

Секция «Текст и дискурс: проблемы анализа и интерпретаций»

Сравнение машинного перевода и перевода, выполняемого человеком, с точки зрения гипотезы лингвистической относительности

Научный руководитель – Яковлев Алексей Александрович

Дудникова Варвара Эдуардовна

Студент (бакалавр)

Российский университет дружбы народов, Факультет физико-математических и естественных наук, Москва, Россия

E-mail: var.dudnikova@gmail.com

Аннотация. Информационные технологии играют в наши дни все большую роль. Еще недавно было не так популярно пользование автоматизированным переводом, поскольку качество оставляло желать лучшего. Сегодня, несмотря на все несовершенство, автоматизированный перевод получил куда большее распространение. Одной из нерешенных проблем остается лишь условная точность (относительность) перевода. Еще во второй четверти XX века американские философы и лингвисты Э. Сепир и Б. Уорф сформулировали гипотезу лингвистической относительности, согласно которой разным языкам соответствуют разные картины мира, т.е. разные системы интерпретации. В статье рассмотрена проблема анализа и интерпретации текста с помощью гипотезы лингвистической относительности.

Ключевые слова: Машинный перевод, автоматизация, гипотеза лингвистической относительности, интерпретация текста.

В наше время, в эпоху компьютерных технологий, всё более востребованным становится перевод с одного языка на другой, так как большое количество информации в интернете, и не только, представлено на иностранных языках, как правило, люди принимают подобную информацию уже обработанной и переведенной на понятный им язык. Но случается и так, что информация попадает в первоизданном виде, например, какая-нибудь книга на языке оригинала. Однако не всегда это произведение имеет перевод. Но не желая отказываться от получения интересующей информации, человек прибегает к помощи систем машинного перевода. Качество машинного перевода оставляет желать лучшего: в большинстве случаев, текст, переведённый машиной, получается недостаточно складным, значения некоторых слов подобраны не совсем корректно.

Это можно объяснить работами Э. Сепира, который рассматривал язык как прежде всего строго организованную систему, все компоненты которой - такие, как звуковой состав, грамматика, словарный фонд, - связаны жесткими иерархическими отношениями. [7, 103] Связь между компонентами системы отдельно взятого языка строится по своим внутренним законам, в результате чего спроецировать систему одного языка на систему другого, не исказив при этом содержательных отношений между компонентами, оказывается невозможным. Приведем цитату из статьи Сепира «Статус лингвистики как науки» (1928): «Два разных языка никогда не бывают столь схожими, чтобы их можно было считать средством выражения одной и той же социальной действительности». [9, 209] Понимая лингвистическую относительность именно как невозможность установить покомпонентные соответствия между системами разных языков, Сепир ввел термин «несоизмеримость» (incommensurability) языков. [7,103]

Машинный перевод действует согласно алгоритмам. И это действительно сложные межъязыковые алгоритмы с применением словарей, правил грамматики, и с уже наработанной базой данных. Но! Перевод текста — это не просто перевод слов с одного языка

на другой, а гораздо сложнее. Поэтому любая автоматизированная интерпретация текста не сравнится с переводом человека! Хотя, ссылаясь на цитату Сепира, и человеку это в полной мере не подвластно.

Несмотря на всё это, использование электронных систем перевода увеличивается. Информация, перевод которой требуется, с каждым днём возрастает, соответственно возрастает актуальность автоматизированных переводов. (но это не означает, что потребность в переводе человеком уменьшается).

Углубимся в историю машинного перевода, который берет начало в XVII веке, когда такими философами, как Лейбниц и Декарт было выдвинуто предположение о существовании некоего кода, соединяющего между собой слова разных языков. Все предположения носили гипотетический характер, и никому не удавалось в действительности создать машину для перевода.

Первые заявления на получение патента на «переводческую машину» были поданы в середине 30-х годов XX века. Одно заявление было подано французским изобретателем Ж. Арцруни, просто создавшим автоматический двуязычный словарь на перфоленте. Другое было сделано советским ученым Петром Троянским, чье изобретение было более детальным. Оно включало в себя, как двуязычный словарь, так и способы работы с грамматическими ролями между двумя языками на основе Эсперанто. Данная система представляла собой три этапа: первый заключался в следующем - редактор (носитель языка) должен был связать слова исходного языка (ИЯ) в логические формы в соответствии с синтаксическими функциями; на втором этапе машина должна была «перевести» все эти формы на язык перевода (ПЯ); а на третьем этапе носитель языка перевода занимался редактированием полученного результата. Его схема оставалась неизвестной до конца 50-х годов, когда появились ЭВМ.

Сейчас машинный перевод делится на три принципиально разных подхода к построению алгоритмов перевода: основанный на правилах [U+202F] (rule-based), статистический, или [U+202F] основанный на статистике [U+202F] (statistical-based) и нейронный перевод. Первый подход является традиционным и использовался большинством разработчиков систем машинного перевода до появления нейросетевого перевода.

Рассмотрим первые два алгоритма с точки зрения гипотезы лингвистической относительности, которой также занимался Б.Уорф, ученик Сепира. Уорф считал, что в концептуализацию входит не только выделяемые в тексте формальные единицы - такие, как отдельные слова и грамматические показатели, - но и избирательность языковых правил, т.е. то, как те или иные единицы могут сочетаться между собой, какой класс единиц возможен, а какой не возможен в той или иной грамматической конструкции и т.д. [6, 46] На этом основании Уорф предложил различать открытые и скрытые грамматические категории: одно и то же значение может в одном языке выражаться регулярно с помощью фиксированного набора грамматических показателей, т.е. быть представленным открытой категорией, а другом языке обнаруживаться лишь косвенно, по наличию тех или иных запретов, и в этом случае можно говорить о скрытой категории. [6, 50] Первые два алгоритма машинного перевода, получается, основаны как раз на теории господина Уорфа.

Проблемы машинного перевода связаны именно с наличием смысловой, качественной уникальности актов речевой деятельности. Машинный перевод воспринимает слова, связи слов, грамматику, лексику, но смысл как нечто составное текста в ближайшее время ему точно не осознать. Но и человеку, чтобы понять смысл текста, надо искусно владеть родным языком и понимать, что 100% интерпретацию получить невозможно. Ведь у каждого языка есть свои погрешности и секреты.

Изучив структуру машинного перевода, его настоящее происхождение, в понимании

лингвистической гипотезы относительности. А также, учитывая, что результат машинного перевода может быть далек от истинного смысла текста и, что электронный переводчик создан не для того, чтобы заменить живых переводчиков, а для того чтобы помочь преодолеть языковые барьеры, обусловленные незнанием иностранных языков, поэтому лучшим переводчиком является человек, несмотря на все гипотезы и теории!

Список используемой литературы:

1. Апресян, Ю. Д. Интегральное описание языка и системная лексикография — М.: Языки славянской культуры, 1995;
2. Булыгина, Т. В, Шмелев, А. Д. Языковая концептуализация мира (на материале русской грамматики, — М.: Языки славянской культуры, 1997;
3. Ванников, Ю. В. Языковая сложность текста как фактор трудности перевода. — М.: Всесоюзный центр переводов, 1988;
4. Федоров А. В., Введение в теорию перевода. — М.: Издательство литературы на иностранных языках, 1953;
5. Федеров, А. В. Основы общей теории перевода. — М.: Высшая школа, 1968;
6. Фрумкина, Р. М. Семантика и категоризация — М.: Наука, 1991;
7. Хворостин, Д. В. Философский смысл гипотезы лингвистической относительности Эдуарда Сепира и Бенджамена Уорфа;
8. Рецкер, Ю.И. Теория перевода и практика перевода. — М.: Международные отношения, 1974;
9. Сепир Э. Избранные труды по языкознанию и культурологии. — М., 1993.