



А.Е. Сериков

## ПРОБЛЕМА АВТОМАТИЧЕСКОГО ПРОИЗВОДСТВА НАРРАТИВОВ

(Самарский университет)

Поскольку, по мнению многих специалистов, тест Тьюринга проверяет не столько наличие сильного интеллекта, сколько способность программы и ее авторов ввести в заблуждение жюри, в последнее время было предложено несколько новых тестов. В частности, Селмер Бринджорд с соавторами предложили в 2001 г. тест, названный ими тестом Лавлейс [1]. Общая идея заключалась в том, что искусственный агент А, созданный человеком Н, должен рассматриваться как интеллектуальный агент человеческого уровня, если он может производить действительно творческие результаты, и если Н не может объяснить, как А это сделал. Позже Марк Ридл утверждал, что такой тест не сможет пройти ни одна программа, потому что, если у Н будет достаточно ресурсов для построения А и достаточно времени для анализа его работы, Н сможет объяснить результаты А. Поэтому он предложил более простую версию теста — тест Лавлейс 2.0. Основная идея заключалась в том, чтобы предложить искусственному агенту создать произведение искусства, отвечающее некоторым требованиям, предъявляемым оценщиком-человеком [7].

В 2019 г. вышла книга Маркуса дю Сотоя «Код креативности», в которой он рассказывает обо всех успешных примерах компьютерного творчества, существовавших на момент написания книги. Один из них — портрет Эдмонда Белами, созданный с помощью алгоритма и проданный на аукционе Christie за 432500 долларов в 2018 году. Еще одним впечатляющим примером творческих программ является алгоритм джазовой импровизации, созданный Франсуа Паше и известный как Continuator, поскольку он может продолжать музыку, начатую живым музыкантом в том же стиле, или, в другой версии, он может сопровождать живого музыканта, угадывая нужный аккорд. Создается впечатление, пишет дю Сотой, что Continuator прошел тест Лавлейс. Есть еще более убедительный пример успеха искусственного интеллекта в создании музыки: в 2016 году алгоритм под названием AIVA стал первой машиной, получившей звание композитора от французской профессиональной ассоциации Société des Auteurs, Compositeurs et Éditeurs de Musique [3].

Однако, как мне кажется, дю Сотой не смог привести ни одного убедительного примера компьютерной программы, способной сочинять интересные истории. Этот факт констатируют и авторы идеи теста Лавлейс. В 2017 году Ридл отметил, что «компьютеры по-прежнему не могут надежно создавать и рассказывать новые истории или понимать истории, рассказанные людьми. Когда компьютеры рассказывают истории с помощью электронной книги или компьютерной игры, они просто пережевывают что-то, написанное человеком» [6]. В 2020 году Селмер Бринджорд публично обсуждал эту тему с Робертом



Марксом и они также пришли к выводу, что на момент обсуждения ни один ИИ не прошел тест Лавлейс в области сочинения историй [5].

В 2009 году Натанаэль Чамберс и Дэн Джурафски создали самообучающуюся систему, извлекающую повествовательные схемы из новостных статей. Схема повествования – это модель событий, которая объединяет несколько типизированных цепочек повествовательных событий вместе. Событие – это глагол вместе с совокупностью его функционально-семантических аргументов. Типизированная повествовательная цепочка событий – это последовательность событий, связанных с ролью определенного типа. Тип автоматически создается как набор лексических единиц, связанных с событиями в цепочке. Модель не предполагает, что набор ролей известен заранее, и она изучает роли одновременно с кластеризацией глаголов. Например, система извлекла из новостной статьи три цепочки событий. Одна цепочка включает глаголы «арестовать» и «обвинять», связанные с типом, представленным существительными «полиция» и «агент», вторая цепочка – глаголами «осуждать» и «приговор», связанными с существительными «судья» и «присяжные», а третья цепочка – глагол «отвечать на обвинение», связанный с существительными «преступник» и «подозреваемый». Путем объединения трех цепочек в одну повествовательную схему система создаст результирующий сюжет в форме: «Полиция (агент) арестовывает и предъявляет обвинение преступнику (подозреваемому); преступник (подозреваемый) признает себя виновным (невиновным); судья (присяжные) осуждает и приговаривает преступника (подозреваемого)» [2].

Брент Харрисон, Кристофер Пурди и Марк Ридл в 2017 году создали систему, автоматически обучающуюся на основе кратких киносюжетов из Википедии и затем генерирующую фабулы на любую заданную тему. В этой модели каждое событие состоит из следующих элементов: субъект, глагол, объект, модификатор. Каждое предложение в исходном корпусе рассказов сокращено до последовательности из 4 слов этой формы. Предложения с более чем одним глаголом или несколькими подлежащими разбиты на несколько событий. Любой новый сюжет, созданный этой системой, также состоит из таких предложений. События – это предложения из 4 слов. Центральным элементом каждого события является глагол, а остальные три слова можно понимать как его аргументы [4].

Автоматический генератор нарративов, представленный в 2019 г. Лили Яо с соавторами, работает на основе еще более простой модели. Система использует существующие корпуса рассказов для автоматического извлечения одного ключевого слова из каждого предложения и строит сюжетную линию как последовательность ключевых слов. То есть событие обозначается отдельным словом. Получив любое слово в качестве заголовка, система использует изученные вероятности для создания новой последовательности слов, а затем на основе полученного таким способом нового сюжета, выстраивает историю, конструируя вокруг ключевых слов наиболее вероятные предложения [9].

Таким образом, в современных работах по автоматическому пониманию и генерации нарративов событие как часть истории часто репрезентируется гла-



голом вместе с его минимальным набором аргументов или просто словом. Соответственно, сюжетная линия обычно определяется как последовательность глаголов или других ключевых слов. Этот тип концептуализации событий предоставляет очень мало возможностей для извлечения и репрезентации здравого смысла и других контекстуальных знаний, что делает автоматически генерируемые истории довольно примитивными.

Итак, сочинение историй оказалось самым трудным из всех творческих заданий, возможных в рамках теста Лавлейс или Лавлейс 2.0. Если допустить, что в области музыки, живописи или абстрактной поэзии тест Лавлейс пройден наилучшими современными алгоритмами, то в области сочинения интересных и осмысленных с точки зрения фабулы рассказов искусственному интеллекту еще очень далеко до людей. Проблема заключается в том, что самообучающиеся нейронные сети пока еще не умеют извлекать из открытых источников и использовать знания о человеческом поведении, основанные на здравом смысле. Лучшие из самообучающихся сетей пока что могут извлекать информацию на уровне отдельных слов, словосочетаний или предложений, используя методы лексического и грамматического анализа. А для извлечения знаний о поведении требуется анализ более высокого уровня, который можно назвать анализом грамматики поведения. Поэтому дальнейшее развитие автоматических систем конструирования нарративов может быть связано с разработкой такой грамматики и ее внедрением в алгоритмы текстового анализа.

### Литература

1. Bringsjord S., Bello P. & Ferrucci D.: Creativity, the Turing Test, and the (Better) Lovelace Test. *Minds and Machines* 11, 3–27 (2001) URL: <https://doi.org/10.1023/A:1011206622741> (дата обращения: 10.04.21).
2. Chambers N., Jurafsky D.: Unsupervised Learning of Narrative Schemas and their Participants. In: *Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP*, pp. 602–610. Association for Computational Linguistics, Suntec, Singapore (2009) URL: <https://www.aclweb.org/anthology/P09-1068.pdf>. (дата обращения: 10.04.21).
3. Du Sautoy M.: *The Creativity Code. Art and Innovation in the Age of AI*. 2nd edn. Harvard University Press, Cambridge, MA (2020). URL: <https://doi.org/10.4159/9780674240407> (дата обращения: 10.04.21).
4. Harrison B., Purdy, C., Riedl, M.: Toward Automated Story Generation with Markov Chain Monte Carlo Methods and Deep Neural Networks. In: *AAAI Publications, Thirteenth Artificial Intelligence and Interactive Digital Entertainment Conference* (2017) URL: <https://aaai.org/ocs/index.php/AIIDE/AIIDE17/paper/view/15908/15241>. (дата обращения: 10.04.21).
5. Mind Matters,: Thinking machines? Has the Lovelace test been passed? (2020) <https://mindmatters.ai/2020/04/thinking-machines-has-the-lovelace-test-been-passed>. (дата обращения: 10.04.21).



6. Riedl M.: Computational Narrative Intelligence: Past, Present, and Future (2017) URL: <https://mark-riedl.medium.com/computational-narrative-intelligence-past-present-and-future-99e58cf25ffa>. (дата обращения: 10.04.21).

7. Riedl M.: The Lovelace 2.0 Test of Artificial Intelligence and Creativity. In: Proceedings of the AAAI Workshop: Beyond the Turing Test, Austin, Texas (2015) URL: <https://arxiv.org/pdf/1410.6142v3.pdf>. (дата обращения: 10.04.21).

8. Yao L., Peng N., Weischedel R., Knight K., Zhao D., Rui Yan R.: Plan-and-Write: Towards Better Automatic Storytelling. In: Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence, pp. 7378–7385. AAAI Press, Palo Alto, CA (2019) URL: <https://www.aaai.org/ojs/index.php/AAAI/article/download/4726/4604>. (дата обращения: 10.04.21).

9. Yao L., Peng N., Weischedel R., Knight K., Zhao D., Yan R.: Plan-and-Write: Towards Better Automatic Storytelling. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33 No. 01: AAAI-19, IAAI-19, EAAI-20, pp. 7378–7385. AAAI Press, Palo Alto, CA (2019) URL: <https://www.aaai.org/ojs/index.php/AAAI/article/download/4726/4604>. (дата обращения: 10.04.21).

А.П. Слащинуна

## О ВАЖНОСТИ БУДДИЙСКИХ ЦЕННОСТЕЙ ДЛЯ ЭТИКИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

(Самарский университет)

Этические проблемы, связанные с созданием искусственного интеллекта (далее ИИ), продолжают занимать умы не только философов-трансгуманистов, но все чаще становятся предметом опасений общественности. Несомненно, трансгуманизация как результат научного прогресса является закономерным путем развития общества, ведь как отмечал французский философ Ги Дебор: «борьба между традицией и прогрессом является основным принципом внутреннего развития культуры исторического общества и продолжается лишь благодаря постоянным победам прогресса» [1, с. 140].

Негативное отношение к ИИ во многом продиктовано эффектом «зловещей долины», характеризующийся неприязнью и отвращением, вызываемым у людей объектами, выглядящими как человек. Однако опасения вызывает не только внешний вид андроидов, копирующих человека, но и этические вопросы, начиная от связанных с проблемой «корабля Тесея» – «будет ли человек, которому в результате многочисленных медицинских имплантаций заменили 99 процентов тела, считаться машиной?» и заканчивая вопросом – «будет ли у ИИ сознание и свободная воля?». На наш взгляд в вопросах этики искусственного интеллекта есть другой, гораздо более насущный вопрос, связанный с разработкой ИИ – «какой должна быть этика ИИ?». Технологии, созданные чело-