

To solve the problem of linear approximation in General, you should find the conditions a minimum sum of the squares of the deviations.

The proposed information technology for detection of periodic components of time series using least squares method. Research of time series for periodic components. The effectiveness of the method was studied by means of computational experiments based on the model of time series and based on the real-life series of Electrocardiograms of laboratory animals. The result of computational experiments analysis method and proposed practical recommendations.

References:

1. Brillindzher D. Vremennyye ryadyi. Obrabotka dannyih i teoriya. M.: Mir, 1980g. – 536p.
2. Zaytsev E.O. Zastosuvannya metodu naymenshih kvadrativ pri obrobsi monitorinhovoyi informatsiyi shlifival'noho obladnannya. /E.O. Zaitsev, V.E. Sydorchuk, I.V. Sydorchuk// Vimiryuvalna ta obchislyuavalna tehnika v tehnologichnih protsesah. – 2015. – №4(53). – P.175-179.
3. Ter-Krikorov A.M., Shabunin M.I. Kurs matematicheskogo analiza. M.: Nauka Gl. red. fiz. mat. Lit., 1988.
4. Vityazev V.V. Spectralno-korrelyatsionniy analiz ravnomernyih ryadov: Uchebnoe posobie – SPb.: Izdatelstvo S.-Peterburgskogo universiteta, 2001.

Прокопов І.В.

студент,

Харківський національний університет радіоелектроніки

СУЧАСНІ ПІДХОДИ ДО МАШИННОГО ПЕРЕКЛАДУ

Системи машинного перекладу можуть використовувати метод перекладу заснований на лінгвістичних правилах. Найбільш підходящі слова з вихідної мови просто замінюються словами переказної мови. Існує твердження, що для успішного вирішення проблеми машинного перекладу необхідно вирішити проблему розуміння тексту на природній мові.

Як правило, метод перекладу заснований на правилах використовує символічне уявлення (посередника), на основі якого створюється текст на переказній мовою. А якщо враховувати природу посередника то можна говорити про інтерлінгвістичний машинний переклад або трансферний машинний переклад. Ці методи вимагають дуже великих словників з морфологічною, синтаксичною і семантичною інформацією і великого набору правил. Сучасні системи машинного перекладу ділять на три великі групи:

- основані на правилах;
- основані на прикладах;
- статистичні.

Системи машинного перекладу засновані на правилах – загальний термін, який позначає системи машинного перекладу на основі лінгвістичній інформації про вихідну і переказну мови. Вони складаються з білінгвістичних словників і граматик, що охоплюють основні семантичні, морфологічні,

синтаксичні закономірності кожної мови. Такий підхід до машинного перекладу ще називають класичним. На основі цих даних вихідний текст послідовно, речення за реченням, перетворюється в текст перекладу. Ці системи протиставляють системам машинного перекладу, які засновані на прикладах. Принцип роботи таких систем – зв'язок структури вхідного і вихідного речення.

Ці системи діляться на три групи:

- системи послівного перекладу;
- трансферні системи;
- інтерлінгвістичні.

Системи послівного перекладу використовуються зараз вкрай рідко через низьку якість перекладу. Слова вихідного тексту перетворюються як є в слова перекладного тексту. Часто таке перетворення відбувається без лематизації і морфологічного аналізу. Це найпростіший метод машинного перекладу. Він використовується для перекладу довгих списків слів, наприклад, каталогів.

Як трансферні системи, так і інтерлінгвістичні, мають одну і ту ж загальну ідею. Для перекладу необхідно мати посередника, який в собі несе сенс перекладного виразу. У інтерлінгвістичних системах посередник не залежить від пари мов, в той час як в трансферних – залежить.

Трансферні системи працюють за дуже простим принципом: до вхідного тексту застосовуються правила, які ставлять у відповідність структури вихідної і переказної мов. Початковий етап роботи включає в себе морфологічний, синтаксичний, а іноді і семантичний аналіз тексту для створення внутрішнього уявлення. Переклад генерується з цього подання з використанням білінгвістичних словників і граматичних правил. Іноді на основі первинного уявлення, яке було отримано з вихідного тексту, будують більш «абстрактне» внутрішнє уявлення. Це робиться для того, щоб акцентувати місця важливі для перекладу і відкинути несуттєві частини тексту. При побудові тексту перекладу перетворення рівнів внутрішніх уявлень відбувається в зворотному порядку.

При використанні цієї стратегії виходить досить висока якість перекладів, з точністю в районі 90% (сильно залежить від мовної пари). Робота будь-якої системи трансферного перекладу складається як мінімум з п'яти частин:

- морфологічний аналіз;
- лексична категоризація;
- лексичний трансфер;
- структурний трансфер;
- морфологічна генерація.

Інтерлінгвістичний машинний переклад – один з класичних підходів до машинного перекладу. Оригінальний текст трансформується в абстрактне уявлення, яке не залежить від мови (на відміну від трансферного перекладу). Перекладний текст створюється на основі цього подання. Можна довести математично, що в рамках цього підходу, створення кожного нового інтерпретатора мови для такої системи буде здешевлювати її, в порівнянні, наприклад, з системою трансферного перекладу. Крім того, в рамках такого

підходу можна реалізувати «переказ тексту», перефразування вихідного тексту в рамках однієї мови.

До сих пір не існує реалізацій такого типу систем, яка б коректно працювала хоча б для двох мов. Багато експертів висловлюють сумніви в можливості реалізації. Найбільша складність для створення подібних систем полягає в проектуванні міжмовної уявлення. Воно повинно бути одночасно абстрактним і незалежним від конкретних мов, але в той же час воно має відображати особливості будь-якої існуючої мови. З іншого боку, в рамках штучного інтелекту, завдання виділення сенсу тексту на даний момент до цих пір не вирішена.

Переклад заснований на прикладах – один з підходів до машинного перекладу, при якому використовується двомовний корпус тексту. Цей корпус тексту під час перекладу використовується як база знань. Передбачається, що люди розкладають вихідний текст на фрази, потім переводять ці фрази, а далі складають перекладний текст з фраз. Причому, переклад фраз зазвичай відбувається за аналогією з попередніми перекладами. Для побудови системи машинного перекладу, заснованої на прикладах потрібно мовний корпус, складений з пар речень. Мовні пари – тексти, що містять речення на однією мовою і відповідні їм речення на іншою мовою.

Переклад, заснований на прикладах, найкраще підходить для таких явищ як фразові дієслова. Значення фразових дієслів сильно залежить від контексту. Фразові дієслова дуже часто зустрічаються в розмовній англійській мові. Вони складаються з дієслова з прийменником або наріччям. Сенс такого виразу неможливо отримати з смислів складових частин. Класичні методи перекладу в даному випадку не застосовуються. Цей метод перекладу можна використовувати для визначення контексту речень.

Статистичний машинний переклад – це метод машинного перекладу. Він використовує порівняння великих об'ємів мовних пар. Статистичний машинний переклад має властивості самонавчання. Чим більше в розпорядженні мається мовних пар і чим точніше вони відповідають один одному, тим краще результат статистичного машинного перекладу. Статистичний машинний переклад оснований на пошуку найбільш вірогідного перекладу речення з використанням даних з двомовних корпусів текстів. У результаті під час виконання перекладу, комп'ютер не оперує лінгвістичними алгоритмами, а вичислює вірогідність застосування того чи іншого слова чи вираз. Слово чи послідовність слів, які мають оптимальну вірогідність, вважаються найбільш відповідними перекладу вихідного тексту та підставляються комп'ютером в одержуваний текст.

В статистичному машинному перекладі становиться задача не перекладу тексту а задача його розшифровки. Стаття написана англійською мовою, насправді являє собою статтю написану російською мовою, але її текст зашифровано (чи спотворено шумом). При такому підході становиться зрозуміло чому, чим далі мови, тим краще працюють статистичні методи, в порівнянні з класичними методами.

Таким чином, для перекладу тексту необхідно знайти спосіб декодування, який використовує природну надмірність, в зв'язку з чим декодування має бути ймовірнісним. Завдання такого декодування полягає в тому, щоб, при цьому повідомленні знайти вихідне повідомлення, якому відповідає найбільша ймовірність.

Робота статистичних систем, так само як і систем заснованих на прикладах відбувається в двох режимах: навчання та експлуатації. У режимі навчання проглядаються паралельні корпусу тексту і обчислюються ймовірності перекладних відповідностей. Будується модель мови перекладу. Тут же визначаються ймовірності кожної n-грами. У режимі експлуатації, для фрази з початкового тексту шукається фраза перекладного тексту, так, що б максимізувати добуток ймовірностей.

Список використаних джерел:

1. Ігітою С.В., Крючкова О.М., Старцев П.Л., Чумак М.В. Проблеми формування словників в системах штучного інтелекту.
2. Чен, Дж.: Паралельна обробка тексту для міжмовної інформаційного пошуку з використанням перекладу статистичної моделі. Магістр природних наук – Режим доступу: <http://www.iro.umontreal.ca/~chen/thesis/node1.html>.
3. Браун, Р.Д.: Автоматично витягнутий тезаурус для міжмовної IR: коли краще гірше, 1-ий семінар по обчислювальній термінології (Computerm), р. 15-21, 1998.

Прокопов І.В.

студент,

Харківський національний університет радіоелектроніки

МОДЕЛІ ВИРІВНЮВАННЯ ТЕКСТУ IBM

Моделі вирівнювань IBM представляють собою послідовність все більш і більш складних моделей, які використовуються в статистичному машинному перекладі для навчання моделі перекладу і моделі вирівнювання. Вони підкріпили більшість статистичних систем машинного перекладу майже двадцять років. Ці моделі пропонують принципову вірогідне формулювання.

Оригінальна робота по статистичному машинного перекладу в IBM запропонує п'ять моделей.

Найпростішою статистичною моделлю перекладу є модель дослівного перекладу. У моделі, відомої як Модель IBM №1, передбачається, що для перекладу речення з однієї мови на іншу досить перевести всі слова (створити «мішок слів»), а розстановку їх у правильному порядку забезпечить модель мови. Єдиним масивом даних, яким оперує Модель №1, є таблиця ймовірностей попарних перекладних відповідностей слів двох мов.

Навчання Моделі №1 проводиться на корпусі паралельних текстів, вирівняному на рівні речень. Математична частина Моделі №1 полягає в наступному. Ймовірність пропозиції оригіналу при даній пропозиції перекладу