

ся: заголовок, який презентує зміст вихідного тексту в концентрованому вигляді; текстова база, яка містить речення, що є «інформаційним ядром» тексту і утримує інформацію, залежну від тематики тексту; онтології, що містять незалежну від тематики тексту інформацію: онтології верхнього рівня (набір змістових категорій, котрі входять до реферату), онтології загальнонаукової лексики (необхідної для синтезу реферативних речень) й онтологій предметних галузей (необхідних для аналізу змісту тексту).

Матеріалом дослідження було обрано реферативні конструкції, які будують тексти індикативних рефератів з чотирьох предметних галузей: економіка, медицина, прикладна лінгвістика та соціоніка. У цілому проаналізовано лінгвістичні характеристики приблизно 5000 реферативних конструкцій та відповідні їм первинні тексти.

Перелік посилань

Панченко Д. І. Моделювання семантичних зв'язків «Текст – Реферат» у системах автоматичного реферування : автореф. дис.... канд. філол. наук : спец. 10.02.21 / Д. І. Панченко. – Х. : Вид-во НУА, 2012.

ЛИНГВИСТИЧЕСКИЕ ПРОБЛЕМЫ ИСПОЛЬЗОВАНИЯ ОТКРЫТОЙ ФОРМЫ ТЕСТОВЫХ ЗАДАНИЙ

Петрасова С. В., Хайрова Н. Ф.

*Национальный технический университет
«Харьковский политехнический институт»,
г. Харьков, ул. Пушкинская, 79/2, тел. 707-63-60,
e-mail: svetapetrasova@gmail.com, nina_khajrova@yahoo.com*

Под образовательной технологией мы понимаем систему научной организации обучения и контроля, создаваемую на основе достижений педагогики, применения новых методов научной организации учебного процесса, технических средств обучения, компьютерных форм организации работы при обучении и проверке знаний [1].

Контроль уровня знаний является неотъемлемой частью процесса обучения. Он обеспечивает обратную связь в системе

«обучаемый – педагог».

Применение заданий в тестовой форме в сочетании с новыми образовательными технологиями позволяет существенно улучшить качество учебного процесса благодаря реализации обучающей, контролирующей, организующей, диагностирующей, воспитательной и мотивирующей функций.

Среди положительных качеств тестирования можно выделить:

- объективность в оценке знаний контролируемого на основе обработки результатов контроля;
- возможность обработки большого количества материала и охвата большего количества испытуемых;
- качественный анализ программ, учебников, тем на основе учета требований к тестам;
- диагностичность теста (предупреждение неуспеваемости, отставания, улучшение методики преподавания и т. д.).

Вместе с тем тестированию, в том числе компьютерному, присущи недостатки, которые можно свести к следующим:

- тесты фиксируют только конечный результат выполнения задания;
- тесты не выявляют причины допущения ошибок (правильный ответ может быть угадан из числа предложенных; неверный ответ может быть обусловлен неправильным прочтением);
- сложность использования открытой формы ответов, связанная с проблемами автоматической обработки ответов на естественном языке, включающими сложность кореферентных связей, выявление сущностей, иерархических отношений, семантических эквивалентов и т. д. [2].

В данном исследовании с целью улучшения качества тестирующих систем и, в частности, разрешения проблемы синонимии предлагается вариант построения семантической сети, включающей отношение семантической эквивалентности.

Семантическая сеть является наиболее перспективным способом формального выражения знаний. Это обусловлено, прежде всего, наглядностью представления знаний в семантических сетях. Также данный способ представления знаний учитывает форму утверждения его семантику [3].

В своем подавляющем большинстве семантические сети разрабатываются экспертами в той или иной предметной области «вручную». Тогда как одним из наиболее полных источников знаний являются такие универсальные средства представления, накопления и передачи информации в человеческом обществе, как

тексты. При этом среди всех текстовых источников толковые словари представляют тексты естественного языка, пожалуй, с наиболее концентрированной смысловой нагрузкой.

В данном исследовании предлагается метод автоматизированного построения фрагмента семантической сети на базе использования знаний толкового словаря английского языка. В разработанной сети определяются отношения классификации, гипонимии, меронимии и гиперонимии и отношения семантической эквивалентности.

Для формализации отношений семантической эквивалентности определяется мера семантической близости, формально определяемая дефинициями глоссариев как отношение мощностей множеств, образованных теоретико-множественным пересечением и объединением множеств терминов дефиниций [4].

В результате исследования разработан и программно реализован алгоритм автоматизированного построения семантической сети, выявляющий вышеприведенные отношения между терминами толкового словаря.

Для определения эффективности разработанного информационно-лингвистического обеспечения исследовалась выборка из ста терминов словаря Microsoft Computer Dictionary, 5th edition. После обработки результатов были определены коэффициенты точности, полноты и ошибки.

Разработанная программная имплементация может быть использована в качестве лингвистической базы для системы тестирования, благодаря которой решается проблема синонимии ответов на естественном языке.

Список литературы

1. *Кабанова Т. А.* Тестирование в современном образовании : учеб. пособие / Т. А. Кабанова, В. А. Новиков. – М. : Высшая школа, 2010. – 381с.
2. *Батешов Е. А.* Основы технологизации компьютерного тестирования : учеб. пособие. – Астана : ТОО «Полиграф-мир», 2011. – 241 с.
3. *Маннинг Кристофер Д.* Введение в информационный поиск : пер. с англ. / Кристофер Д. Маннинг, Прабхакар Рагхаван, Хайнрих Шютце. – М. : ООО «И.Д.Вильямс», 2011. – 528 с.
4. *Метод автоматизированного построения семантической сети терминов учебной дисциплины / Л. А. Федорченко,*

Н. Ф. Хайрова, А. И. Довнаръ, С. О. Булгаков // *Радіоелектронні і комп'ютерні системи.* – 2011. – № 4 (52). – С. 1–7.

СИСТЕМИ ПОШУКУ ПЛАГІАТУ В УМОВАХ ІНТЕРНЕТ-ПРОСТОРУ

Побіженко В. В., Побіженко І. О.

*Харківська державна академія культури
м. Харків, Бурсацький узвіз, 4, тел. 731-32-82,
e-mail: vlad1603@mail.ru*

Особливе місце в змінах соціальних систем останньої чверті ХХ-го – початку ХХІ століть посідає перехід до нової системи суспільних відносин, заснованої на цінності кожної людини. Суспільство висуває нові вимоги до особистості, детермінуючи той набір життєвих стратегій та якостей, що дозволить їй максимально реалізувати свій потенціал.

Формування й розвиток творчого потенціалу та підвищення активності молоді пов'язані з функціонуванням соціальних інститутів, особливе місце серед яких посідає освіта. Від неї залежать можливості здобуття знань, саморозвитку, самодіяльності, актуалізації інтелектуального потенціалу майбутнього фахівця.

Водночас сучасне інформаційно-комунікаційне середовище містить багато спокус для молоді, серед яких спокуса позичити (скопювати) чужу наукову працю без посилань. Для більшості студентів написання самостійних робіт зводиться до суто технічного пошуку інформації: копіювання чи сканування матеріалу з підручника, посібника, статті (зокрема і з Інтернету) із подальшим його роздрукуванням. При цьому, автор не турбує себе навіть редагуванням тексту, написанням речень, які встановлюють зв'язок між окремими частинами запозиченого тексту, а іноді навіть не читає власного «твору». Нині комп'ютерні технології розширили межі плагіату: потрібну роботу можна знайти в Інтернеті.

Існують різні методи пошуку плагіату в наукових документах, які ґрунтуються на загальних методах пошуку інформації. Перший метод пошуку рядка розробили Майкл Рабин та Річард Карп у 1987 році. Цей метод шукає шаблон, який у тексті використовує хешування.

Найвідомішим методом обробки дублікатів у веб-пошуці, ретельно викладеним Андрієм Бродером зі співавторами в 1997 році, є метод «шинглів». Для підвищення ймовірності того, щоб у результаті невеликої зміни тексту контрольна сума не змінилася, можна спробувати вибрати з тексту декілька підстрок. Шингл (від