

Соловчук Александр Михайлович
УО "Брестский государственный университет имени А.С. Пушкина"
Преподаватель-стажер

-

Технологии машинного перевода
solovchuk_aleksandr@mail.ru

А. М. СОЛОВЧУК

Брест, БрГУ имени А. С. Пушкина

ТЕХНОЛОГИИ МАШИННОГО ПЕРЕВОДА

Машинный перевод на основе правил (Rule-Based Machine Translation) — это общий термин, который относится к системам машинного перевода на основе лингвистической информации об исходном и переведенных языках. Они состоят из двуязычных словарей и грамматик, охватывающих основные семантические, морфологические, синтаксические законы каждого языка. Такой подход к машинному переводу также называют классическим. На основании этих данных исходный текст последовательно, в соответствии с предложениями, преобразуется в текст перевода. Эти системы отличаются от систем машинного перевода, которые основаны на примерах. Принцип работы таких систем заключается во взаимосвязи структуры входных и выходных предложений.

RBMT системы делятся на три группы:

- системы пословного перевода;
- трансферные системы (Transfer) — преобразуют структуры входного языка в грамматические конструкции выходного языка;
- интерлингвистические системы (Interlingua)- промежуточный язык описания смысла.

Основным преимуществом является большой охват текстов с приемлемым уровнем качества перевода, а также низкая стоимость начальной разработки и модернизации.

Особенности RBMT систем:

- преимущества: синтаксическая и морфологическая точность; стабильность и предсказуемость результата; возможность настройки на предметную область.
- недостатки: трудоемкость и длительность разработки, необходимость поддерживать и актуализировать лингвистические бд; «машинный акцент» при переводе.

Машинный перевод на основе примеров (англ. Example-based machine translation, EBMT) — Это метод машинного перевода, который часто характеризуется использованием двуязычного корпуса с параллельными текстами в качестве основной базы знаний для перевода. Фактически это перевод по аналогии, который можно рассматривать как применение метода рассуждения, основанного на прецедентах, к машинному обучению.

Основой данного перевода является идея перевода по аналогии. Что касается процесса перевода человека, идея о том, что перевод осуществляется по аналогии, является отказом от идеи, что люди

переводят предложения с использованием глубокого лингвистического анализа. Вместо этого эта идея основана на убеждении, что люди переводят, сначала разбирая предложения в конкретные фразы, затем переводя эти фразы, и, наконец, путем правильного объединения этих фрагментов в одно длинное предложение. Переводы фраз выполняются аналогично предыдущим переводам. Принцип перевода по аналогии закодирован в машинном переводе на основе примеров через примеры перевода, которые используются для обучения такой системы. Другие подходы к машинному переводу, в том числе статистический машинный перевод, для изучения процесса перевода также используют двуязычный корпус.

Нейронный машинный перевод (англ. NeuralMachineTranslation, NMT) — это подход к машинному переводу, в котором используется большая искусственная нейронная сеть. Он отличается от методов машинного перевода, основанных на статистике фраз, которые используют отдельно разработанные подкомпоненты[1].

Модели NMT используют продвинутые и функциональные тренировки. Их работа требует лишь небольшой части памяти по сравнению с традиционными системами статистического машинного перевода (SMT). Кроме того, в отличие от традиционных систем перевода, все части нейронной модели перевода обучаются вместе (от начала до конца), чтобы максимизировать эффективность перевода

Список использованной литературы

1. «Википедия — свободная энциклопедия» [Электронный ресурс]. – Режим доступа http://ru.wikipedia.ru/wiki/Машинный_перевод – Дата доступа :20.03.20