

## МАШИННЫЙ ПЕРЕВОД

До Тхи Хань

Томский политехнический университет  
634050, Россия, г. Томск, пр-т Ленина, 30  
E-mail: thd1@tpu.ru

### Машинный перевод: фантастика или реальность?

Заменят ли компьютеры переводчиков? Если да, то когда? Если нет, то почему? Между этими двумя противоположными точками зрения находится мир возможностей, которые входят в категорию МП (машинного перевода). Качественный машинный перевод уже стал классикой научной фантастики, однако в отличие от экспериментов по телепортации, необходимые для машинного перевода технологии уже существуют, и систем машинного перевода появилось уже весьма не мало и они продолжают совершенствоваться. Захотят ли переводчики и переводческие бюро работать с новыми технологиями? И вообще будут ли работать новые технологии? И самое важное, может ли начинающий сейчас карьеру переводчика надеяться на то, что в предстоящую четверть века от его карьеры останется хоть что-то смутно напоминающее работу переводчика?

Данная статья предлагает внести некоторую ясность в вопросе, волнующем умы как профессиональных лингвистов, так и обывателей, далеких от лингвистических споров.

Что же такое машинный перевод и как могут помочь переводчику в его работе?

### Машинный перевод

Машинный перевод – процесс перевода текстов (письменных, а в идеале и устных) с одного естественного языка на другой полностью специальной компьютерной программой. Так же называется направление научных исследований, связанных с построением подобных систем.

Сразу стоит обратить внимание на фразу «в идеале». Системы машинного, или автоматического перевода не идеальны. Даже письменный машинный перевод далек от совершенства, чего уж говорить об устном. При использовании систем машинного перевода без использования человека не обойтись, в силу чего различаются несколько видов взаимодействия человека и компьютера.

Формы организации взаимодействия ЭВМ и человека при машинном переводе:

- с постредактированием: исходный текст перерабатывается машиной, а человек-редактор исправляет результат;
- с предредактированием: человек приспосабливает текст к обработке машиной (устраняет возможные неоднозначные прочтения, упрощает и размечает текст), после чего начинается программная обработка;

- с интерредактированием: человек вмешивается в работу системы перевода, разрешая трудные случаи;

- смешанные системы (например, одновременно с пред- и постредактированием).

На данный момент ни одна система машинного перевода не способна выполнить полноценный письменный перевод, который бы требовал минимальной редактуры, или не требовал её вовсе. Всё дело в принципе работы подобных программ.

Два основных алгоритма машинного перевода: основанный на правилах; основанный на статистике. В первом случае программа использует заложенные в неё грамматические, морфологические и синтаксические правила, во втором – большой объем языковых пар (тексты на двух разных языках, содержащие соответствующие друг другу предложения). Первый принцип используется в большинстве автономных систем (ПРОМТ и т. д.), по второму принципу работают онлайн-системы перевода: Google Translate, Яндекс.Перевод и Lingvo.Pro компании АBBYY. Исходя из опыта использования подобных систем, можно сделать вывод, что статистические системы демонстрируют намного более удовлетворительные результаты. Переводчик Google действительно выдаёт относительно вменяемый текст, причем можно даже говорить о сохранении тематики. Минусом статистического принципа является, во-первых, необходимость наличия подключения к интернету, во-вторых, зависимость от базы данных. Крупные концерны, такие как Google, располагают огромной базой данных, которую можно использовать для операции перевода, однако не все компании являются Гуглом.

Являясь преимуществом для текстов технической тематики, для художественных текстов стандартизация не применима, т. к. в художественном тексте наиболее важным критерием как раз таки является новизна, нестандартность использования языковых средств и экспрессивность. Таким образом, несмотря на прогресс систем автоматического перевода, главенствующая роль всё же принадлежит человеку.

Кроме того, результат работы автоматической системы перевода может стать прекрасным поводом для смеха, а если не повезёт, то и для беды. К счастью, большинство людей не готово доверить машине перевод важных документов, а вот веселья от существующих вариантов переводов программой генерируется предостаточно. Один из канонических примеров раннего этапа развития машинного перевода: фраза из Евангелия от Матфея (Матф. 26:41) на латинском Spiritus quidem

promptus est caro autem infirma была переведена на русский следующим образом: «Водка крепка, а закуска протухла»

И ещё кое-что: автоматический перевод не стоит путать с автоматизированным. Самое главное отличие – при автоматизированном переводе переводчиком является именно человек. При этом используются специальные программы, позволяющие улучшить качество перевода. К ним могут относиться:

- электронные словари (ABBYY Lingvo, Multitran, Multilex, GoldenDict и т. д.);
- программы проверки правописания;
- программы проверки грамматики;
- терминологические базы данных;
- программы автоматизированного перевода (OmegaT, Trados).

#### **Что можно и что нельзя переводить машинным способом**

Главное преимущество машинного перевода в том, что он позволяет быстро справиться с очень большими объёмами текста и поэтому иногда оказывается экономически выгоднее перевода вручную. При этом следует помнить, что качество машинного перевода всегда будет уступать человеческому. Поэтому использовать его целесообразно лишь в определённых случаях

Во-первых, машинным способом могут переводиться материалы для внутреннего пользования, например, когда требуется в общих чертах понять содержание сайта, статей или писем на иностранном языке или найти сообщения на ту или иную тему в прессе на нескольких языках мира. Во-вторых, технические и узкоспециальные тексты, которые затем пойдут на редактирование специалистам по данной тематике – в этом случае машинный перевод используется в качестве подстрочника, на основе которого технический специалист будет создавать финальный текст, опираясь на свои знания в предметной области.

Многие типы материалов в принципе не предназначены для машинного перевода. Так, нельзя доверять машине тексты, где неточность перевода может поставить под угрозу здоровье человека, работоспособность сложного прибора или крупный контракт – сэкономленное время здесь не оправдывает риска. Любые документы, подразумевающие юридическую ответственность (дого-

воры, гарантийные обязательства), требуют контроля человека. Машинный перевод непригоден для маркетинговых материалов, где текст фактически переосмысливается в новом культурном контексте и создается заново.

В целом, приемлемого качества можно ожидать при переводе строго формализованных технических текстов, в то время как художественные и рекламные тексты машинному переводу не поддаются.

#### **Заключение**

За более чем полвека своего существования машинный перевод из полуфантастического проекта стал реальностью. Вопрос качества работы систем машинного перевода стоит на стыке двух дисциплин: информационных технологий и лингвистики, причем ни одна из них пока не готова предложить решения для системы перевода, способной конкурировать с разумом человека. Компьютер вряд ли когда-нибудь заменит переводчика во всех областях, однако машинный перевод уже активно применяется как полезный инструмент, пусть и подходящий для решения строго определённых задач. Ни в IT, ни в лингвистике до сих пор не раскрыт вопрос природы человеческой речи, глубинный принцип её формирования не ясен, а потому о системах, способных полностью заменить человека мы можем только мечтать, что, возможно, и к лучшему. Для того чтобы понять, в каких случаях стоит и в каких не стоит прибегать к машинному переводу, и чтобы полученный результат не разочаровал, важно знать принципы работы МП, его ограничения и область его применимости.

#### **Литература**

1. Кузнецов П.С., Ляпунов А.А., Реформатский А.А. Основные проблемы машинного перевода. Вопросы языкознания, 1956, № 5.
2. Кулагина О.С. Исследования по машинному переводу. М., Наука, 1979 г.
3. Кулагина О. С. О современном состоянии машинного перевода// Математические вопросы кибернетики, вып. 3, М.: Наука, 1991 г., стр. 5 – 50. Библиография из 140 названий. ISBN 5-02-014323-5.
4. Будущее машинного перевода // Компьютерра № 21, 05 июня 2002.