

К ПРОБЛЕМЕ КАЧЕСТВА МАШИННОГО ПЕРЕВОДА ТЕКСТА

Матросов С.В., студент

Научный руководитель — Соловьева Ю.А., к.т.н., доцент

Сибирский государственный индустриальный университет, г.Новокузнецк

E-mail: shamanisheee@list.ru

Понятие машинного перевода как действия, фактически направленного на преобразование текста, который был описан на одном естественном языке, в текст-эквивалент ему по составу, структуре и содержанию на другом естественном языке, естественном или же синтетическом, возникло вскоре после создания электронных словарей [1]. Появление подобных средств, помогающих в и облегчающих поиски нужного слова, вместе с тем предлагая множество других, не менее полезных функций, стало отправной точкой в истории перевода текста уже с помощью машины, а не только силами человека. Впрочем, машинным переводом называется и сам результат описанного ранее действия, а также направление научных исследований, связанное с построением и изучением подобных систем.

Сам по себе перевод – деятельность комплексная и многогранная, и данный процесс нельзя описать как банальную замену одного языка другим. В процессе перевода происходит взаимодействие различных культур, опыта личности, разных менталитетов, традиций и всевозможные культурные установки. Не стоит забывать и о том, что перевод также ценится и как источник информации об языках, участвующих в процессе перевода, а также о культурах, к которым данные языки принадлежат. Современная теория перевода, основываясь на многолетнем опыте, исходит из идеи того, что перевод является средством общения в той же мере, в какой им является и сам язык. Потому первостепенной целью лингвистического и коммуникационного исследования перевода является создание коммуникативной модели перевода, которая будет в первую очередь направлена на облегчение общения между носителями разных естественных языков, представителями различных стран и культур.

Следуя теории коммуникации двух разных культур, в процессе перевода сопоставляются не просто два разных языка, но и сами культуры в широком понимании этого слова. Межкультурные различия могут вызвать определенный казус: понятное представителям одного языка и культуры, к которой он относится, может вызвать непониманием у иностранного получателя текста перевода, причем подобное характерно не только для двух разных естественных языков, но даже и в рамках различных диалектов одного и того же языка! Таким образом, задача перевода включает в себя не столько отыскание идентичных, эквивалентных соответствий для передачи той коммуникативной установки, которую подразумевал отправитель, но и предоставление необходимых для понимания получателю дополнений и пояснений так, чтобы реакция иноязычного получателя была аналогично реакции получателя сообщения на исходном языке. Лишь тогда перевод можно будет называть тождественным, а, значит, и более понятным и возможным к применению.

В настоящее время проблема низкого качества машинного перевода еще не решена в полном объеме. Для ее решения необходимы не только комплексные и глубокие теоретические исследования, но и довольно большая научно-экспериментальная работа. Начало ей, конечно же, положено – последние несколько лет различными компаниями, работающими в данной предметной области, ведется изучение и совершенствование гибридной [2] модели перевода текста – той, что, по сути, объединяет в себе лучшие черты, которыми были характерны ее предшественники, не только расширяя при этом само «поле деятельности», но добавляя

и дополнительные статистические шаги в процессе перевода, служащие для повышения его качества.

Построение качественного машинного перевода текстов любой направленности попросту немыслимо без использования комплексов словарей, которые включают в себя семантические единицы различных уровней: слова, словосочетания, предложения. Ведущее место должны занимать именно словосочетания, так как в текстах (будь то технические или даже художественные) наименования понятий очень редко выражаются отдельными словами, чаще всего являясь целыми словосочетаниями или даже предложениями. В процессе перевода при этом должно соблюдаться правило предпочтения, которое говорит о том, что, прежде всего, следует вычленять из исходного текста и переводить семантические единицы высшего уровня, то есть словосочетания и предложения, а к единицам более низких уровней обращаться лишь в том случае, если нет иной возможности выполнить точный перевод.

Исследования в данной предметной области стоит отнести к разряду фундаментальных исследований. Главная причина их возникновения и существования в наше время не столько в том, что это позволит в будущем снизить расходы на перевод всевозможных иностранных текстов и облегчить процесс обмена научной, художественной и деловой информацией между различными странами мира [3], но и в том, что они являются одним из наиболее эффективных путей поиска технических решений в области проблем автоматической обработки данных, уровень развития методов и средств которой оказывает непосредственное влияние на рост или снижение темпов научно-технического прогресса.

Высока и их теоретическая значимость в виду того, что методы построения систем машинного перевода текста изучены недостаточно хорошо, а, значит имеется возможность более глубокого и подробного их исследования. Прикладная ценность полученных результатов заключается в их возможном использовании в изучении и решении проблем, возникающих в данной предметной области, которые диктуются потребностями конечных пользователей, работающих в данной предметной области и могут изменяться в зависимости от конкретной решаемой прикладной задачи.

Можно с уверенностью констатировать следующий факт: несмотря на то, что на протяжении последних лет ведутся активные исследования в предметной области, связанной с обработкой и машинным переводом текстов, даже построение хорошей системы, близкой по своему качеству к тому, что способен производить человек-профессионал, нельзя будет считать завершающей стадией исследования, так как всегда будет иметься ряд менее приоритетных задач, требующих своевременного решения.

Список литературы:

1. С.В. Матросов, К.С. Порошин. Основные этапы развития систем машинного перевода текста / Информационно-телекоммуникационные системы и технологии (ИТСиТ-2014): Материалы Всероссийской научно-практической конференции, г. Кемерово, 16-17 октября 2014 г.; Кузбас. гос. техн. ун-т им. Т.Ф. Горбачева. – Кемерово, 2014. – 464 с. (стр. 270-271).
2. Статистические и гибридные методы перевода в технологиях компании PROMT - Control Engineering Russia [Электронный ресурс] / URL: <http://www.controlengrussia.com/innovatsii/statisticheskie-i-gibridny-e-metody-perevoda-v-tehnologiyah-kompanii-promt/> (Дата обращения: 07.04.2015).
3. С.В. Матросов, К.С. Порошин. Анализ систем машинного перевода текста (доклад) / Перспективы развития информационных технологий: Труды Всероссийской молодежной научно-практической конференции, г. Кемерово, 29-30 мая 2014 г. – Кемерово, 2014. – 401 с. (стр. 249-250).