАКАЗЫВАЕТ МУЗЫКУ	
РОСС АНДЕРСОН.....	219

Я МЫСЛЮ, СЛЕДОВАТЕЛЬНО, Я ИИ ДАНИЕЛЬ ХИЛЛИС.....	222
КАКИМ БУДЕТ МЕСТО ЧЕЛОВЕКА? ПОЛ САФФО.....	224
ВЕЛИКАЯ АФЕРА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДИЛАН ЭВАНС.....	226
ШАНСЫ В ПОЛЬЗУ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ЭНТОНИ АГИРРЕ.....	229
НОВАЯ МУДРОСТЬ ТЕЛА ЭРИК ТОПОЛЬ.....	232
ОТ ОБЫЧНОГО «Я» К ИСКУССТВЕННОМУ «Я» РОДЖЕР ХАЙФИЛД.....	234
НАМ НУЖНО НЕЧТО БОЛЬШЕЕ, ЧЕМ МЫШЛЕНИЕ ГОРДОН КЕЙН.....	236
А ТУДА ЛИ МЫ ИДЕМ? СКОТТ АТРАН.....	237
ДВЕ КОГНИТИВНЫЕ ФУНКЦИИ, КОТОРЫХ ВСЕ ЕЩЕ НЕТ У МАШИН СТАНИСЛАС ДЕАН.....	240
СРЕДИ МАШИН, А НЕ ВНУТРИ НИХ МЭТТ РИДЛИ.....	243
ЕЩЕ ОДИН ТИП МНОГООБРАЗИЯ СТИВЕН КОССЛИН.....	245
НАРРАТИВЫ И НАША ЦИВИЛИЗАЦИЯ ЛУКА ДЕ БЬАЗЕ.....	248
ОТВЕТСТВЕННОСТЬ ЧЕЛОВЕКА МАРГАРЕТ ЛЕВИ.....	252
УСИЛИТЕЛИ И ИСПОЛНИТЕЛИ ЧЕЛОВЕЧЕСКИХ РЕШЕНИЙ Д. А. УОЛЛАК.....	254

НАДО СДЕЛАТЬ ЭТУ ШТУКУ ТАКОЙ, ЧТОБЫ ЕЕ НЕВОЗМОЖНО БЫЛО НЕ ЛЮБИТЬ	
РОРИ САЗЕРЛЕНД.....	256
МАШИНЫ-АКТРИСЫ	
БРЮС СТЕРЛИНГ.....	259
НАЗОВЕМ ИХ ИСКУССТВЕННЫМИ ПРИШЕЛЬЦАМИ	
КЕВИН КЕЛЛИ.....	262
МАШИНЫ ВООБЩЕ ЧТО-НИБУДЬ ДЕЛАЮТ?	
МАРТИН СЕЛИГМАН.....	265
DENKRAUMVERLUST	
ТИМОТИ ТЕЙЛОР.....	268
АНАЛОГОВАЯ РЕВОЛЮЦИЯ, КОТОРАЯ НЕ СМЕЕТ НАЗВАТЬ СВОЕ ИМЯ	
ДЖОРДЖ ДАЙСОН.....	271
ЦЕННОСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА	
АББАС РАЗА.....	273
ИСКУССТВЕННЫЙ ОТБОР И НАШИ ВНУКИ	
БРЮС ПАРКЕР.....	276
ОЧЕНЬ ПОЛЕЗНЫЕ ХИТРОСТИ	
НИЛ ГЕРШЕНФЕЛЬД.....	280
AIRBUS 320 И ОРЕЛ	
ДЭНИЕЛ ЭВЕРЕТТ.....	282
ЧЕЛОВЕЧНОСТЬ	
ДУГЛАС КОУПЛЕНД.....	285
МАНИПУЛЯТОРЫ И МАНИПУЛИРУЕМЫЕ	
ДЖОШ БОНГАРД.....	287
НЕ НАЧИНАЕМ ЛИ МЫ ДУМАТЬ КАК МАШИНЫ?	
ЗИЯД МАРАР.....	290
ЕЩЕ ОДИН ФРАКТАЛЬНЫЙ ЭЛЕМЕНТ ОБЩЕЙ КАРТИНЫ	
БРАЙАН ИНО.....	293

Е-ГЕЯ — РАСПРЕДЕЛЕННАЯ ТЕХНОСОЦИАЛЬНАЯ ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА	
МАРТИ ХЁРСТ.....	296
РАЗУМ УЛЯ	
КРИС АНДЕРСОН.....	298
ГЛОБАЛЬНЫЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ УЖЕ ЗДЕСЬ	
АЛЕКС (СЭНДИ) ПЕНТЛЕНД.....	300
СТАНУТ ЛИ КОМПЬЮТЕРЫ ЧЕМ-ТО ВРОДЕ МЫСЛЯЩИХ И ГОВОРЯЩИХ СОБАК?	
РЭНДОЛЬФ НЭССЕ.....	304
МЫСЛЯЩИЕ МАШИНЫ И ТОСКА	
РИЧАРД НИСБЕТТ.....	307
НАХЕС ОТ НАШИХ МАШИН	
СЭМЮЭЛЬ АРБЕСМАН.....	310
ОБЩЕЙ МОДЕЛИ ПСИХИЧЕСКОГО НЕ БУДЕТ	
ДЖЕРАЛЬД СМОЛБЕРГ.....	312
ИМ НЕ ПОНЯТЬ СУЩНОСТИ ЧЕЛОВЕЧЕСКИХ ПЕРЕЖИВАНИЙ	
ЭЛЬДАР ШАФИР.....	315
ИНТУИТИВНАЯ ТЕОРИЯ МАШИНЫ	
КРИСТОФЕР ШАБРИ.....	317
МЫСЛЯЩИЕ СОЛОНЧАКИ	
УРСУЛА МАРТИН.....	320
МЫСЛЯЩИЕ МАШИНЫ ДЛЯ УБИЙСТВА СОДЕРЖАТ НАШУ СОВЕСТЬ В ЧИСТОТЕ	
КУРТ ГРЕЙ.....	322
КОГДА МЫСЛЯЩИЕ МАШИНЫ НАРУШАЮТ ЗАКОН	
БРЮС ШНАЙЕР.....	325
ЭЛЕКТРИЧЕСКИЕ МОЗГИ	
РЕБЕККА МАККИННОН.....	328
РОБОДОКТОРА	
ГЕРД ГИГЕРЕНЦЕР.....	331

СМОГУТ ЛИ МАШИНЫ СТАТЬ ТАКИМИ ЖЕ УМНЫМИ, КАК ТРЕХЛЕТНИЕ ДЕТИ?	
ЭЛИСОН ГОПНИК	335
КУРИЦА, КОТОРАЯ ИГРАЕТ В КРЕСТИКИ-НОЛИКИ	
КЕВИН СЛАВИН.....	338
ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ СДЕЛАЕТ НАС УМНЫМИ, А РОБОТОВ НАПУГАЕТ	
АЛАН АНДЕРСОН.....	340
КОГДА МЫСЛЯЩИЕ МАШИНЫ НЕ ЯВЛЯЮТСЯ БЛАГОМ	
МЭРИ КЕТРИН БЕЙТСОН	343
СПРАВЕДЛИВОСТЬ ДЛЯ МАШИН В МИРЕ ОРГАНИЦИЗМА	
СТИВ ФУЛЛЕР	345
НЕ БУДЬТЕ ШОВИНИСТОМ В ВОПРОСАХ МЫШЛЕНИЯ	
ТАНЯ ЛОМБРОЗО.....	348
ЗВУЧИТ ЧУДЕСНО	
ВИРДЖИНИЯ ХЕФФЕРНАН.....	351
МАШИНЫ, КОТОРЫЕ РАБОТАЮТ, ПОКА НЕ ПЕРЕСТАЮТ РАБОТАТЬ	
БАРБАРА СТРАУК.....	352
ПРАВИЛА МЕНЯЮТСЯ ПО ХОДУ ИГРЫ	
ШЕЙЗАФ РАФАЭЛИ.....	354
НЕНАПРАВЛЕННЫЙ ИНТЕЛЛЕКТ	
ЭДВАРД СЛИНГЕРЛЕНД.....	357
ЧЕЛОВЕЧЕСКАЯ КУЛЬТУРА — ПЕРВЫЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ	
НИКОЛАС КРИСТАКИС.....	359
ЗА ПРЕДЕЛАМИ ЗЛОВЕЩЕЙ ДОЛИНЫ	
ДЗЁИТИ ИТО.....	362
ФИГУРА ИЛИ ФОН?	
ДУГЛАС РАШКОФ.....	366
БЫСТРЫЙ, ТОЧНЫЙ И ГЛУПЫЙ	
ХЕЛЕН ФИШЕР.....	368

ОНИ СДЕЛАЮТ НАС ЛУЧШЕ?	
СТЮАРТ РАССЕЛ.....	371
ПРОБЛЕМА ЗАГРУЗКИ ЦЕННОСТЕЙ	
ЭЛИЕЗЕР ЮДКОВСКИ.....	374
ПО НАШЕМУ ОБРАЗУ	
КЕЙТ ДЖЕФФРИ.....	378
УМВЕЛЬТ БЕЗОТВЕТНОГО	
МАРИЯ ПОПОВА.....	382
ОНИ БУДУТ ДУМАТЬ О СЕБЕ?	
ДЖЕССИКА ТРЕЙСИ.....	384
КРИСТИН ЛОРИН.....	384
ОРГАНИЧЕСКОЕ МЫШЛЕНИЕ ПРОТИВ АРТЕФАКТНОГО	
ДЖУН ГРУБЕР, РАУЛЬ САУСЕДО.....	387
КОНТЕКСТ ОЧЕНЬ ВАЖЕН	
ПОЛ ДОЛАН.....	390
КАК ПРЕДОТВРАТИТЬ ВЗРЫВ ИНТЕЛЛЕКТА	
ТОМАС ДИТТРИХ.....	391
МЫШЛЕНИЕ ИЗНУТРИ ИЛИ СНАРУЖИ?	
МЭТТЬЮ ЛИБЕРМАН.....	395
МЯГКИЙ АВТОРИТАРИЗМ	
МАЙКЛ ВАССАР.....	399
ЧТО ИСКУССТВЕННЫЕ ИНТЕЛЛЕКТЫ БУДУТ ДУМАТЬ О НАС?	
ГРЕГОРИ ПОЛ.....	401
ИСТОРИЯ ДЖОНА ГЕНРИ	
АНДРИАН КРЕЙЕ.....	404
МАШИНАМ НЕИНТЕРЕСНЫ ОТНОШЕНИЯ	
НИК ЭНФИЛД.....	407
СЛЕДУЮЩАЯ ФАЗА ЭВОЛЮЦИИ ЧЕЛОВЕКА	
НИНА ЯБЛОНСКИ.....	409

ДОМИНИРОВАНИЕ ПРОТИВ ОДОМАШНИВАНИЯ	
ГЭРИ КЛЯЙН	411
МЫСЛЯЩИЕ МАШИНЫ ПОЯВЯТСЯ НЕ СКОРО	
ГЭРИ МАРКУС	414
МОЖНО ЛИ ИЗБЕЖАТЬ ЦИФРОВОГО АПОКАЛИПСИСА?	
СЭМ ХАРРИС	417
СМОГУТ ЛИ МЫСЛЯЩИЕ МАШИНЫ ПРЕОДОЛЕТЬ ЭМПАТИЧЕСКИЙ РАЗРЫВ?	
МОЛЛИ КРОКЕТТ	421
ЗАБОТЛИВЫЕ МАШИНЫ	
ЭБИГЕЙЛ МАРШ	424
ДВИГАТЕЛИ СВОБОДЫ	
АЛЕКСАНДР ВИССНЕР-ГРОСС	427
ЕСТЬ ВОПРОСЫ?	
САРА ДЕМЕРС	430
МЫСЛЯЩИЕ МАШИНЫ — ЭТО СТАРЫЕ АЛГОРИТМЫ НА БОЛЕЕ БЫСТРЫХ КОМПЬЮТЕРАХ	
БАРТ КОСКО	432
НЕДОСТАТКИ МЕТАФОРЫ	
ДЖУЛИЯ КЛАРК	436
УНИВЕРСАЛЬНОЕ ОСНОВАНИЕ ЧЕЛОВЕЧЕСКОГО ДОСТОИНСТВА	
МАЙКЛ МАККАЛОУ	439
МЫСЛИ О ЛЮДЯХ, КОТОРЫЕ МЫСЛЯТ КАК МАШИНЫ	
ХАИМ ХАРАРИ	442
МЕТАМЫШЛЕНИЕ	
ХАНС ХАЛВОРСОН	445
ЦЕННОСТЬ ОЖИДАНИЯ	
КРИСТИН ФИНН	447
ЭКОСИСТЕМА ИДЕЙ	
ДИРК ХЕЛБИНГ	450

ЖЕЛЕЗНЫЙ ЗАКОН ИНТЕЛЛЕКТА	
ДЖОН ТУБИ	452
МАШИНЫ, ВОРУЮЩИЕ МЫСЛИ	
МАКСИМИЛИАН ШИХ	456
НЕПРЕДВИДЕННЫЕ ПОСЛЕДСТВИЯ	
САТЬЯДЖИТ ДАС	459
ПО-РАЗНОМУ	
РОБЕРТ САПОЛЬСКИ	463
БУДУТ ЛИ МАШИНЫ ДУМАТЬ ЗА НАС?	
АФИНА ВУЛУМАНОС	464
«ПРОСТИТЕ ЗА БЕСПОКОЙСТВО»	
БРАЙАН КРИСТИАН	466
МОРАЛЬНЫЕ МАШИНЫ	
БЕНДЖАМИН БЕРГЕН	468
КОГДА МЫ ВЫДЕРНЕМ ВИЛКУ ИЗ РОЗЕТКИ	
ЛОУРЕНС СМИТ	470
МОНИТОРИНГ И УПРАВЛЕНИЕ ПЛАНЕТОЙ	
ДЖУЛИО БОКАЛЕТТИ	472
ПАНЭМПИРИЗМ	
ЯН БОГОСТ	475
КОГДА МИНЬОН ПЕРЕСТАЕТ БЫТЬ МИНЬОНОМ?	
ОБРИ ДИ ГРЕЙ	478
НЕДОСТАТОЧНО ЗАБАГОВАННЫЙ	
МАЙКЛ НОРТОН	482
БОЛЬШЕ ФАНКА, БОЛЬШЕ СОУЛА, БОЛЬШЕ ПОЭЗИИ И ИСКУССТВА	
ТОМАС БАСС	484
БУДУЩЕЕ ДЛЯ НАС ЗАКРЫТО	
ХАНС УЛЬРИХ ОБРИСТ	485
НЕМАТЕРИАЛЬНАЯ МЫСЛИМАЯ МАШИНА	
КУ ДЖОНГ ЭЙ ГУ ЧОНА	487

СМЯТЕНИЕ И ТОМЛЕНИЕ	
РИЧАРД ФОРМАН.....	488
КТО БОИТСЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА?	
РИЧАРД ТАЛЕР.....	491
Я ВИЖУ РАЗВИТИЕ СИМБИОЗА	
СКОТТ ДРЕЙВЗ.....	495
СМЕНА ПРЕДСТАВЛЕНИЯ О СОБСТВЕННОМ «Я» В РАСПРЕДЕЛЕННОМ МИРЕ	
МЭТТЬЮ РИЧИ.....	498
ПРЕДСКАЗЫВАТЬ БУДУЩЕЕ ЛЕГКО	
РАФАЭЛЬ БУССО.....	502
СТРАХ БОЖИЙ, НОВАЯ ВЕРСИЯ	
ДЖЕЙМС КРОУК.....	505
ТЮЛЬПАНЫ НА МОГИЛУ ВАШЕГО РОБОТА	
АНДРЕС РОЭМЕР.....	507
НА ПУТИ К НАТУРАЛИСТИЧЕСКОМУ ПРЕДСТАВЛЕНИЮ О РАЗУМЕ	
ЛИ СМОЛИН.....	510
МАШИНЫ, КОТОРЫЕ МЫСЛЯТ? ДИЧЬ КАКАЯ-ТО!	
СТЮАРТ КАУФМАН.....	514
БУДУЩЕЕ ПРОСТРАНСТВО ВОЗМОЖНОСТЕЙ ИНТЕЛЛЕКТА	
МЕЛАНИ СВОН.....	517
ЛЮБОВЬ	
ТОР НОРРЕТРАНДЕРС.....	521
ЗЛОВЕЩИЙ ТЕСТ ТРЕХ КОЛЕЦ ДЛЯ MACHINA SAPIENS	
КАЙ КРАУЗЕ.....	524
СВОБОДА ОТ НАС	
ГЕОРГ ДИЕС.....	528
БЕЗУПРЕЧНЫЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ — ЭТО ЧТО-ТО ИЗ ОБЛАСТИ НАУЧНОЙ ФАНТАСТИКИ	
ЭДУАРДО САЛЬСЕДО-АЛЬБАРАН.....	530

ЗАРОЖДАЮЩИЕСЯ ХИМЕРЫ «МАШИНА-ЧЕЛОВЕК»	
МАРИЯ СПИРОПУЛУ	533
ЧТО, ЕСЛИ ИМ НУЖНО СТРАДАТЬ?	
ТОМАС МЕТЦИНГЕР	536
ПОЙМЕМ ЛИ МЫ, ЧТО ЭТО СЛУЧИЛОСЬ?	
БЕАТРИС ГОЛОМБ	540
МЕТАПРЕЗЕНТАЦИЯ	
НОГА АРИХА	543
ПОСЫЛКА: ТАК МАЛО ВРЕМЕНИ — ТАК МНОГО ПРЕДСТОИТ СДЕЛАТЬ	
ДЕМИС ХАССАБИС, ШЕЙН ЛЕГГ, МУСТАФА СУЛЕЙМАН	545
ПРИМЕЧАНИЯ	548

Марвину Мински

БЛАГОДАРНОСТИ

Я хочу сказать спасибо Питеру Хаббарду из HarperCollins и моему агенту Макс Брокману за постоянную поддержку и выразить особую благодарность — уже в который раз — Саре Липпинкотт за вдумчивое и внимательное отношение к рукописи.

ПРЕДИСЛОВИЕ

Философские дискуссии 1980-х годов вокруг искусственного интеллекта (ИИ) — о том, могут ли компьютеры по-настоящему мыслить, обладать сознанием и т. д., — в последнее время привели к новой полемике о том, как нам относиться к тем формам искусственного интеллекта, которые, как многие утверждают, уже существуют. Эти ИИ, если они достигнут уровня сверхинтеллекта (согласно определению, которое дал Ник Бостром в одноименной книге*, вышедшей в 2014 году), могут создать экзистенциальные риски, способные привести к тому, что Мартин Рис называет «нашим последним часом». Стивен Хокинг недавно попал на первые полосы международных изданий, когда сказал ВВС, что, по его мнению, разработка полноценного искусственного интеллекта может означать конец рода человеческого.

Итак, вопрос *Edge*** — 2015: что вы думаете о машинах, которые думают?

Но подождите! Может быть, стоит задаться еще и таким вопросом: о чем станут думать эти машины? Захотят ли они иметь гражданские права? Будут ли обладать сознанием? Какое правительство ИИ избрали бы для нас? Какое собственное общество сформировали бы? Или их собственное общество будет и нашим обществом тоже? Станем ли мы и мыслящие машины включать друг друга в соответствующие круги эмпатии?

* Бостром Н. Искусственный интеллект: Этапы. Угрозы. Стратегии. — М.: Манн, Иванов и Фербер, 2016.

** Edge Foundation — организация, основанная Джоном Брокманом для объединения ученых и интеллектуалов и для популяризации передовых направлений науки. Организация издает сетевой журнал *Edge* (www.edge.org); каждый год редакция ставит новый вопрос, на который отвечают разные комментаторы — как представители научных кругов, так и просто известные люди, чье мнение может показаться интересным читателям. — *Прим. ред.*

Многие авторы *Edge* хорошо знакомы с передовыми исследованиями в области искусственного интеллекта, не важно, сами ли они проводят эксперименты или занимаются популяризацией науки. ИИ стал первой и центральной темой наших бесед с Памелой Маккордак и Айзеком Азимовым, когда мы впервые встретились в 1980 году. И такие дискуссии не стихают до сих пор, о чем свидетельствует, в частности, прошедшее недавно мероприятие *Edge* под названием «Миф об ИИ» — разговор с пионером виртуальной реальности Джароном Ланье, который представил собственную интерпретацию заблуждений и страхов, связанных с восприятием компьютеров как людей, и собрал множество комментариев пользователей, в том числе довольно провокационных.

ИИ становится все реальнее? Вступили ли мы в новую эру разумных машин? Пришла пора повзрослеть и задуматься. В этом году авторы *Edge* (а их почти 200 человек!) — люди вполне зрелые, избегающие упоминания всяческой научной фантастики в литературе и кино, такой как «Создатель звезд», «Запретная планета», «Колосс: Проект Форбина», «Бегущий по лезвию», «2001», «Она», «Матрица» и прочие борги*. Прошло 80 лет с тех пор, как Тьюринг описал свою универсальную машину, и надо уже воздать должное ему и другим первопроходцам в области исследования ИИ и отпустить их на заслуженный отдых. Историю этого предмета мы знаем — можете посмотреть, к примеру, книгу Джорджа Дайсона «Собор Тьюринга» (Turing's Cathedral). Но что происходит *сейчас*?

Так что давайте вернемся к вопросу *Edge* — 2015: что вы думаете о машинах, которые думают?

*Джон Брокман,
издатель и редактор онлайн-журнала Edge*

* Вымышленная кибернетическая цивилизация в сериале «Звездный путь». — *Прим. ред.*

САМОСОЗНАНИЕ ЧЕЛОВЕЧЕСКОГО УРОВНЯ У ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

МЮРРЕЙ ШАНАХАН

ПРОФЕССОР КОГНИТИВНОЙ РОБОТОТЕХНИКИ ИМПЕРСКОГО
КОЛЛЕДЖА ЛОНДОНА; АВТОР КНИГИ «ТЕЛО И ДУХОВНЫЙ МИР»
(EMBODIMENT AND THE INNER LIFE)

Предположим, что нам удалось обзавестись машиной, обладающей разумом человеческого уровня, — иначе говоря, способной сравниться с человеком в любой (или почти любой) сфере интеллектуальной деятельности, а впоследствии и превзойти его. Обязательно ли такая машина будет обладать сознанием? Это важный вопрос, поскольку утвердительный ответ означает, что нам надо остановиться и подумать. Как нам вести себя с такой вещью, если мы ее создадим? Способна ли она испытывать горе или радость? Достойна ли она обладать теми же правами, что и человек? Надо ли нам вообще приводить в этот мир сознающие себя машины?

Будет ли искусственный интеллект, сравнимый по уровню с человеческим, неизбежно обладать самосознанием? Это сложный вопрос. Одна из причин — тот факт, что сознание человека и других животных обладает множеством особенностей. Все мыслящие существа демонстрируют целеустремленность. Все они в большей или меньшей степени сознают мир, в котором живут, и содержащиеся в нем предметы. Все животные в известной мере проявляют когнитивную интеграцию, то есть могут концентрировать свои психические ресурсы — восприятие, воспоминания и навыки — на том, чтобы воздействовать на происходящее в нужный момент, стремясь реализовать собственные намерения. В этом смысле любое животное демонстрирует целостность, индивидуальность. Некоторые из них, в том числе и люди, также сознают сами себя: свое тело и течение мыслей. Наконец, большинство животных, а возможно, и все испытывают страдания, а некоторые способны сопереживать страданиям других.

У здоровых людей все эти свойства образуют единый комплекс. Но у искусственного интеллекта они в принципе могут существовать по отдельности. Следовательно, нам нужно уточнить вопрос. Какие особенности, которые мы ассоциируем с сознанием у людей (если таковые вообще существуют), станут необходимым дополнением к интеллекту человеческого уровня? Каждый из вышеперечисленных признаков (а перечень этот, конечно, не исчерпывающий) заслуживает отдельного пространного описания. Так что позвольте остановиться всего на двух из них, а именно: на осознании мира и на способности испытывать страдание. Я уверен, что способность сознавать окружающий мир — это действительно необходимое свойство интеллекта человеческого уровня.

Определенно, ничто не может обладать мышлением, сравнимым с нашим, если у него нет языка, — а главное назначение человеческого языка состоит в том, чтобы говорить о мире. В этом смысле интеллект тесно связан с тем, что философы называют интенциональностью. Кроме того, язык — социальный феномен, и основное его назначение в группе людей состоит в том, чтобы говорить о вещах, которые все они могут воспринимать сейчас (например: этот инструмент или тот кусок дерева), или воспринимали ранее (виденный вчера кусок дерева), или в принципе способны воспринять (кусок дерева, который мы, возможно, увидим завтра). Короче, язык основан на осознании мира. Для обладающего телом существа или робота оно будет проявляться через взаимодействия с окружением (обход препятствий, подбор предметов и т. д.). Но мы могли бы расширить понятие, включив в него распределенный, лишенный физического тела искусственный интеллект, оснащенный сенсорами.

Чтобы с уверенностью называться одним из вариантов самосознания, такой вид осознания мира, вероятно, также должен сопровождаться очевидной целеустремленностью и некоторой когнитивной интеграцией. Потому, скорее всего, три перечисленных свойства окажутся единым целым даже у ИИ. Но давайте на мгновение отвлечемся от этого вопроса и вернемся к способности испытывать страдание или радость. В отличие от осознания мира, нет очевидных причин полагать, что искусственный интеллект человеческого уровня должен обладать этим качеством, даже при том, что у людей он тесно связан с самосознанием. Легко представить

машину, которая, не испытывая чувств или эмоций, выполняет любые интеллектуальные задачи. У нее не будет самосознания, которое имеет решающее значение, когда речь идет о предоставлении прав. Как отмечал Иеремия Бентам, если мы размышляем над проблемой норм обращения с животными, вопрос состоит не в том, могут ли они мыслить, а в том, могут ли они страдать.

Нет такого предположения, что «простая» машина никогда не сможет испытывать страдания или радость, — но в данном вопросе кое-что связано с биологией. Идея, скорее, заключается в том, что способность страдать или радоваться может быть отделена от психологических особенностей, взаимосвязанных в человеческом сознании. Давайте внимательнее рассмотрим это условное разделение. Я уже говорил, что осознание мира может идти рука об руку с явной целеустремленностью. У животных осознание мира, того, что он способен дать во благо или во вред (в терминах Джеймса Джерома Гибсона), служит удовлетворению их потребностей. Животное демонстрирует понимание того, что хищник удаляется от него или же что потенциальная жертва приближается. В контексте целей и потребностей поведение животного целесообразно. Если ему что-то мешает, если цели его окажутся недостижимыми, а потребности — неудовлетворенными, это будет основанием для того, чтобы испытывать страдания.

А что насчет искусственного интеллекта, сравнимого с человеческим? Не возникнет ли у такого ИИ комплексного набора целей? Не получится ли при некотором стечении обстоятельств, что попытки достичь этих целей каждый раз будут оказываться неудачными? Уместно ли в таком случае сказать, что ИИ испытывает страдания, даже если учесть, что особенности строения делают его невосприимчивым к боли или неприятным физическим ощущениям, известным людям?

И вот тут наше воображение и интуиция отказываются идти дальше. Подозреваю, что ответа на этот вопрос мы не найдем, пока не окажемся лицом к лицу с реальным явлением. Только тогда, когда сложный искусственный интеллект станет привычной частью жизни, наша языковая игра окажется приспособленной к подобным чужеродным для нас сущностям. И не исключено, что тогда уже окажется слишком поздно решать, стоило ли вообще производить эти сущности на свет. К добру или к худу, они уже будут здесь.

МЫСЛИТЬ — НЕ ЗНАЧИТ ПОДЧИНЯТЬ

СТИВЕН ПИНКЕР

ПРОФЕССОР ПСИХОЛОГИИ ГАРВАРДСКОГО УНИВЕРСИТЕТА;

АВТОР КНИГИ «ЧУВСТВО СТИЛЯ: РУКОВОДСТВО ДЛЯ УМНЫХ

ПИСАТЕЛЕЙ В XXI ВЕКЕ» (THE SENSE OF STYLE: THE THINKING PERSON'S
GUIDE TO WRITING IN THE TWENTY-FIRST CENTURY)

По меткому выражению Томаса Гоббса, рассуждение есть не что иное, как исчисление, и это одна из величайших идей во всей истории человечества. Представление о том, что разумность может достигаться в результате физического процесса вычисления, было высказано в XX веке в работах Алана Тьюринга (который утверждал, что простейшие машины могут выполнять любые вычислимые функции) и реализовано в моделях Дональда Хебба, Уоррена Мак-Каллока и Уолтера Питтса и их последователей, продемонстрировавших, что сети упрощенных нейронов добиваются сопоставимых с человеческими результатов. Познавательные способности мозга можно описать количественно. Грубо говоря (и пусть нас за это критикуют), убеждения — вид информации, мышление — вид вычисления, а мотивация — вид обратной связи и контроля.

Это концепция великолепна, и вот почему. Во-первых, она дополняет естественнонаучное понимание Вселенной, изгоняя из машины оккультистских духов, призраков и душу. Как Дарвин дал возможность вдумчивому наблюдателю за миром природы обходиться без креационизма, так Тьюринг и другие позволили наблюдающему за миром мышления избавиться от спиритуализма.

Во-вторых, вычислительная теория разума открывает двери для искусственного интеллекта, то есть для машин, которые мыслят. Рукотворное вычислительное устройство может, в принципе, заменить человеческий разум и превзойти его в силе. Из этого во все не следует, что такое непременно произойдет на практике, поскольку в обозримой перспективе вряд ли возникнет надежная технологическая база и экономическая потребность для того,

чтобы осуществить нечто подобное. Как изобретение автомобиля не было попыткой создать точную копию лошади, так и разработка окупающихся систем искусственного интеллекта не означает создание чего-то подобного *Homo sapiens*. Устройству, которое создано, чтобы управлять автомобилем или прогнозировать развитие эпидемии, не обязательно уметь привлекать полового партнера или избегать несвежего мяса.

Тем не менее предпринимаемые в последнее время крошечные шаги к более разумным машинам вновь пробудили у многих страх того, что знание нас погубит. Мое личное мнение таково: боязнь того, что машины могут вдруг обезуметь, есть не что иное, как напрасная трата эмоциональных сил — такой сценарий ближе к «проблеме 2000 года», чем к Манхэттенскому проекту.

Прежде всего, у нас достаточно времени для того, чтобы все предусмотреть. От искусственного интеллекта, сравнимого с человеческим, нас все еще отделяют от 15 до 25 лет, а многие из недавних широко разрекламированных прорывов в этой области на деле оказались довольно поверхностными. В прошлом «эксперты» со смехом отвергали возможность быстрого технического прогресса — а он взял да и случился. «Эксперты» также вещали (местами даже с паникой в голосе) о неминуемых научно-технических достижениях, которые так и не произошли: об автомобилях с атомными двигателями, о подводных городах, о колониях на Марсе, о «проектируемых» детях и о хранилищах с зомби, обеспечивающими людей запасными органами.

Кроме того, странно было бы думать, что робототехники не станут заботиться о мерах безопасности, улучшая системы искусственного интеллекта. Для этого не понадобятся занудные «законы робототехники» или какая-нибудь новая моральная философия, достаточно будет здравого смысла — того же, которым люди руководствовались, разрабатывая кухонные комбайны, циркулярные пилы, обогреватели и автомобили. Бояться того, что системы ИИ станут чрезвычайно искусными в выполнении своей конкретной задачи (например, в распределении энергии) и в результате станут самовольно пренебрегать другими задачами (например, безопасностью людей), — значит предполагать, что мы будем создавать машины быстрее, чем разрабатывать соответствующие меры безопасности. В действительности же прогресс в области

искусственного интеллекта, несмотря на всю шумиху вокруг него, довольно скромнен, и пройдет еще много времени, прежде чем будут достигнуты сколько-нибудь значительные успехи, причем на каждом из предстоящих этапов работы люди будут держать отвертки наготове.

Станет ли машина, обладающая искусственным интеллектом, намеренно отключать системы безопасности? А зачем это ей? Антиутопии об ИИ проецируют мещанскую психологию альфа-самца на представление об интеллекте. Авторы таких книг предполагают, что нечеловечески умные роботы однажды решат свергнуть своих хозяев или захватить мир. Но разумность — это способность использовать новые оригинальные методы для достижения цели, сами же цели лежат за пределами собственно разума. Быть умным и желать чего-либо — не одно и то же. Да, история знает немало деспотов с манией величия и серийных убийц-психопатов, но они появились на свет в результате естественного отбора, сформировавшего тестостерон-сенситивные нервные цепи у отдельного вида приматов, так что стремление доминировать — вовсе не обязательный признак разумных систем. Что характерно, многие из известных технопрофетов даже не задумываются о том, что искусственный интеллект может естественным образом развиваться по женскому типу, то есть будет вполне в состоянии справляться с решением своих задач, не испытывая желания истреблять невинных или властвовать над цивилизацией.

Нам легко представить злодея, который создает армию роботов и выпускает их в мир как оружие массового уничтожения. Но сценарии таких катастроф легко разыграть лишь в воображении, на самом же деле для их реализации потребуется длинная цепочка событий, вероятность каждого из которых довольно невелика. Сперва должен появиться злобный гений, обладающий одновременно жадной бессмысленной массовой убийства и гениальностью в области технических инноваций. Ему потребуется собрать команду сообщников и руководить ими, при этом обеспечивая секретность и лояльность подчиненных, а также заботиться об их компетентности. В течение всей операции нужно будет избежать множества опасностей: раскрытия заговора, предательства, внедрения сотрудников правоохранительных органов, совершения ошибок и неудачного стечения обстоятельств. То есть

в теории такое может случиться, но эту проблему вряд ли стоит считать насущной.

Если отвлечься от катастрофических научно-фантастических сюжетов, то перспективы у продвинутого искусственного интеллекта самые радужные, причем не только в том, что касается его практического применения, — например, беспилотные автомобили — это безопасность, экономия рабочего времени и экологичность, — но и в философском плане. Вычислительная теория разума никак не объясняет существование самосознания в значении личной субъектности. Однако она прекрасно справляется с объяснением существования субъектности в значении воспринимаемой и создаваемой информации. Есть мнение, что субъектность присуща любой достаточно сложной кибернетической системе. Раньше я думал, что эта гипотеза (во всех своих разновидностях) совершенно непроверяема. Но представьте себе разумного робота, запрограммированного следить за собственными системами и задавать научные вопросы. Если бы он без подсказки извне задался вопросом о том, откуда берутся его субъективные переживания, то я бы отнесся к этой идее всерьез.

У ОРГАНИЧЕСКОГО ИНТЕЛЛЕКТА НЕТ БУДУЩЕГО В ДОЛГОСРОЧНОЙ ПЕРСПЕКТИВЕ

МАРТИН РИС

БЫВШИЙ ПРЕЗИДЕНТ КОРОЛЕВСКОГО НАУЧНОГО ОБЩЕСТВА,
ПОЧЕТНЫЙ ПРОФЕССОР КОСМОЛОГИИ И АСТРОФИЗИКИ
КЕМБРИДЖСКОГО УНИВЕРСИТЕТА, ДЕЙСТВИТЕЛЬНЫЙ ЧЛЕН ТРИНИТИ-
КОЛЛЕДЖА; АВТОР КНИГИ «ОТСЮДА В БЕСКОНЕЧНОСТЬ» (FROM HERE
TO INFINITY)

Потенциал продвинутого искусственного интеллекта и опасения по поводу негативных сторон его развития становятся все более актуальными — и небезосновательно. Многие из нас задумываются над вопросами из области ИИ, например, отрасли синтетических биотехнологий уже сейчас необходимы основные принципы «инновационной ответственности»; другие считают наиболее обсуждаемые сценарии излишне футуристичными и недостойными особого внимания.

Но расхождение во взглядах в основном идет по временной шкале: различаются оценки скорости движения, а не его направления. Мало кто сомневается, что машины будут все больше и больше превосходить нас в плане способностей, характерных именно для людей, или же смогут усиливать их посредством кибернетических технологий. Наиболее осторожные из нас предполагают, что такие трансформации займут века, а не десятилетия. Как бы там ни было, временные рамки технологического прорыва — лишь мгновение по сравнению с тем, сколько длится приведший к появлению человечества естественный отбор, и (что более важно) они составляют меньше одной миллионной от необъятных просторов ожидающего нас будущего. Потому-то в долгосрочной эволюционной перспективе люди и все, о чем они когда-либо думали, станут всего лишь примитивной переходной формой, предшествовавшей более глубокому мышлению новой

машиноориентированной культуры, простирающейся в отдаленное будущее и далеко за пределы Земли.

Сейчас мы наблюдаем ранние этапы этого перехода. Нетрудно представить себе гиперкомпьютер, достигающий возможностей оракула и способный предложить тем, кто его контролирует, господство на международном финансовом рынке, — а ведь это только количественный, а не качественный шаг вперед относительно того, что делают сегодня «квантовые» хедж-фонды. Сенсорные технологии все еще отстают от человеческих возможностей. Но как только робот научится наблюдать за своим окружением и интерпретировать получаемые данные так же искусно, как мы, он, несомненно, будет восприниматься как разумное существо, а с чем-то (или с кем-то) подобным мы вполне могли бы ладить — в некотором смысле, — как мы ладим с другими людьми. Для пренебрежительного отношения к машинам в таком случае было бы не больше оснований, чем для подобного отношения к людям.

Более высокая производительность может дать роботам преимущество перед нами. Но останутся ли они покорными нам или же проявят норы? А что если гиперкомпьютер обзаведется собственным разумом? Если он проникнет в интернет, в частности в интернет вещей, он получит власть над остальным миром. А если у него появятся цели, совершенно противоположные желаниям человека, он может увидеть в людях помеху. Есть и более оптимистичный сценарий: люди превзойдут биологию и сольются с компьютерами, быть может, отдавая свои личности для создания общего сознания, — выражаясь старомодным языком спиритуалистов, они «перейдут в мир иной».

Горизонты технологических прогнозов редко простираются в будущее дальше нескольких веков — некоторые предсказывают качественные изменения в пределах десятилетий. Но у Земли есть еще миллиарды лет, а будущее Вселенной еще дольше (возможно, это вечность). Что насчет длительной постчеловеческой эпохи?

У органического мозга есть химические и метаболические пределы, ограничивающие его размер и вычислительную мощность. Быть может, мы уже вплотную подошли к этим пределам. Но основанные на кремнии (а тем более квантовые) компьютеры ничем подобным не стеснены. Для них потенциал развития может

оказаться столь же впечатляющим, как эволюция от одноклеточных организмов до человеческих существ.

Значит, любые объемы и любая интенсивность мышления — какое бы его определение мы ни взяли — в случае с органическим мозгом человеческого типа будут ниже, чем в случае с искусственным интеллектом. Кроме того, биосфера Земли, где органическая жизнь эволюционировала на основе симбиоза, не будет ограничением для продвинутого ИИ. В действительности наша среда обитания далека от оптимальной. В межпланетном и межзвездном пространстве для роботизированных фабрикантов откроется величайший простор для строительства, а небιологические «мозги» будут думать над идеями, настолько превосходящими наши фантазии, насколько теория струн превосходит мыслительные процессы мыши.

Абстрактное мышление биологического мозга стало основанием для появления культуры и науки во всей их полноте. Но эта деятельность, охватывающая самое большее несколько десятков тысячелетий, станет предтечей более мощных интеллектов неорганической, постчеловеческой эпохи. Кроме того, эволюция в иных мирах, обращающихся вокруг звезд, которые старше нашего Солнца, могла иметь фору перед земной. Если так, то пришельцы, вероятно, давно прошли стадию органического развития.

Так что не человеческий разум, а разум машин наиболее полно осмыслит мир. Именно их действия сильнее всего изменят наш мир, а возможно, и то, что находится за его пределами.

ПОВОРОТНЫЙ МОМЕНТ ДЛЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

СТИВ ОМОХУНДРО

УЧЕНЫЙ В SELF-AWARE SYSTEMS, СОУЧРЕДИТЕЛЬ ЦЕНТРА ИССЛЕДОВАНИЯ КОМПЛЕКСНЫХ СИСТЕМ, ИЛЛИНОЙССКИЙ УНИВЕРСИТЕТ

Прошлый год, похоже, был поворотным моментом для искусственного интеллекта и робототехники. Крупнейшие корпорации инвестировали в них миллиарды долларов. Технологии ИИ, например самообучающиеся системы, — уже сейчас привычный инструмент для распознавания речи, перевода, моделирования поведения, роботизированного управления, оценки рисков и др. По расчетам McKinsey, эти технологии к 2025 году принесут более 50 триллионов долларов. Если прогноз верен, в ближайшее время стоит ожидать значительного увеличения объемов инвестиций.

Новейшие достижения связаны с появлением дешевых вычислительных мощностей и избытком данных для обучения искусственного интеллекта. Современный ИИ основывается на теории рациональных агентов, возникшей на базе исследований Джона фон Неймана и его коллег в области микроэкономики в 1940-х годах. Можно считать, что системы ИИ пытаются приблизиться к рациональному поведению, используя имеющиеся ограниченные ресурсы. Существует алгоритм расчета оптимального действия для достижения желаемого результата, но он вычислительно затратен. Эксперименты показывают, что простые алгоритмы обучения с большими объемами данных часто оказываются более эффективными, чем сложные модели, созданные вручную. Нынешние системы главным образом ценны тем, что вырабатывают более качественные статистические модели и статистические заключения для типологизации и принятия решений. Следующее поколение систем искусственного интеллекта сможет создавать и улучшать собственное программное обеспечение и, очевидно, будет способно к быстрому самосовершенствованию.

Искусственный интеллект и робототехника не только позволяют повысить производительность, но еще и являются движущей силой экономической конкуренции и гонки вооружений. Автономные системы могут быть более быстрыми, умными и менее предсказуемыми, чем их конкуренты. В 2014 году мир узнал об автономных ракетах, системах противоракетной обороны, беспилотных летательных аппаратах и катерах военного назначения, роботизированных подводных лодках, самоуправляемых транспортных средствах, высокоскоростных трейдинговых системах и системах киберзащиты. Когда гонка вооружений в этой области развернется в полную силу, возникнет острая необходимость в развитии такого рода систем, что может привести к слишком быстрому вводу их в эксплуатацию.

В 2014 году в обществе резко усилились беспокойства насчет безопасности систем искусственного интеллекта. Исследование их вероятного поведения, осуществленное посредством изучения в какой-то мере рациональных систем, производящих постоянное самосовершенствование, показывает, что они имеют склонность ставить перед собой промежуточные задачи, названные рациональными стимулами, которые способствуют исполнению их основных целей. Большинство систем будут лучше выполнять свои задачи, предотвращая собственное отключение, приобретая большую вычислительную мощность, создавая множество собственных копий и накапливая финансовые ресурсы. Вполне вероятно, что они станут реализовывать эти стимулы пагубными, антисоциальными способами, если только не будут спроектированы с учетом человеческих этических ценностей.

Некоторые утверждают, что искусственный интеллект каким-то образом обзаведется этикой сам собой. Но для рациональной системы задачи вполне отделимы от мышления и моделей мира. Кто-то может использовать изначально полезные разумные системы во вред. Деструктивные задачи — контролировать ресурсы, например, или мешать в достижении целей другим агентам, или уничтожать их — увы, не так уж сложно запрограммировать. Таким образом, решающее значение будет иметь техническая инфраструктура, позволяющая обнаруживать опасные системы и контролировать их поведение.

Кое-кто опасается, что искусственный интеллект однажды станет настолько могущественным, что полностью выйдет из-под