

УДК 81`42

Елена Сергеевна Солнцева, к. филол. н.
Санкт-Петербургский государственный
университет, Санкт-Петербург, Россия
э-почта: e.solntseva@spbu.ru

Elena Sergeevna Solntseva, Cand. of Sc. (Philology)
Saint Petersburg State University, St. Petersburg,
Russia
e-mail: e.solntseva@spbu.ru

КОГЕРЕНТНОСТЬ ЗАПРОСОВ И ОТВЕТОВ В РАБОТЕ НЕЙРОСЕТИ

Данный доклад посвящен актуальному вопросу взаимодействия человека и искусственного интеллекта, а именно изучению степени согласованности запроса и ответа при работе с нейросетью.

Ключевые слова: когерентность запроса и ответа; искусственный интеллект; MyDoc-Chat GPT-4o от Open AI; ответ нейронной сети; промпт для чат-бота.

PROMPT-RESPONSE COHERENCE IN NEURAL NETWORK

The report touches upon a breakthrough issue of interaction between a human and an artificial intelligence, and in particular regards coherence between a prompt and a neural network response.

Key words: coherence of the prompt and response; artificial intelligence; MyDoc-Chat GPT-4o by OpenAI; neural network response; prompt for a chatbot.

Искусственный интеллект (ИИ) стремительно завоевывает позиции на рынке переводческих, образовательных и даже креативных услуг, становясь сегодня одной из самых актуальных областей исследования. Примечательно и то, что эта область является в высшей степени междисциплинарной, обеспечивая устойчивую связь достижений самых разных наук и технологий, в частности, биофизики, информатики, математики, нейрофизиологии, психологии, лингвистики и др. Данное исследование посвящено особенностям коммуникации с искусственным интеллектом, а именно исследованию когерентных отношений между запросом (промптом) и ответом нейронной сети.

Нейронные сети, на базе которых работает искусственный интеллект, представляют собой новый особый способ опосредованной коммуникации. Системы искусственного интеллекта имитируют когнитивные функции человека. В процессе такого общения происходит регулярное обращение к базам данных, подобно тому, как человек в ходе коммуникации, обмена информацией, непрерывно обращается к различным участкам мозга, к тем областям, где хранятся знания о различных предметах, явлениях и процессах. В связи с этим коммуникация между ИИ и человеком становится максимально приближенной к общению двух людей.

Работа искусственного интеллекта основана на работе нейронной сети. В головном мозге человека каждый отдельный нейрон может строить тысячи связей с другими нейронами, и таким образом они образуют нейронную сеть. Чем больше сигналов посылается между двумя нейронами, тем сильнее становится связь, и поэтому с каждым новым опытом и каждым памятным событием или фактом мозг немного перестраивает свою физическую структуру. Основой работоспособности нейронной сети становится связь между различными этапами передачи сигнала.

В исследованиях текста и дискурса эта связь между объектами внешнего мира и их воплощением в высказываниях разной продолжительности получила название «когерентность». Под когерентностью дискурса понимается когнитивно обусловленная контекстуальная связанность, соположенность компонентов внутри одного дискурса. Несколько лингвистических, текстовых, а также несколько нелингвистических элементов (изображений, звука, гипертекста и др.) сочетаются в пределах одного дискурса, если соположены в одном контексте, то есть объединены, в частности, одной темой, или автором, или изданием, или любым другим контекстом, для которого соположение именно данных определенных элементов с когнитивной точки зрения релевантно и оправдано. В рамках настоящего исследования изучается когерентность дискурса между ИИ и человеком как структурная и семантическая согласованность между запросом (промптом) к нейросети и ее ответом.

На сегодняшний день на рынке представлены нейросетевые искусственные интеллекты GPT-4o от Open AI, Grok от Илона Маска, Gemini от Google, GigaChat при поддержке Сбера, YaGPT3 от Яндекса, DeepSeek, разработанная в Китае, и ряд

других. Для работы с ними в коммуникацию встраивается виртуальный помощник, собеседник, чат-бот, который передает запрос от человека к ИИ. В качестве чат-ботов популярность приобрели ruGPT, ChatGPT, MyDoc-Chat и другие.

В качестве материала данного исследования чат-боту MyDoc-Chat GPT-4o от Open AI задавались различные запросы на составление текстов. Целью стало проследить семантические и структурные характеристики ответа и сравнить их соответствие/несоответствие первоначальному запросу. Объем материала исследования составил 50 запросов и 50 ответов на них. Предполагаемой аудиторией выступали студенты, преподаватели или переводчики. Основным методом исследования стал метод контент-анализа.

Запросы (промпты) включали, в частности, такие формулировки как *«Составь тест по английскому языку для студентов российского вуза 1 курса, включи 4 части...»*; *«Подготовь презентацию по тексту...»*; *«Напиши стайлгайд для перевода сайтов на русский язык»* и др.

Ответы на запросы характеризовались некоторыми особенностями. С точки зрения структуры, полученные ответы на запрос обладали четкой и логичной последовательностью: присутствовало разделение на параграфы, тематические пункты и подпункты, логические смысловые части отделялись друг от друга визуально (графически) и семантически. При этом структура отдельно в запросе не задавалась. Выбор количества и содержания частей, релевантных для рассмотрения заданной темы, определялся автоматически на основе типичной структуры заданного жанром содержания (слайды для презентации, вопросы для теста и пр.).

С другой стороны, с точки зрения семантического наполнения, ответы явно не соответствовали запросам в том, что касается глубины и объема освещаемого материала. Ряд формулировок был поверхностным и наиболее общим, использовались слова и выражения, которые без труда можно поставить в любой другой схожий контекст, то есть слова и выражения, неспецифические для заданной темы. Так, для запроса *«напиши стайлгайд...»* было рекомендовано, среди прочего, *«определить общий тон общения (формальный, неформальный, дружелюбный, профессиональный и т. д.) и соблюдать его»*.

Улучшения семантического качества ответов на запросы наблюдались при дроблении запроса на более мелкие компоненты. Так, например, если вместо общего запроса на создание стайлгайда предлагался более конкретный запрос: *«опиши, каким должен быть стиль речи для сайта компании»*, то ответ становился более семантически насыщенным. В частности, на наиболее узкий специфический запрос *«приведи конкретные слова на русском языке, которые можно использовать на сайте компании»* анализируемая нейронная сеть выдала ответ в виде перечня подходящих слов, сгруппированных структурно в зависимости от типичных разделов стандартного сайта компании. Например, раздел *«Главная страница»*: *«Добро пожаловать»*, *«Наша миссия»*, *«Узнать больше»* и т. п.

Подводя итог проведенного исследования, можно отметить следующую тенденцию в когерентных отношениях «запрос-ответ» для MyDoc-Chat GPT-4o от Open AI: чем специфичнее и конкретнее формулируемый запрос, тем более глубоким и насыщенным становится ответ. Для получения полноценного семантически наполненного ответа на запрос результат будет качественнее с точки зрения семантики, если формулировать запрос не на широкую задачу, в целом, глобально, а на более узкую, специфическую, локальную.

Перспективой настоящего исследования может стать дальнейшее изучение соотношения семантики и структуры запросов к нейросети и ответов на них, а также рассмотрение подобной проблематики на материале разноструктурных языков.