

МЕТОД АВТОМАТИЧНОГО ПЕРЕКЛАДУ ТЕКСТУ З ВИКОРИСТАННЯМ КОНТЕКСТНО-ВІЛЬНИХ ГРАМАТИК

Яремчук В.В.

Науковий керівник – доц., к.т.н. Грищук Т.В.

Автоматичний переклад – це дія, яка виконується на комп'ютері, яка направлена на перетворення тексту на одній природній мові в еквівалентний за змістом текст на іншій мові, а також це результат такої дії. Проблема роботи сучасних систем машинного перекладу полягає у тому, що вони дають можливість накопичувати пам'ять перекладу, в якій зберігаються оригінальні та перекладені сегменти (речення, фрази тощо), але не дозволяють інтегрувати в них власні граматичні правила, за якими описуються фрази у вигляді контекстно-вільних граматик. Метою даної роботи є побудова моделі процесу перекладу тексту з використанням контекстно-вільних граматик.

Постановка задачі: задамо дві граматики $G1 = (V1_N, V1_T, S, P1)$ та $G2 = (V2_N, V2_T, S, P2)$, де V_T – множина термінальних символів, V_N – множина нетермінальних символів, причому $V_N \cap V_T = \emptyset$, S – початковий символ (всі речення виводяться з нього), P – множина правил виводу (або правил генерування). Задачею дослідження є розробка методу пошуку відповідності речення α з граматики $G1$, що містить елементи множини V_T , в граматичі $G2$. Введемо обмеження: початковий символ, нетермінали – назви команд (перший рівень дерева граматичи) є мовонезалежними. Таке обмеження дає можливість звузити область пошуку, якщо граматика описує більше ніж одну команду. Друге обмеження – другий рівень правил граматичи має складатись лише з нетермінальних символів. Це необхідно у випадках коли однаковий смисл може бути описаний різними реченнями.

Алгоритм перекладу за розробленим методом наступний: 1) на основі вхідної послідовності отримуємо назву команди та послідовність нетерміналів другого рівня в граматичі $G1$; 2) в граматичі $G2$ виділяємо піддерево, починаючи з вершини – назви команди та вершин другого рівня – назв нетерміналів вхідної послідовності; 3) обходимо отримане дерево, генеруючи всі можливі речення. Таким чином розроблений метод може бути використаний для побудови систем автоматичного перекладу на основі контекстно-вільних граматик та дозволяє отримати всі альтернативні варіанти фрази за одним екземпляром речення, перекладати речення з однієї мови на іншу, інтегрувати «знання», що зберігаються у вигляді граматик в сучасні системи комп'ютерного перекладу текстів.

1. Машинный перевод. Энциклопедия Кругосвет [Електроний ресурс] – Режим доступу: <http://www.krugosvet.ru> - Назва з екрану.
2. Хомский Н. Введение в формальный анализ естественных языков // Н. Хомский, Дж. Миллер. – М.: Едитоиал УРСС, 2003. – 64 с.