

Прерванный диалог: как и почему тексты эпохи нейросетей порывают с традицией

Б. В. Орехов
(НИУ ВШЭ, ИРЛИ РАН)

Заметный в последние годы технологический прогресс в области компьютерного порождения текстов заставил по-новому взглянуть на проблематику природы нарратива и в целом литературы. Порождение художественного текста — одна из стандартных задач, на которых проверяется успешность решений в области искусственного интеллекта (ИИ). Часть этих задач успешно решена (например, способность играть в шахматы лучше человека), но с порождением текста художественных жанров пока есть заметные проблемы (стандартизированные жанры в целом компьютером освоены).

Успешность решения задач в ИИ-инженерии определяется исчислимыми критериями, но в отличие от игры в шахматы (или любую другую игру с нулевой суммой) такие критерии в области текстопорождения определить непросто. Специалистов в основном сошлись на оценке мимикрии, то есть степени сходства компьютерного и человеческого текста.

Благодаря широкой доступности текстовых генераторов у филологов копится опыт в оценке текстовой продукции искусственного происхождения.

«Детские болезни», которые сопровождали такие тексты в начале развития области, постепенно уходят. К ним можно было отнести неграмматичность, неестественность синтаксиса, несловарные слова, слишком быстрое переключение контекста.

Современные тексты, порождаемые нейросетями на основе архитектуры «Трансформер», грамматичны, состоят из словарных слов, достаточно долго развивают один контекст. Но они все еще по множеству параметров далеки от текстов, которые квалифицированный читатель готов признать художественными, и тем более — приносящими эстетическое удовольствие.

Еще совсем недавно текстовые генераторы не были способны порождать связную историю. Теперь такая задача им по силам, более того, по запросу пользователя история может быть модифицирована под схему того или иного жанра.

И все же повествовательный текст сильно отличается от «литературного», и в докладе я планирую показать эту разницу на неочевидном материале: в отличие от привычных нам художественных текстов естественного происхождения, в текстах ИИ отсутствует момент литературной игры, взаимодействия с традицией и трансформации под ее влиянием.

В профессиональной среде отсутствует рефлексия над этой проблемой, за чем, по моему мнению, стоит целый комплекс затруднений, связанный в том числе со способами преподавания литературы и выстраивания школьной программы. Способ представления литературы в учебном поле как череды вершин не подразумевает идеи связи между ключевыми текстами традиции, поэтому и разработчики нейронных сетей не ставят вопроса о включении этого элемента в порождающую программу.