

МАШИННЫЙ ПЕРЕВОД В MICROSOFT OFFICE WORD

к.ф.-м.н., доцент, Виктор Сергеевич Дацко. Филиал «Протвино» ГОУ ВПО МО
«Международный университет природы, общества и человека “Дубна”»

Обосновываются достоинства машинного перевода. Приведён пример такого перевода.

MACHINE TRANSLATION MICROSOFT OFFICE WORD

PhD, Associate Professor, V.S. Datsko «International University of Human, Society and Nature
“Dubna” », “Protvino” branch

Argue accomplishments of machine translation. An example of such a denote.

За последние 50-60 лет знание английского языка стало необходимостью не только для чтения научно-технической литературы, но и в обыденной жизни. Это объясняется не только развитием экономики в англоязычных странах, но и простотой (можно даже сказать примитивизмом) английского языка. В связи с такой популярностью знанием английского языка, в настоящее время, обладают весьма многие представители неанглоязычного населения. Казалось бы, что в таких условиях нет никакой проблемы в переводе английского текста – бери и читай! Однако существуют несколько факторов осложняющих перевод.

Есть значительная разница в переводе художественного и научно-технического текста. Машинный перевод научно-технических и деловых текстов (про художественные пока что говорить не приходится) может быть построен на основе использования разнообразных словарей. В первом случае необходимо максимально точно передать стиль автора, а во втором максимально точно описать научно-техническую проблему или конструкцию. На пути такого описания встречается специфическая проблема, которую, например К. К. Агузарова, иллюстрирует так: встречаются общелитературные английские слова, употребляемые в специальном значении. К таким словам относятся, например: **jacket** – *куртка* и, вместе с тем – *кожух*; **jar** – *кувшин* или *конденсатор*; to load - *нагружать* и *заряжать*. Или встречается употребление одного слова в разных специальных значениях, в зависимости от отрасли знания; слово **pocket** - **карман**, например, имеет следующие специальные значения: **воздушная яма** (в авиации); **окружение** (в военном деле); **мертвая зона** (в радио); **гнездо месторождения** (в геологии); **кабельный канал** (в электротехнике). Таким образом, и при переводе художественного и при переводе научно-технического текста приходится использовать различные словари для наилучшего переложения информации на другой язык. Необходимость использования словарей и, соответственно, поиск в них нужных слов значительно замедляет скорость перевода даже для квалифицированного переводчика. А ведь обычно, если используется набор словарей, то речь идёт не просто о чтении текста, а и о его записи. А это значит, что переведённый текст необходимо ещё и набрать на клавиатуре компьютера в каком-либо редакторе. Если же перевод выполняет неквалифицированный работник, то оба указанные факторы значительно замедляют работу.

В то же время, поскольку всё равно для набора текста используется компьютер, целесообразно использовать какую-нибудь программу машинного перевода. В этом случае переводимый текст нет необходимости набирать на клавиатуре, он требует только редакторской правки, которую необходимо выполнить в любом случае. Объём работы по такой правке различен в зависимости от качества работы программы машинного перевода. Однако опыт показывает, что для неквалифицированного переводчика даже простая программа машинного перевода существенно облегчает работу.

Microsoft Office Word как наиболее распространённый текстовый редактор предусматривает, за некоторую доплату, использование встроенной программы машинного перевода. Более того, есть возможность использовать средства машинного перевода без встроенной программы. Правда, при этом нет возможности переводить большие объёмы текста,

кроме того, для перевода даже отдельного слова редактор обращается к интернету, что представляет определённые неудобства (как мне кажется – значительные). В то же время в редакторе запасены возможности для создания собственной программы машинного перевода средствами Visual Basic for Applications (VBA, Visual Basic для приложений) — немного упрощённой реализации языка программирования Visual Basic, встроенной в линейку продуктов Microsoft Office.

В настоящее время существует достаточно большое количество программ машинного перевода более или менее отличающихся друг от друга. Сравнительный анализ их работы – тема отдельной статьи, например, в браузере Google можно пользоваться программой “Переводчик” для перевода не очень больших текстов и сравнительно неплохого качества. Все они требуют после себя редакторской правки, поэтому даже самая простая программа не сильно уступает уже существующим. Самостоятельно написанная программа в любой момент может быть улучшена и при этом не требует никаких дополнительных финансовых затрат.

Логика работы такой программы и её структура очень просты. Читается слово переводимого текста, затем читается строка словаря и находится соответствие слов. Если в строке словаря переводимое слово содержится, то в этой строке есть и слово - перевод. Если в прочитанной строке словаря искомого переводимого слова нет – читается следующая строка словаря и исследуется соответствие. Процесс повторяется пока нужное соответствие не находится после чего переведённое слово заносится в массив, который периодически записывается в файл перевода, печатаемый в конце работы. Разумеется, процесс подразумевает некоторое количество служебных программ: чтения из файла, записи в файл и так далее.

В общем, структура реальной программы выглядит следующим образом:

```
'
'      I - menu
'      I
'      I      I-статья-Разбивка---I
'      I      I      I-Слово
'      I      I
'      I - чтнтекст----I      I-Предложение
'      I      I-книжка-Разбивка---I
'      I      I-Слово
'      I      I-Разнослов-Словод
'      I      I
' anghel ----I - srchw -----I-base
'      I
'      I - rowout
'
' -----72
'
' menu      - заказ режимов работы
' чтнтекст  - считывание строки переводимого текста, разделение ее по словам и предложениям
' srchw     - поиск английского слова в словаре
' rowout    - формирование строки из переведенных слов
' base      - выделение корня английского слова
' статья    - работа с небольшим текстом
' книжка    - работа с большим текстом
' Разбивка  - разбиение прочитанного абзаца на предложения и слова
' Предложение - функция, разбиение прочитанного абзаца на предложения
' Слово     - функция, разбиение прочитанного абзаца на слова
' Разнослов - разбиение прочитанной строки словаря на слова
' Словод    - функция, выделение слова в строке словаря
'
```

В программе “menu” определяется надо ли в случае ненайденного соответствия запрашивать оператора – какое слово подставлять в перевод вместо данного английского? Если задан режим без запросов, то вместо ненайденного слова-перевода будет подставлено исходное английское слово. Здесь же определяются файл, содержащий исходный текст, файл содержащий словарь, файл, куда будет записан перевод. Также в menu определяется размер переводимого текста, если это статья – файл целесообразно сначала прочитать word-ом, а получившийся текст макрокомандой читать в программе “статья”. Если файл большой – например книга, то лучше его читать абзац за абзацем из файла и далее читать в программе “книжка”.

Программа перевода существует в виде макрокоманды Word и может быть запущена, как только в ней возникнет потребность. Опыт показывает, что результат её работы не уступает или, во всяком случае, мало уступает другим программам машинного перевода. Есть некоторые надежды на улучшение качества перевода с её помощью, связанные с применением обучающихся алгоритмов, а качество нынешней работы можно оценить из следующего примера. Первый абзац отредактированный перевод, второй – результат машинного перевода (выделены курсивом).

Если используется "наилучшее приближение наблюдений" как критерий оценки λ , то угловое значение красного смещения для сверхкомпактных источников радиоизлучения (Джексон и Dodgeson 1997) подразумевает $\lambda < 0$. Имеется противоречие между двумя космологическими критериями для Вселенной с большими красными смещениями: t - z соотношение дает $\lambda > 0$, в то время, как θ - z соотношение требует $\lambda < 0$ для модели с наилучшим приближением данных. Неопределенность, окружает космологическую "стандартную свечу" и "стандартные правила", мы полагаем, что нынешний энтузиазм по поводу положительной космологической константы может быть преждевременен.

Альтернативный сценарий: QSSC

*Расстояние dL измерено в Мpc. Здесь мы использовали систему обозначений и *definitions* Perlmutter, и*

Al (1999). Количество M_{has} упомянутый этими авторами как " константа Хаббла освобождает ' абсолютный

Величина. Второй срок(термин) на правильной(правой) ручной стороне этого уравнения - стандартная яркость

Расстояние (dL) срок(термин), который может быть вычислен для QSSC. Изделие(программа) $H_0 dL$ независимо от Константа Хаббла. Последний срок(термин) обозначает уменьшение силы света, произведенное межгалактической присыпкой, и может

Быть явно напишите как:

Опыт работы с программой показывает, что целесообразно вводить в программу анализ словосочетаний (аналог так называемой "естественной грамотности", когда человек пишет грамотно, не задумываясь о соответствующих правилах просто потому, что много раз читал нечто похожее в других источниках). Такой анализ должен привести к лучшей смысловой передаче информации при переводе текста.

Решение проблем автоматической обработки информации, в принципе, подразумевает организацию больших коллективов, использующих мощную электронно-техническую базу, совершенные математические и статистические методы обработки информации.

Однако на примере развития электронно-вычислительной техники можно видеть, что обеспечение массовости даёт мощный стимул к совершенствованию. Представляется, что срабатывает простая логическая связь: - массовость обеспечивает экономическую рентабельность, которая стимулирует постепенное совершенствование, сопровождающееся более или менее крупными "скачками" – рационализациями, изобретениями, а иногда прорывами - открытиями. Кроме того, этот путь обеспечивает данное направление развития кадрами. В этом смысле полезно знакомить студенческую среду с соответствующими проблемами и методами их решения в процессе образовательного процесса (под лозунгом – "Это вы можете!").

Распространение "самодетельного" конструирования программ машинного перевода повышает вероятность обнаружения перспективных решений в построении программ перевода, которое позволит в будущем сэкономить средства на переводах иностранных текстов и облегчить процесс обмена научно технической и деловой информацией. Исследования по машинному переводу необходимы как один из наиболее эффективных путей поиска технических решений по ряду проблем автоматической обработки информации, а уровень развития средств и методов автоматической обработки информации оказывает непосредственное влияние на темпы научно-технического прогресса.