

## МЕТОД ПІДТРИМКИ ДІАЛОГУ НА ОСНОВІ ОБРАЗНОЇ ІНДЕКСАЦІЇ ПРИРОДНО-МОВНОГО КОНТЕНТУ

Вінницький національний технічний університет

### Анотація

Запропоновано метод підтримки діалогової системи, яка здатна давати відповіді на запитання користувача не позбавлені сенсу в залежності від вміщеного навчального контенту. Даний метод оснований на моделі бази знань із нечітким відношенням сенсу. Достовірність роботи розробленої моделі перевірено за допомогою методу експертних оцінок.

**Ключові слова:** діалогова система, навчальний контент, нечітке відношення, метод експертних оцінок.

### Abstract

The method of dialogue support system that is able to answer users' questions are not deprived of sense according to published educational content. This method is based on the model of the knowledge base of fuzzy relation of meaning. The reliability of the developed model tested by the method of expert evaluations.

**Keywords:** dialogue system, educational content, fuzzy relation, method of expert estimates.

В умовах світу, що розвивається все більшої актуальності набуває розробка діалогових систем, що могла б дати відповідь на запитання користувача у певній предметній області, максимально наближену до відповіді людини-експерта.

Метою роботи є розробка математичного забезпечення методу підтримки діалогової системи, яка має здатність перетворити певний вхідний навчальний контент на формальну базу знань, представити відповідь користувача у вигляді, сумісному з базою знань, згенерувати відповідь на надати її у вигляді максимально наближеному до відповіді людини-експерта. Достовірність роботи перевіримо за допомогою методу експертної оцінки.

При розробці даного методу потрібно брати до уваги, що система працює у межах одного навчального тексту. Для реалізації поставленої мети будемо використовувати нечіткі відношення. Вхідний текст представимо у вигляді матриці сили зв'язків між словоформами у тексті, де рядки (головне слово пари) та стовпці (підлегле слово). Особливість підходу полягає у врахуванні сили зв'язків лише між значимими словоформами. Для демонстрації наглядності дану матрицю було отримано вручну, проте існують програмні лінгвістичні технології, наприклад, DKPro [1] – у цьому випадку потрібно розраховувати на визначення зв'язків з похибкою до 10%.

Основу матрицю позначимо як  $Q$ . На основі початкової матриці із зануленими комірками, сформуємо матрицю запитання –  $R$ . Наступним кроком буде визначення матриці  $P$ , яка надасть відповідь по вхідному тексту  $Q$  на запитання користувача  $R$ . Для обробки матриць будемо використовувати функції належності, запропоновані у [2].

Для отримання відповіді застосуємо формули композиції нечітких відношень «MAX-MIN» [3]

$$\mu_P(< i_l, i_j >) = \max_{i_k \in I} \{ \min\{\mu_Q(< i_l, i_k >), \mu_R(< i_k, i_j >)\} \}. \quad (1)$$

Застосувавши дану формулу (1), ми отримаємо нову матрицю  $P$ , яка буде містити різні числові значення з  $\mu_p$  в тих комірках, які найімовірніше будуть відповіддю на запитання користувача.

Для визначення відповіді пропонується відсортувати за формальним критерієм список пар «головне та залежне слово», наприклад за силою зв'язків між відповідними словоформами. Тоді набір пар, які мають найбільше чисельне значення і перевищують певну задану межу сформує відповідь.

Експертною оцінкою вважають процедуру отримання оцінки певної задачі (проблеми) на основі думки спеціалістів (експертів) з метою подальшого прийняття рішення [4]. Існує дві групи експертних оцінок: індивідуальні оцінки (основані на використанні незалежних одна від одної думок

окремих експертів) та колективні оцінки (ґрунтуються на використанні колективної думки експертів). Розрізняють такі способи вимірювання об'єктів: ранжування; попарне порівняння; безпосередня оцінка [5,6,7].

Було проведено експеримент, в результаті якого було отримано результат. Для визначення адекватності запропонованих у [8] методу та автоматизованої системи підтримки діалогу потрібно обрати метод оцінки відповідей. Особливістю розробленої системи є те, що відповіді на питання користувачів надаються у вигляді обраних фрагментів (слів, словосполучень) з кількох речень закладеного в систему навчального текстового контенту. Проаналізувавши існуючі методи експертної оцінки релевантості відповіді, обираємо бальний метод встановлення пріоритетів. Після проведення опитування, було підтверджено високий рівень адекватності роботи методу підтримки діалогової системи на основі текстового навчального контенту.

Отримані чисельні результати оцінок експертів дають нам можливість застосування методів параметричної ідентифікації, що в свою чергу покращить результати роботи запропонованого методу.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Natural Language Processing: Integration of Automatic and Manual Analysis [Electronic resource]. – Technischen Universität Darmstadt, 2014. – Available at: \www/URL: <http://tuprints.ulb.tu-darmstadt.de/4151/1/rec-thesis-final.pdf>. – 21.04.2015.
2. Бісікало О.В. Формальні методи образного аналізу та синтезу природно-мовних конструкцій : монографія [Текст] / О. В. Бісікало // – Вінниця : ВНТУ, 2013. – 316 с. – ISBN 978-966-641-528-1.
3. Штовба С.Д. Введение в теорию нечетких множеств и нечеткую логику /С.Д. Штовба // [Електронний ресурс]. – Доступ: \www/URL: [http://matlab.exponenta.ru/fuzzylogic/book1/1\\_5.php](http://matlab.exponenta.ru/fuzzylogic/book1/1_5.php).
4. Peter D. Grogono. A Review of Expert Systems Evaluation Techniques [Електронний ресурс] / Peter D. Grogono, Alun D. Preece, Rajjan Shinghal, and Ching Y. Suen // American Association for Artificial Intelligence in technical report - american association for artificial intelligence ws. 1994. – pp.113-118 – Режим доступу: <https://www.aaai.org/Papers/Workshops/1993/WS-93-05/WS93-05-016.pdf>.
5. Wilcoxon F. Individual Comparisons by Ranking Methods [Електронний ресурс] / Frank Wilcoxon // Biometrics Bulletin, Vol. 1, No. 6. – 1945. – pp. 80-83. – Режим доступу: <http://sci2s.ugr.es/keel/pdf/algorithm/articulo/wilcoxon1945.pdf>.
6. Метод попарних порівнянь (ПС). Лекція №6. [Електронний ресурс] – Режим доступу: <http://ecsocman.hse.ru/data/466/641/1219/ch6.pdf>.
7. Allen M.J. Direct Method to Assess Student Learning [Електронний ресурс] / Mary J. Allen // WASC/AAHE Collaborative Workshop on Building Learner-Centered Institutions-Developing Institutional Strategies for Accessing and Improving Student Learning. – 2004. – 8 p. – Режим доступу: <http://programs.honolulu.hawaii.edu/intranet/sites/programs.honolulu.hawaii.edu/intranet/files/assessment-methods.pdf>.
8. Лісовенко А.І. Підтримка діалогу з навчальним контентом [Електронний ресурс] / А.І. Лісовенко, О.В. Бісікало, О.В. Яхимович, С.С. Траченко // Адаптивні технології управління навчанням: матеріали першої міжнародної конференції. Одеса, 23-25 вересня 2015 р. – Одеса, 2015. – С. 97-100. – Режим доступу: <https://drive.google.com/file/d/0B34KZFqaGoAyOVRGTlc2cIJ0eUk/view>.

**Лісовенко Анна Ігорівна** – аспірант факультету комп'ютерних систем та автоматики, Вінницького національного технічного університету, Вінниця, e-mail: [sweettee@ukr.net](mailto:sweettee@ukr.net).

Науковий керівник: **Бісікало Олег Володимирович** – д-р технічних наук, професор, декан факультету комп'ютерних систем та автоматики, Вінницького національного технічного університету, м. Вінниця.

**Lisovenko Anna I.** – graduate student of faculty of Computer Control Systems and Automatic Equipment, Vinnitsa National Technical University, Vinnitsa, e-mail: [sweettee@ukr.net](mailto:sweettee@ukr.net).

Supervisor: Bisikalo Oleg V. – Dr.Sc. (Eng.), Professor, Dean of faculty of Computer Control Systems and Automatic Equipment, Vinnitsa National Technical University, Vinnitsa.