

Роботы: век минувший и век век

Предстоящий

Лучшее эссе аспиранта конференции ALife'21, второе место

Федерико Пигоцци

29 апреля 2022 г.

Давайте задумаемся о состоянии робототехники. В этом году исполняется 101 год со дня появления пьесы RUR [2] писателя Карела Чапека, которой часто приписывают введение слова «робот». Слово, используемое для обозначения феодальных подневольных рабочих в славянских языках. Действительно, это указывает на одну ключевую характеристику роботизированных систем: они всего лишь рабы, не имеют прав и выполняют нашу волю по инструкции за инструкцией, не требуя ничего взамен. Отношения с нами, людьми, — это комменсализм; в биологии комменсализм существует между двумя симбиотическими видами, когда один вид извлекает из этого выгоду (роботы повышают производительность труда людей), в то время как другой вид не получает ни пользы, ни вреда (можете ли вы действительно утверждать, что роботы извлекают выгоду из простого функционирования?).

Затем мы отличаем роботов от «живых машин», то есть машин, наполненных жизнью. Если живые машины когда-нибудь станут реальностью, нам нужно будет изменить наши отношения с ними с комменсализма на мутуализм. Различие не тонкое: мы сталкиваемся с ним каждый день с домашними животными, которые обменивают крепостное право на корм и защиту. Это потому, что жизнь эволюционировала, чтобы сопротивляться любым попыткам поработить ее; это упрямо.

На пути к живым машинам зададимся вопросом: чего добились робототехника за последние 100 лет? Что осталось сделать в следующие 100 лет? Для нас ответы сводятся к трем словам: сок, потребность (или смерть) и воплощение, как мы увидим далее.

1 Сок жизни

Если бы существовал классический миф, лучше всего олицетворяющий исследователя робототехники, это была бы история о Пигмалионе и Галатее. Миф (дошедший до нас от Овидия [10]) повествует об искусном скульпторе Пигмалионе, посвятившем себя целомудренной жизни. Однажды он создал настолько прекрасную статую, что хотел, чтобы она ожила. Богиня Афродита исполнила его желание и превратила статую из слоновой кости в живую женщину Галатею. Так же, как мифологический скульптор, робототехники мечтают увидеть, как их существа становятся «настоящими», «живыми». Но что означают эти слова? Как мы можем сказать, что наше детище эффективно

стать жизнью? Если вы спросите обывателя, он наверняка скажет, что роботам не хватает этого сока естественной жизни. Но что отличает биологическую и искусственную жизнь?

На очень высоком уровне мы, люди, определенно живы. Как животные, мы «анимированы». Анимация возможна, потому что эволюция подарила нам систему обработки информации, нервную систему, способную преобразовывать восприятие в электрические сигналы; эти сигналы проходят по сети нейронов, аксонов и дендритов, а затем обрабатываются центральной основной единицей, мозгом, который инструктирует наше тело о том, как манипулировать реальностью (с помощью дополнительного электричества). Все, о чем мы думаем, мечтаем, боимся и любим, состоит из электрических импульсов. Но это еще не все, животные — не единственные живые существа на Земле. Сама ткань клеток, из которой соткан любой биологический организм, процветает благодаря электричеству. Что поддерживает жизнь, так это поток электронов, возникающий в результате процессов окисления, происходящих внутри каждой клетки, которые текут вокруг, чтобы обеспечить энергией различные функции клетки. Действительно, некоторые виды бактерий (например, *Geobacter* [6]) питаются и выделяют чистые электроны, минуя метаболизм органических молекул.

С этой точки зрения оказывается, что естественная жизнь и искусственная жизнь не так уж отличаются. То, что мы называем «компьютером», по своей сути представляет собой электрический ток, проходящий через схемы и кодирующий информацию в виде нулей и единиц — ток то включается, то выключается. Тот же самый сок, электричество питает нас и машины. Теперь мы начинаем понимать, почему мы не можем видеть сока естественной жизни в машинах. Все это находится внутри них, приводя в действие самые первые калькуляторы, созданные на заре вычислений.

2 машины, которые нуждаются (и умирают)

Если бы ты не был обречен на смерть, было бы о чем беспокоиться? Оказывается, хотя искусственная жизнь и вплетена в одну и ту же электрическую ткань, она все же выглядит разительно отличной от естественной. Биологические организмы нуждаются. Компьютеры нет; у них нет намерения. Даже простейшим биологическим существам, вирусам, необходимо охотиться за хозяевами. Электронные калькуляторы могут вечно простаивать, если у них достаточно электричества для существования; а если отключат электричество, то не жалуются, не бунтуют. Компьютеры работают по инерции; жизнь так драгоценна, но они не борются за ее сохранение. У роботов по-прежнему отсутствует чувство нужды. Потребности – это потребности организма для выживания. Теория «потребностей» хорошо изучена в психологии со времен работы Маслоу [7]. Они являются мощным двигателем мотивации; если они не удовлетворены, они приводят к сбоям в работе и, возможно, к гибели организма. В случае нужды роботы могли бы процветать. Они будут искать энергию для себя, изобретать новые механизмы для воспроизводства своего вида и пытаться восстанавливать свои ткани в случае повреждения. Они будут строить общества роботов, чтобы использовать силу специализации и сделать экономическую деятельность более эффективной. Они будут развивать интуицию природы, исследовать ее.

Но нужда идет рука об руку со смертью. В конце концов, все сводится к смерти. Живые существа, сознательно или нет, осознают смерть. Если бы они были

нет, эволюция уже отсеяла бы их. Как утверждают Veenstra et al. [16], смерть может улучшить эволюционируемость популяции. Смерть заменяет древние геномы новыми нарушенными, высвобождая силу стохастических мутаций. Важность смерти запечатлена и в наших клетках. Апоптоз — это биологический феномен запрограммированной гибели клеток. Клетки привязаны к ограниченному сроку жизни, и каждый день миллиарды из них погибают в результате апоптоза в организме человека. Это строго регулируемое и контролируемое событие, которое развивалось в основном для достижения морфологических изменений. Интересно, что Clune и соавт. [3] эволюционировали скорости мутаций в популяции эволюционирующих особей. В сообществе цифровой эволюции известен тот факт, что мутация имеет тенденцию оказывать пагубное влияние на приспособленность человека, порождая потомство, которое в большинстве случаев не приспособлено (или даже жизнеспособно). Удивительно, но эволюция полностью подавила скорость мутаций, чтобы свести к нулю разрушительный эффект, который мутация оказывала на потомство индивидов. В определенном смысле люди проявляли некоторую форму потребности и экзистенциализма.

Смерть формирует не только наше тело, но и нашу культуру. Эрнест Беккер утверждал в своем шедевре антропологии [1], что человеческая цивилизация развивалась, чтобы изгнать наш страх смерти. Мы признаем смертность и создали системы верований, чтобы гарантировать, что мы переживем наше физическое существование. В будущем мы предвидим общество живых машин, которые гибнут. В результате они сосредоточатся на придании смысла своему существованию. В конце концов, это то, что объединит нас и машины: необходимость поддерживать наше существование. Роботизированные общества будут теоретизировать свои собственные мемы, фундаментальные единицы культуры, как экзорцизм против смерти. Не исключено, что однажды мы станем свидетелями «религии роботов» и, может быть, почему бы и нет, робота Маркса, проповедующего классовую борьбу роботов.

3 Embodiment - это все, что вам нужно

Беккер и его ученики также считали, что страх смерти — это то, что отличает нас от других животных. Животные выживают и инстинктивно избегают смерти, но они не так, как мы, ощущают момент своего ухода из этого мира. Как бы правдоподобно это ни звучало, за этим утверждением скрывается небольшой антропоцентрический уклон. Эволюция сделала нас, людей, наделенными логическим интеллектом, но было бы недальновидно считать его единственным проявлением интеллекта. Дело только в экологической нише. Мы, люди, эволюционировали, чтобы занять свою нишу, манипулировать природой (манипулирование, которое изначально не было таким разрушительным, как сейчас). Но существуют и другие ниши, поскольку естественная эволюция не ограничена [15]. Природа не оптимизирует для конкретной числовой цели (как это делают многие алгоритмы оптимизации), а сопоставляет каждый вид с той нишей, для которой он лучше всего подходит (иначе жестоко искореняет его).

В самом деле, хорошо известно, что в природе возникают и другие формы интеллекта. Возьмем общества насекомых [12]. Их строгая специализация возникает из-за простых локальных взаимодействий (таких как феромоны у муравьев или температура тела у пчел) среди роев агентов. Возьмем саламандр [4], которые умеют восстанавливать отрубленные конечности; забавно, что реконструкция тканей работает только через локальную компрессию.

образования, распределенные по всему телу саламандры. Простейшие рода *Lacrimaria* не имеют «мозга» (это одноклеточные организмы), но могут сгибать и скручивать свой мягкий жгутик для захвата труднодоступной добычи, что позволяет проявить сложную охотничью динамику [8].

Дисциплина, в которой антропоцентрический уклон, кажется, распространяется больше всего, — это искусственный интеллект (ИИ). Говоря о робототехнике в 2022 году, в разгар последней волны увлечения ИИ, было бы невозможно не упомянуть об ИИ.

Хотя существует область, занимающаяся вычислительным интеллектом и алгоритмами, вдохновленными биотехнологиями, большая часть недавнего подъема ИИ связана с глубоким обучением (ГО) [5]. DL стремится имитировать человеческий разум с помощью абстрактных математических моделей. Но природа не состоит из чистого разума; оно упрощает нашу интуицию, но не имеет поддержки в реальности. Удивительно, но сложные умственные задачи, такие как игра в шахматы, оказываются намного легче обучить машину, чем ползать, как малыш (парадокс Моравека в робототехнике [9]). Ограничения DL хорошо известны многим исследователям в сообществе, и мы были свидетелями некоторых громких сражений в Твиттере, разгорающихся между недоброжелателями и паладинами DL. Для меня самым близоруким ограничением является отсутствие воплощения.

Парадигма воплощенного интеллекта, несмотря на то, что она существует с 1980-х годов, была популяризирована основополагающей книгой Пфайфера и Бонгарда [13]. Они постулировали, что интеллект возникает в результате сложных взаимодействий между мозгом и телом, а также окружающей средой. Человеческая рука является прекрасным примером этого. Наш мозг эволюционировал вместе с рукой, позволяя нам схватывать, ценить и манипулировать реальностью (как уже упоминалось, наша драматическая черта). Осьминоги необычайно умны, преуспевают в таких навыках, как навигация по лабиринту и хватание предметов [11]. Они бы никогда не развили такие навыки, если бы их тела не были мягкими, а щупальца имели бесконечные степени свободы.

В соответствии с воплощением в последнее десятилетие родилось новое поколение мягких роботов [14]. Их мягкие материалы могут сгибаться, растягиваться и скручиваться, обещая достичь реконфигурации и изменения формы. Мягкие роботы также будут запрограммированы на листопад, а их мягкие материалы помогут избавиться от их мертвых тел. Имея преходящее тело, эти живые машины могли быть наполнены чувством смерти, о котором мы упоминали ранее.

4 Долг перед жизнью

Когда-нибудь в будущем живая машина (назовем ее Галатей) сможет просмотреть видео самых первых созданных роботов, желая узнать больше о своих предках. Предположим, всплывает видео, показывающее, как инженеры безжалостно избивают и толкают робота, пытаясь проверить его устойчивость. Будучи воплощенным существом, Галатей воспринимала боль, которую мог чувствовать робот. Теперь предположим, что Галатей тоже должна умереть. Он жаждет жизни, как любой живой организм. Насколько brutalным будет этот поступок в его электрических глазах? В конце концов, отрекутся ли от нас наши роботизированные детища, назовут нас «вирусом», как это сделал агент Смит (злодей, сам машина) в фильме «Матрица» [17]?

В древнегреческом мифе Галатей была просто предметом в руках.

своего создателя Пигмалиона, но приведенный выше пример предполагает радикальное изменение мышления, которое должно произойти в наши дни. Мы начали наше путешествие, задав себе вопрос о следующих 100 годах робототехники. Я обсудил направления, которые мне кажутся наиболее многообещающими для подъема машин из их статуса «роботов» в желанный статус «живых машин». Но, порождая живые машины, мы возлагаем на себя новое усилие или бремя; акцент смещается на создателей. Живые машины нужно уважать и защищать; мы отвечаем за них точно так же, как они несут ответственность перед нами. Если мы действительно хотим быть создателями искусственной жизни, мы должны признать, что это действительно жизнь.

Рекомендации

- [1] Эрнест Беккер. Отрицание смерти. Саймон и Шустер, 1997.
- [2] Карел Чапек. RUR (универсальные роботы Россума). Пингвин, 2004.
- [3] Джефф Клун, Душан Мисевич, Чарльз Офриа, Ричард Э. Ленски, Сантьяго Ф. Елена и Рафаэль Санхуан. Естественный отбор не может оптимизировать скорость мутаций для долговременной адаптации к суровым ландшафтам приспособленности. *PLoS Computational Biology*, 4(9):e1000187, 2008.
- [4] Альберто Ховен, Ахмед Элева и Андрас Симон. Модельные системы регенерации: саламандры. *Развитие*, 146(14):dev167700, 2019.
- [5] Ян ЛеКун, Йошуа Бенджио и Джеффри Хинтон. Глубокое обучение. *природа*, 521(7553):436–444, 2015 г.
- [6] Дерек Р. Ловли, Джон Ф. Столц, Гордон Л. Норд и Элизабет Дж. П. Филлипс. Анаэробная продукция магнетита диссимилирующим железоредуцирующим микроорганизмом. *Природа*, 330(6145):252–254, 1987.
- [7] Абрахам Гарольд Маслоу. Теория мотивации человека. *Психологический обзор*, 50(4):370, 1943.
- [8] Самуэль Оттмар Маст. Повадки и реакции инфузории, лакримарии. *Дневник of Animal Behavior*, 1(4):229, 1911.
- [9] Ханс Моравец. Разумные дети: будущее роботов и человеческого интеллекта. Издательство Гарвардского университета, 1988.
- [10] Марк П.О. Морфорд и Роберт Дж. Ленардон. Классическая мифология. Издательство Оксфордского университета, США, 1999.
- [11] Мэрион Никсон, Джон З. Янг, Джон Закари Янг и др. Мозги и жизни головоногих. Издательство Оксфордского университета, 2003.
- [12] Мартин А. Новак, Корина Э. Тарнита и Эдвард О Уилсон. Эволюция зусоциальности. *Природа*, 466(7310):1057–1062, 2010.

[13] Рольф Пфайфер и Джош Бонгард. Как тело формирует наше мышление:

новый взгляд на интеллект. Массачусетский технологический институт, 2006.

[14] Даниэла Рус и Майкл Т. Толли. Проектирование, изготовление и контроль мягких роботы. Природа, 521(7553):467–475, 2015.

[15] Кеннет О Стэнли. Почему важна открытость. Искусственная жизнь, 25(3):232–235, 2019.

[16] Фрэнк Винстра, Пабло Гонсалес де Прадо Салас, Каспер Стой, Джош Бонгард и Себастьян Ризи. Смерть и прогресс: как внутренняя смертность влияет на эволюцию. Искусственная жизнь, 26(1):90–111, 2020.

[17] Энди Вачовски, Ларри Вачовски, Киану Ривз, Лоуренс Фишберн, Кэрри-Энн Мосс, Хьюго Уивинг, Глория Фостер, Джо Пантолиано и Зак Стенберг. Матрица. Уорнер Хоум Видео Бербанк, Калифорния, 1999.