

**Т. В. Бусел**

## ИСКУССТВО ПЕРЕВОДА В ЭПОХУ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Перевод как искусство был и остается актуальной интеллектуальной и творческой задачей. До наступления эпохи искусственного интеллекта слово “переводчик” ассоциировалось исключительно с человеком. Перевод научных статей, художественных произведений и фильмов, технических и юридических документов выполняли профессиональные переводчики. Однако разработка и

внедрение технологий искусственного интеллекта оказали огромное влияние на индустрию переводов, в частности, на машинный перевод. Возможности машинного перевода тесно связаны с совершенствованием моделирования процессов понимания и генерации текстов на естественном языке, что традиционно относится к классу проблем искусственного интеллекта.

Нейронный машинный перевод очень востребован в современном обществе, поскольку разработка качественных и высокопроизводительных систем способствует разрешению важнейших социально-коммуникативных задач, ежедневно ими пользуются более полумиллиарда человек. Согласно отчету ассоциации Research Industry Network росту мирового рынка машинного перевода способствует целый ряд факторов:

- 1) увеличение спроса на переводческие услуги, они играют ключевую роль в преодолении культурных и языковых барьеров, а также в повышении эффективности работы международных компаний и организаций;

- 2) растущий объем различных видов цифрового контента: текстового, игрового, видео, аудио, графического и пользовательского;

- 3) совершенствование технологий машинного обучения (так называемого «искусственного интеллекта») и обработки информации на естественном языке (Natural Language Processing – NLP);

- 4) всё большее распространение облачных услуг машинного перевода (DeepL, Google Cloud Translation API, Alibaba Cloud Machine Translation, IBM Watson Language Translator, SDL Language Cloud Translation Management).

Развитие направления машинного перевода представляет большой научный интерес, что обусловлено комплексностью и сложностью компьютерного моделирования данного процесса. Как вид языковой деятельности перевод затрагивает все уровни языка – от распознавания графем (и фонем при переводе устной речи) до передачи смысла высказывания и текста. Кроме того, для перевода характерна обратная связь и возможность сразу проверить теоретическую гипотезу об устройстве тех или иных языковых уровней и эффективности предлагаемых алгоритмов. Эта специфическая черта перевода привлекает внимание отечественных и зарубежных ученых, в результате чего продолжают возникать все новые теории автоматизации перевода и формализации языковых данных и процессов.

С помощью современных технологических решений можно получить связный и точный текст на другом языке, как правило, не требующий редактирования человеком, за считанные секунды. Эту возможность нам дает применение инструментов на базе искусственного интеллекта, в частности, самообучающихся нейронных сетей, разработкой которых занимаются ведущие научные центры и университеты мира: Эдинбургский университет в Шотландии, Массачусетский, Стэнфордский и Колумбийский университеты в США, Монреальский университет в Канаде и МГУ имени М. В. Ломоносова. Последние несколько лет компании, являющиеся мировыми лидерами в области информационных технологий OpenAI, Promt, Facebook AI Research Lab, Google и Microsoft внедряют технологии нейронных сетей в системы машинного перевода.

Разработчики нейронных сетей осознают уязвимость искусственного интеллекта и стремятся максимально приблизить ИИ-перевод к человеческому. Одним из самых многообещающих последних достижений в области NLP является использование трансформеров (transformers), новой архитектуры нейронной сети, позволяющей достичь самых современных результатов при решении задач машинного перевода. Особенностью трансформеров является их способность одновременно анализировать все части предложения, что значительно ускоряет процесс обработки и улучшает понимание контекста.

Как правило, система сначала кодирует исходное предложение в абстрактный набор чисел – а потом декодирует из чисел слова, но уже на другом языке. «Дочитав» предложение до конца, причем одновременно и слева направо, и справа налево, нейросеть начинает декодировать – и предсказывать слова перевода, причем каждое предсказанное слово используется для предсказания следующего слова. Тем самым для более точного выбора слов для перевода используется контекст всего исходного предложения, а также контекст всех предыдущих предсказанных слов. Именно в этом основное отличие нейронных систем от систем, основанных на использовании правил и статистических методов, в которых контекст ограничен предыдущими 4–7 словами. И это более похоже на то, как переводит человек.

Системы с архитектурой последнего поколения Transformer, как правило, оснащены так называемым «механизмом внимания» (attention mechanism), который при предсказании следующего слова как бы фокусируется на одном или нескольких словах исходного предложения, складывая эту информацию с закодированным полным контекстом. Контроль внимания в процессе перевода представляет собой сложный многоуровневый процесс, который в работе искусственного интеллекта и в деятельности переводчика реализуется совершенно по-разному. Если внимание искусственного интеллекта – это «встроенные алгоритмы, призванные передать информацию через систему кодирования и декодирования за счет многоуровневых математических функций, то внимание переводчика – это способность концентрироваться на разных задачах, как лингвистических, так и внелингвистических, одновременно и успешно их выполнять» (Шебаршина, 2021).

Раньше системы машинного перевода строились под конкретные языки, а в настоящее время одним из основных требований к подобному программному обеспечению является мультязычность. Кроме того, очень востребованы наиболее сложные в этой отрасли системы по распознаванию и синхронному переводу человеческой речи.

По данным, опубликованным в научном журнале New Scientist, в 2022 была создана и успешно протестирована инновационная модель ИИ, которая может осуществлять перевод с 204 языков. Модель, получившая название No Language Left Behind (NLLB), опережает систему Google Translate, которая в настоящее время работает на 133 языках, и систему Microsoft Translator, которая обслуживает 110 языков. Модель NLLB поддерживает такие языки, как ачехский, на котором говорят 4,7 миллиона человек в Индонезии и Малайзии,

и фриульский, на котором говорят всего 600 000 человек в северо-восточной Италии, а также может выполнять высококачественные переводы на 55 африканских языков. Несмотря на свое название, No Language Left Behind охватывает лишь незначительную часть из почти 7000 языков, существующих во всем мире.

Большие языковые модели, подобные NLLB, обучаются на огромных объемах разнообразных текстовых данных, таких как книги, статьи, веб-страницы и т.д. В процессе обучения модель учится понимать языковые закономерности, структуру предложений и контекст. Обучение больших языковых моделей – это сложный и ресурсоемкий процесс, требующий комбинации передовых алгоритмов машинного обучения и больших объемов данных.

Языковая модель NLLB использует трансферное обучение, что означает, что она может применять полученные знания из одной задачи, чтобы обучиться выполнению другой задачи. Например, модель NLLB, обученная переводить с английского на французский язык, может использовать полученные знания для перевода с английского на испанский язык.

По мнению представителей научного сообщества и Ассоциации переводческих компаний, качество нейронного машинного перевода достигло такого уровня, что применение данных технологий коренным образом изменило профессию переводчика и всю отрасль переводов и локализации в целом. Авторы исследования «The future of employment: how susceptible are jobs to computerisation» Б. Фрей и М. А. Осборн, опубликованного на сайте Оксфордского университета, полагают, что благодаря скорости, качеству и гарантии конфиденциальности машинный перевод будет востребован во всех отраслях, в том числе – в бизнесе и науке, а профессиональные переводчики все чаще будут заниматься постредактированием и готовить данные для обучения современных систем машинного перевода.

Эксперты полагают, что под воздействием современных интеллектуальных технологий профессия переводчика не исчезнет, но будет видоизменяться. ИИ войдет в инструментарий профессионалов, поэтому целесообразно развивать технологии искусственного интеллекта в РФ и на их основе создавать современные эффективные системы машинного перевода.