

є називний відмінок однини, прикметників і дієприкметників – називний відмінок чоловічого роду, в одній словниковій статті подаються дієслова доконаного виду тощо). Якщо термін являє собою словосполучення, то він подається за алфавітом першого слова, а не стрижневого (абстрактна лексика, абстрактні іменники). Якщо термін має пов'язані з ним терміни меншого семантичного обсягу, то вони подаються в тій же словниковій статті (функція слова, функція комунікативна). Якщо термін має декілька значень, то всі вони наводяться з абзацу в тій же статті під цифрами 1, 2 і т.д. Граматична характеристика термінів узгоджена з основними законами літературної мови.

Джерельна база словника – сучасна нормативна термінологія, зафіксована в науковій і навчальній літературі, лексикографічних працях.

Своєрідною особливістю словника є те, що в додатках до нього подаються контрольнотренувальні вправи та завдання для тестувань з різних розділів граматики, якими передбачено не тільки відтворення мовних знань, одержаних у школі, але і вироблення підходів до рідної мови, як безцінної народної скарбниці. Цьому сприятимуть мікротексти гуманістичного, науково-технічного, народознавчого, та культурологічного характеру, які розглядаються в пізнавальному та прикладному аспектах. Безперечно, їх опрацювання допоможе зміцнити навички стилістичного використання певного слова або словосполучення, добирати потрібне слово для усного та писемного мовлення, оволодівати культурою мовленнєвого спілкування, збагачувати активний запас слів.

Робота з перекладними словниками сприяє збагаченню активного запасу слів студентів, а також виробленню навичок добирати потрібне слово, що є одним із вирішальних чинників інтеграції технічних і гуманітарних знань. Зазначені словники розглянуті Радою Укртрансвузу і рекомендовані як довідникові навчально-методичні посібники для студентів вищих навчальних закладів системи Міністерства транспорту України.

Все це розвиває лінгвістичне мислення, формує навички користування певними комунікативними моделями відповідно до мети, зумовленої як внутрішніми потребами, так і зовнішніми обставинами, що є одним із вирішальних чинників національної самобутності народу.

АЛГОРИТМ КОМП'ЮТЕРНОГО ПЕРЕКЛАДУ

© Леонід Сніцарук, 2000
ДУ “Львівська політехніка”

Розвиток міжнародних відносин, супроводжуваний бурхливим розвитком комп'ютерних технологій, зокрема мережі INTERNET, викликав появу програм автоматичного перекладу текстів, які постійно вдосконалюються і досягли вже достатньо високого рівня.

Абсолютна більшість цих програм використовує загальномовні словники, хоча останні версії (наприклад, Magic Goody) включають спеціалізовані словники, такі як медичні, правничі, різногалузеві технічні і т.п.

Алгоритм автоматичного перекладу побудований так, що відшукуються відповідники слів і здійснюється відповідне граматичне їх опрацювання. Останні версії програм пропонують поруч з підібраним базовим словом інші його значення. Під час читання тексту користувач має можливість увійти у спеціалізовані словники і відшукати там інші відповідники слова, що перекладається, оскільки алгоритм автоматичного перекладу цього не передбачає. Таким чином, можна одержати точний переклад як слів,

так і речень, фраз. Очевидно, що такий переклад не є автоматичним, займає багато часу і вимагає високої кваліфікації перекладача і фахових знань у галузі, до якої відноситься текст, що перекладається.

Побіжний аналіз роботи програм автоматичного перекладу текстів дозволяє запропонувати алгоритм, який повинен забезпечити суттєво вищий рівень та оперативність виконання перекладу.

Відповідно до структури пропонованого алгоритму здійснюється лексичний аналіз іменників оригінального тексту. Для цього може бути ефективно використаний реферат статті, де завжди є ключові слова. У разі його відсутності аналізується весь текст з метою виявлення ключових слів. Критерієм віднесення слів до ключових може бути їх частота у даному тексті. 3-4 найуживаніших іменники, які мають фахові ремарки, вказують ними галузь, якої стосується даний текст. Інший варіант виявлення ключових слів - безпосередньо за частотою фахових ремарок. Найкращий результат слід очікувати від поєднання цих двох методів. На основі цього аналізу робиться висновок про основну та другорядні галузі, до яких відноситься даний текст. Подальша робота програми полягає у підборі відповідників словам оригінального тексту як, в першу чергу, у словнику загальної лексики, так і в словнику (словниках), що стосується (стосуються) визначеної алгоритмом галузі. Наприклад, відносячи текст до галузі медицини, програма використовуватиме для перекладу словник загальної лексики та медичний. Оскільки багато спеціалізованих словників мають свій внутрішній розподіл термінів за вузькою спеціалізацією (багаторівневу спеціалізацію), то в разі виявлення підпрограмою аналізу характерних термінів, програма використовуватиме в першу чергу терміни цієї спеціалізації. Рівні спеціалізації можуть бути легко виявлені на стадії виявлення ключових слів (див. вищенаведений метод - за частотою фахових ремарок). Наприклад, у спеціалізованому медичному словнику, можливе використання термінів хірургії, терапії, офтальмології і т.п. Результат аналізу може бути проконтрольований користувачем або вказаний ним перед запуском програми перекладу тексту, тобто оминаючи етап аналізу.

Можливий дещо інший варіант алгоритму, якщо використовувати повний словник, на зразок словника Webster [1] чи Oxford. У цих словниках слова, що стосуються певних галузей, мають відповідні позначки (ремарки) і виділені в окремі групи. Наприклад, слово "plate" має відповідники: 1) пластинка, дощинка; 2) тарілка; 3) столове срібло; 4) плита, лист; 5) гравюра, естамп (мистецтво); 6) уклейка; 7) екслібрис; 8) друкарська форма (поліграфія); 9) призовий кубок; 10) перегони на приз (спорт); 11) вставна щелепа (медицина); 12) анод лампи (радіо).

Таким чином, алгоритм перекладу у цьому варіанті навіть дещо спрощується, оскільки не потрібно робити переходи в спеціалізовані словники - підбір слова-відповідника вестиметься в межах гнізда.

Очевидно, що для забезпечення максимальної точності перекладу доцільно вводити кілька рівнів спеціалізації в кожній галузі. Вище вже наводився приклад за медичним словником. У радіоелектроніці це можуть бути: приймачі, передавачі, телевізійна техніка, радіолокація, транзистори, вакуумні прилади, квантова техніка тощо.

Взаємна кореляція слів у стійких словосполучень в текстах, особливо вузькоспеціалізованих, значно вища, ніж окремих слів. Ця кореляція зростає з поглибленням рівня спеціалізації. На глибоких рівнях спеціалізації великий відсоток словосполучень притаманний лише цій спеціалізації. Для ефективного використання

словосполучень необхідно створити комп'ютерні перекладні словники з розвинутою системою багаторівневої спеціалізації. В таких словниках одночасно з базовими словами повинні бути гнізда, побудовані на цьому слові. Для цього у словосполученнях потрібно зберігати природний порядок слів, у якому вони вживаються. Наприклад, у словнику з радіоелектроніки базове слово "корекція" матиме гніздо, в якому в алфавітному порядку будуть наведені словосполучення "корекція спотворень; корекція високих частот; корекція зі стиранням помилок" і т.п. Для базового слова "кореляційний" - аналогічно: "кореляційний аналіз; кореляційний декодер; кореляційний інтеграл" і т.п.

За такими принципами побудований двомовний перекладний словник з радіоелектроніки [2], його російсько-українська частина.

Таким чином, запропонований алгоритм комп'ютерного перекладу з використанням нових словників словосполучень з розвинутою системою багаторівневої спеціалізації повинен забезпечити високий рівень автоматичного перекладу спеціалізованих текстів.

1. *Webster's Encyclopedic Unabridged Dictionary of the English Language. Gramercy Books, N.Y., New Jersey. 1989, 1854 pgs.* 2. *Російсько-український та українсько-російський словник з радіоелектроніки. Б.Рицар, К.Семенистий, І.Кочан. Логос, Львів, 1995, 608с.*

УКЛАДАННЯ СЛОВНИКІВ ЗА ТЕХНОЛОГІЄЮ ЛЕКСИКОГРАФІЧНОЇ БАЗИ ДАНИХ

© Ігор Кульчицький, Василь Костирко, 2000

Львівська комерційна академія

В даній роботі розглянуто застосування технології укладання словників з використанням баз даних на персональних комп'ютерах [1]. Основу цієї технології складає лексикографічна база даних (ЛБД), що розглядається як спеціалізована надбудова над стандартною системою керування базами даних. Укладання певного словника за цією технологією має такі чотири етапи:

- побудова абстрактної моделі цього словника;
- генерація структури та прикладного програмного забезпечення лексикографічної бази даних;
- наповнення лексикографічної бази даних;
- конвертування лексикографічної бази даних до належного кінцевого вигляду.

Побудова абстрактної моделі словника базується на таких його визначальних рисах:

• **Членоване розміщення матеріалу.** Основною композиційною і комунікативною одиницею словника є незалежний відрізок тексту, що називається **словниковою статтею**. Їх сукупність складає **основу** словника.

• **Ліва та права частини.** Кожну словникову статтю можна розділити на ліву та праву частини. Ліва частина — це будь-яка одиниця мови, що є об'єктом лексикографування і вноситься у словник. Назвемо її лексикографічною одиницею, а сукупність всіх лексикографічних одиниць словника його реестром. Праву частину словникової статті творять співвіднесені з лексикографічною одиницею лексикографічні параметри. Лексикографічний параметр — це деякий квант лінгвістичної інформації, що має самостійний інтерес для користувача і є словниковим відображенням структурних рис мови (наприклад, орфографічний параметр, графічна довжина слова, наголос, вимова, поділ на склади, частина мови, переклади, глумачення тощо).