

РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ СИНТАКСИЧНОГО АНАЛІЗУ РЕЧЕННЯ

Цікал І. В.

Науковий керівник – доц., к. т. н. Бісікало О. В.

Проблема розуміння машинами «природної мови» постала чи не з їх появи. Сучасні обчислювальні можливості дозволяють застосувати широкий клас математичних методів аналізу неструктурованих даних для обробки великих масивів документів, ефективно вирішуючи завдання пошуку інформації, класифікації, кластерного аналізу, виявлення прихованих закономірностей та інші. На жаль, впровадження математичних методів в обробку тексту проходить на тлі відсталості власне лінгвістичної складової алгоритмів, що не дозволяє досягти високої якості роботи прикладних систем.

На зараз переважає думка, що лінгвістичні алгоритми є ненадійними, слабо масштабованими і надто повільними для вирішення реальних завдань. Але синтаксичний аналіз це та область прикладної лінгвістики, що розвивається. Мета такого аналізу – автоматична побудова функціонального дерева фрази, тобто знаходження взаємозалежностей між різнорівневими елементами речення. Вважається, що маючи успішно побудоване функціональне дерево фрази, можна виділити з речення смислові елементи: логічний суб'єкт, логічний предикат, прямі і непрямі додатки і різні види обставин. Існує велика кількість різних підходів до синтаксичному аналізу текстів. Але відомі алгоритми промислово реалізовані тільки в автоматичних перекладачах і до цих пір викликають безліч нарікань за недостатню точність. Тому існує проблема створення більш точного та надійного аналізатора речення, що дозволить «спілкування» з машиною природною мовою.

В основу нашого алгоритму ліг так званий асоціативний метод. При аналізі використовується контекстно-вільна граматика для формалізації природної мови. Формальна граматика або просто граматика в теорії формальних мов – це спосіб опису формальної мови, тобто виділення деякої підмножини з множини всіх слів деякого кінцевого алфавіта.

Існують два основних методи синтаксичного аналізу, а саме: висхідний і низхідний. У своїй роботі ми спробуємо поєднати їх результати, що підвищить точність аналізу.

Звісно мають місце й інші методи та алгоритми синтаксичного аналізу, яких на даний момент досить багато і вони знайшли своє використання, але в основі нашої методики лежить поняття «штучного інтелекту», тобто ми плануємо дати машині можливість «навчатись», як це робить людина. Користуючись таким підходом ми надіємось отримати хороші результати.