

## ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ПЕРЕВІРКИ ВІДПОВІДЕЙ В ІНТЕЛЕКТУАЛЬНІЙ АВТОМАТИЗОВАНІЙ СИСТЕМІ КОНТРОЛЮ ЗНАНЬ

<sup>1</sup>Миколаївський національний університет імені В.О. Сухомлинського

Досліджено процес перевірки відповідей в системах тестування на запитання «відкритого» типу (коротка текстова відповідь, есе). Проведено аналіз автоматизованих систем, в яких реалізовано цей процес. Визначено обмеження їх застосування для контролю рівня знань з технічних дисциплін. Аргументовано актуальність досліджень щодо застосування методів «нечіткого» порівняння рядків в задачах перевірки відповідей, поданих у довільній текстовій формі. Запропоновано інформаційну технологію перевірки таких відповідей інтелектуальної автоматизованої системи контролю рівня знань у вигляді структурної поетапної схеми процесу обробки вхідної відповіді. Наведено обчислювальні алгоритми кожного етапу обробки, виконано їх реалізацію мовою С#. На першому етапі виконується «нормалізація» слів відповіді та еталону (правильної відповіді на запитання, яка зберігається в базі даних). Результатом першого етапу є рядкові масиви: перший — набір слів відповіді, другий — слова еталону (не враховуються слова довжиною меншою за чотири символи). Другий етап передбачає циклічний виклик функції знаходження довжини найбільшої загальної підпоследовності слів (LCS) масивів відповідей та еталону. Подано блок-схему процедури обчислення LCS, яка базується на рекурсивному алгоритмі, запропонованому Хіршбергом. Порівняння кожного слова вхідної відповіді з усіма словами еталону, включаючи синоніми, дозволяє знайти довжину LCS навіть у разі незбігу порядку слів у відповіді та еталоні, що є перевагою запропонованого підходу. На третьому етапі відбувається обчислення загального показника подібності відповіді та еталону шляхом знаходження суми LCS окремих слів. Останній етап — формування результату перевірки відповіді на основі значення показника подібності (встановлюється в залежності від вимог: високий рівень збігу — 50 %, достатній — 30 %, низький — 10 %). Проведено випробування запропонованої інформаційної технології обробки відповідей, поданих у довільній текстовій формі. Виконано порівняння отриманих результатів з відстанню Левенштейна та латентно-семантичним аналізом. Запропонована ІТ дає кращий результат під час перевірки відповідей, в яких вживаються слова синоніми.

Кількість неправильних спрацювань для 50 варіантів відповідей різної довжини (від 10 до 200 символів) склала 4 %. В результаті випробування встановлено рекомендовану довжину відповіді та еталону (не більше 200 символів), за якої досягається найбільша точність. Сформовано напрямки майбутніх досліджень: підвищення ефективності алгоритму шляхом введення додаткового етапу обробки, визначення «загальної» міри подібності відповіді та еталону на основі коефіцієнта Жаккара; реалізація автоматизованої інтелектуальної системи контролю рівня знань на базі технології клієнт-сервер; приведення результату перевірки у відносну шкалу оцінювання рівня знань.

**Ключові слова:** інтелектуальна система, тестовий контроль знань, відповіді у довільній текстовій формі, «нечітке» порівняння рядків, довжина найбільшої загальної підпоследовності, алгоритм Хіршберга.

### Вступ

Стрімкий розвиток інформаційних технологій, інформатизація системи освіти, розповсюдження дистанційного навчання визначають актуальність розробки автоматизованих систем для контролю рівня знань, які передбачають інтелектуалізацію процесу тестування.

Однією з визначальних ознак інтелектуальної системи є вміння розв'язувати складні задачі, які погано формалізуються, для яких можуть бути характерні невизначеність та динамічність вихідних даних і знань. Оскільки процес перевірки відповідей «відкритого» типу («запитання-коротка відповідь», «есе»), які передбачають отримання від користувача введення відповіді у довільній текстовій формі, відноситься до слабкоформалізованої задачі, то систему, в якій реалізується цей

процес, можна вважати інтелектуальною.

Більшість існуючих систем тестування реалізують завдання «закритої» форми (вибір одної або декількох правильних відповідей з множини, встановлення послідовності, зіставлення тощо). В системах, які підтримують перевірку відповідей на запитання «відкритого» типу (Moodle [1], SunRay TestOfficePro [2]), встановлення правильності отриманих даних виконується із залученням експерта (викладача) або передбачає наявність точного збігу з еталонним варіантом відповіді. Новітньою інтелектуальною системою тестування є платформа «WriteToLearn» [3] для оцінювання письмових відповідей у вигляді висновків, резюме, есе, переказів, використовуючи техніку узагальнення та механізмів зворотного зв'язку з метою розвитку таких умінь як «розуміння читання» та «навички письма». Її технологічною основою є латентний семантичний аналіз (LSA), статистична теорія вивчення мови та комп'ютерна модель, яка вимірює смислову схожість слів та документів з точністю, що наближаються до людських суджень. Крім LSA, «WriteToLearn» використовує статистичну класифікацію, методи аналізу, обчислювальної лінгвістики, коректорів орфографії та граматики. Схожою є лінгвістична система Versant [4], яка використовує латентно-семантичний аналіз для вивчення значення слів та аналізує кожний текст на предмет вмісту, стилю та механіки письма. Системи «WriteToLearn», «Versant» мають певну специфіку використання, вони є лінгвістичними, спрямованими на вивчення мов, формування стилістичної вправності написання текстів, що визначає обмеженість їх застосування в якості інтелектуальних систем оцінки рівня знань з технічних дисциплін. Останні вимагають практичної спрямованості відповіді, самостійності в конструюванні відповіді (рішення): відтворити визначення, надати характеристику, коротке пояснення, написати рівняння, теорему, блок програмного коду тощо.

Отже, актуальним є завдання розробки системи контролю рівня знань з технічних дисциплін, яка б виконувала автоматизовану обробку текстових відповідей «відкритого» типу.

Дослідженнями питань розробки, застосування автоматизованих інтелектуальних систем контролю рівня знань займалися науковці: О. І. Комарницька, Ю. Ю. Черепанова, Ю. П. Кондратенко, І. Д. Рудінський, А. Varlybayev, A. Sharipbay, H. Dreher, R. Williams, D. Pérez-Marín та інші [5]—[12].

О. І. Комарницькою в інтелектуальній системі оцінки знань застосовано метод латентного семантичного аналізу, який дозволяє на підставі оцінки кореляції між словами та текстами зробити висновок про ступінь близькості змісту цих слів чи групи слів [5]. Використано метрику Левенштейна, що дозволило встановити ступінь відповідності тексту відповіді тому тексту, що міститься у базі даних предметної сфери. Для формування загальної оцінки відповіді на тестові завдання автором запропоновано комплексний показник, у якому враховано наявність у відповіді слів, присутніх у зразку, відповідність структур зразка і відповіді (порядку слів).

Ю. Ю. Черепанова запропонувала метод тестування теоретичних знань, оснований на побудові та аналізі концептуальних моделей відповідей з використанням тезауруса семантичних полів [6]. Запропонована автором структура тезауруса дозволяє використати метод аналізу правильності відповіді за ключовими словами за рахунок автоматизованої побудови списку ключових слів. При цьому як ключові слова можуть бути використані парадигматичні поля слів, що входять в еталон відповіді.

В зазначених роботах малодослідженим є застосування методів «нечіткого» порівняння рядків для оцінки відповідей «відкритої форми». За результатами попередніх досліджень [13], [14], визначено, що використання таких методів в системах тестування дозволяє ефективно здійснювати перевірку відповідей, поданих у текстовій формі шляхом обчислення загальної найбільшої підпоследовності.

*Метою роботи* є розробка інформаційної технології перевірки відповідей в інтелектуальній автоматизованій системі контролю рівня знань. Необхідно сформулювати інформаційну технологію процесу тестування шляхом описання обчислювальних алгоритмів кожного етапу обробки вхідної відповіді.

### Постановка задачі

Нехай  $X[n]$  масив слів відповіді, поданої у текстовій формі, — масив слів еталону (правильної відповіді). Необхідно встановити, чи надана користувачем відповідь у довільній текстовій формі відповідає еталону, оцінити ступінь «схожості» масивів та прийняти рішення на основі загального показника «подібності».

Для розв'язання поставленої задачі запропоновано «порівнювати» масиви  $X[n]$  та  $Y[m]$  шля-

хом визначення довжини найбільшої спільної підпоследовності (longest common subsequence, LCS) кожного слова масиву  $X[n]$  з усіма елементами масиву  $Y[m]$ . Якщо сума усіх довжин LCS для  $x_i, i=1\dots n$  та  $y_j, j=1\dots m$  є більшою або дорівнює  $k$  відсотків довжини еталону, то відповідь вважається правильною

$$\sum_{i=1}^n \sum_{j=1}^m LCS(x_i, y_j) \geq \frac{k}{100\%} \sum_{j=1}^m length(y_j).$$

Значення  $k$  встановлюється в залежності від вимог (високий рівень збігу — 50 %, достатній — 30 %, низький — 10 %)

Об'єкт дослідження — процес перевірки відповіді, поданої у довільній текстовій формі. Предмет дослідження — інформаційна технологія для реалізації процесу автоматизованої перевірки відповідей «відкритого» типу. Методи дослідження — методи аналізу та порівняння, формалізації, «нечіткого» порівняння рядків, динамічного програмування.

### Результати дослідження

Інформаційну технологію процесу перевірки відповіді в інтелектуальній системі контролю рівня знань показано на рис. 1.

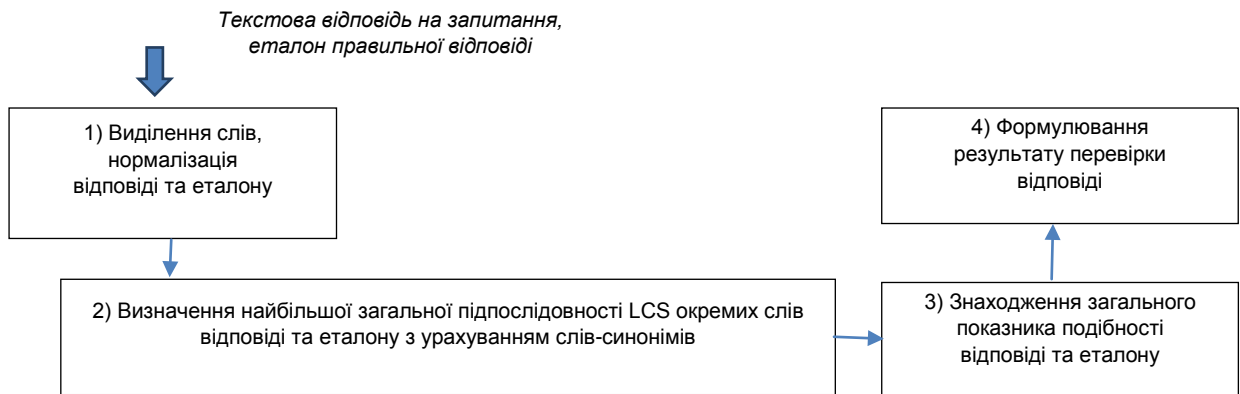


Рис. 1. Інформаційна технологія процесу перевірки відповіді в інтелектуальній системі контролю рівня знань

Перший етап «Виділення слів, нормалізація відповіді та еталону» застосовується для розбиття відповіді та еталону на масив окремих слів, приведення їх в єдиний «нижній» регістр, видалення всіх символів відмінних від пробілу, цифр та букв. «Словами» вважаються всі последовності символів без пробілів, які мають довжину  $\geq 4$ . Цим самим видаляються майже всі прийменники, сполучники тощо.

Лістинг коду етапу нормалізації мовою C#:

```

public string[] Normalize(string text)
{ var res_string = new StringBuilder(200);
  var words = new List<string>();
  var res_register = text.ToLower();
  foreach (var s in res_register)
  { if ((s == ' ') || (char.IsLetterOrDigit(s) == true))
    res_string.Append(s); }
  var word = res_string.ToString().Split(' ');
  foreach (var w in word)
  { if (w.Length >= 4) words.Add(w);
  }
  return words.ToArray();}
  
```

Другий етап «Визначення найбільшої загальної підпоследовності LCS окремих слів відповіді та еталону» базується на основі модифікації методу динамічного програмування пошуку довжини найбільшої загальної підпоследовності, запропонованого Хіршбергом [15]—[17]. Блок-схему алгоритму визначення довжини найбільшої загальної підпоследовності відповіді та еталону показано на рис. 2.

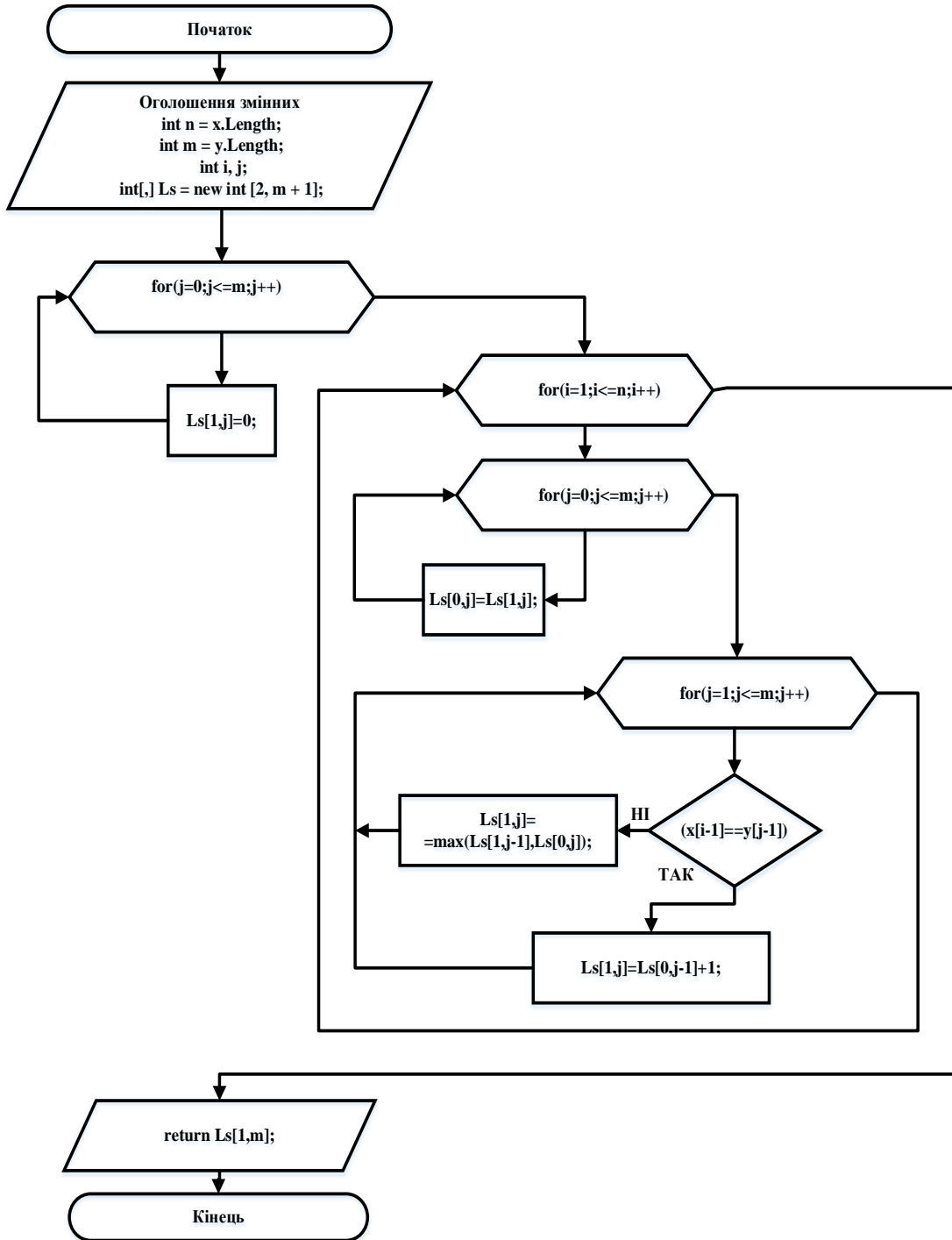


Рис. 2. Блок-схема алгоритму визначення довжини найбільшої загальної підпослідовності відповіді та еталону

Перевагою методу Хіршберга у порівнянні з методом Вагнера–Фішера, який застосовується для обчислення відстані Левенштейна, є лінійна просторова складність. Метод Вагнера–Фішера вимагає квадратичних витрат часу та пам’яті, тобто часова та просторова складність складає  $\Theta(n \cdot m)$ . Метод Хіршберга квадратичний за просторовою складністю  $\Theta(n \cdot m)$ , але лінійний відносно витрат пам’яті  $\Theta(n + m)$ . Ефективність цього методу з погляду просторової складності зумовило його вибір для запропонованої інформаційної технології. Метод Хіршберга дозволяє ітеративно обчислювати довжини префіксів рядків, які послідовно збільшуються, а не відстань між ними.

Оскільки порядок слів у відповіді на запитання та еталоні може відрізнятись, здійснюється перевірка кожного слова відповіді з усіма словами еталону. В результаті порівняння повертається максимальне значення довжини найбільшої загальної підпослідовності слова відповіді та усіх слів

еталону. Перевагою запропонованої інформаційної технології є врахування синонімів, які вказуються в еталоні в дужках після основного слова. Це дозволило підвищити ефективність процесу перевірки у разі надання користувачем відповіді у вигляді близьких за змістом слів. При цьому до загального розміру рядка еталону довжина синонімів не враховується, оскільки саме на основі порівняння значення LCS з довжиною еталона приймається рішення щодо класифікації відповіді.

На третьому етапі «Знаходження загального показника подібності відповіді та еталону» визначається довжина загальної підпоследовності для всього речення шляхом додавання LCS усіх слів. Четвертий етап «Формулювання результату перевірки» відповіді передбачає порівняння довжини загальної підпоследовності всієї відповіді із довжиною еталону.

Якщо довжина загальної підпоследовності відповіді більше або дорівнює  $k$  довжини еталону, то надана відповідь вважається правильною (значення показнику подібності можна встановити від 10 % до 50 %). Лістинг коду четвертого етапу мовою C#:

```
public void Porivnya_Click(object sender, EventArgs e)
{
    string[] mas_vid = Normalize(vidpovid.Text);
    string[] mas_etalon = Normalize(etalon.Text);
    int razm_stroki = etalon.Text.Length;
    int kil = 0, sum_kil=0, k=2, max_kil;
    for (int i=0; i< mas_vid.Length; i++)
    {
        max_kil = 0;
        for (int j = 0; j < mas_etalon.Length; j++)
        {
            kil = FindLS(mas_vid[i], mas_etalon[j]);
            if (max_kil < kil) max_kil = kil;
        }
        if (max_kil>=3) ( sum_kil += max_kil);
    }
    if (sum_kil>= razm_stroki/k) result.Text="відповідь вірна";}

```

Для тестування ІТ розроблено додаток «FuzzyComparer», інтерфейс головної форми якого зображено на рис. 3.

В табл. 1 подано результати тестування модуля перевірки відповідей автоматизованої інтелектуальної системи контролю рівня знань.

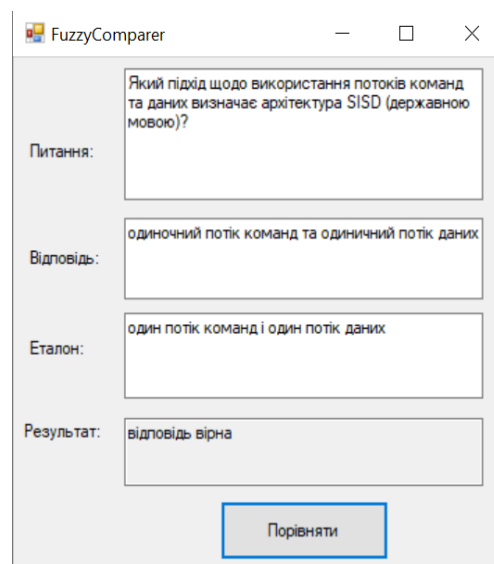


Рис. 3. Інтерфейс головної форми додатку «FuzzyComparer»

Таблиця 1

Результати тестування ІТ перевірки відповідей

Запитання	Answer	Pattern	LD	LCS	$\overline{LD}$	$LCS_{\overline{LD}}$	LSA	$\alpha$	$\Delta$
Назвіть комп'ютер і ПЗ, які є частиною ЛКМ і підтримують ефективне звернення комп'ютерів локальної мережі до Інтернету.	сервер проксі, (13)	Проксі-сервер, (13)	13	12	0	+12	1	0,92	-0,08
Який підхід щодо використання потоків команд та даних визначає архітектура SISD (державною мовою)?	одиночний потік команд та одиночний потік даних, (47)	один потік команд і один потік даних, (36)	12	29	24	+5	0,88	0,81	-0,07
Яка функція в Windows відповідає за створення потоку (у мові C++)?	createthraed, (12)	CreateThread(), (14)	4	10	10	0	0,65	0,71	+0,06
Призначення механізму синхронізації процесів і потоків	узгодження роботи потоків під час звернення до загальних ресурсів, (65)	організація узгодженої (дружньої) роботи з загальними ресурсами, (48)	44	31	4	+27	0,75	0,65	-0,1
Що таке мережа Інтернет?	всесвітня мережа до складу якої входять локальні, глобальні та інші мережі, (74)	глобальна (всесвітня) мережа, яка є сполученням (об'єднанням) локальних, регіональних і глобальних мереж, (76)	56	41	20	+21	0,4	0,54	+0,14

Продовження табл. 1

Запитання	Answer	Pattern	LD	LCS	$\overline{LD}$	$LCS_{\overline{LD}}$	LSA	$\alpha$	$\Delta$
Що таке багатозадачність, види багатозадачності?	Одночасне виконання двох та більше задач в рамках операційної системи. Класифікація видів багатозадачності: кооперативна, істина, (128)	Властивість (можливість) операційної системи або середовища програмування одночасного (паралельного) виконання декількох (двох та більше) потоків (процесів, задач). Види (класифікація) багатозадачності: істинна, невітисняюча (кооперативна), витісняюча, (141)	103	105	45	+58	0,43	0,74	+0,31

*Примітки:* Answer — довжина рядка текстової відповідь, наданої користувачем; Pattern — довжина рядка правильної відповіді (еталону); LD — значення відстані Левенштейна; LCS — значення LCS, визначене за запропонованою ІТ;  $\overline{LD} = Pattern - LD$ ;  $LCS_{\overline{LD}} = LCS - \overline{LD}$ ; LSA — коефіцієнт узгодженості відповіді та еталону за технологією LSA, діапазон [-1; 1];  $\alpha = LCS/pattern$  — коефіцієнт узгодженості для запропонованої ІТ;  $\Delta = \alpha - LCS$ .

Для порівняння в табл. 1 подані значення обчислення відстані Левенштейна для тих же вихідних даних (знайдено за допомогою онлайн ресурсу <https://calculatorium.ru/text/levenshtein-distance>). Оскільки LCS та редакційна відстань пов'язані між собою (чим більше LCS, тим менше операцій потрібно для перетворення), то для порівняння результатів вводиться стовпець (четвертий за порядковим номером), який містить різницю довжини еталону та відстані Левенштейна. Значення цього стовпця порівнюються зі значенням LCS, яке обчислено з використанням запропонованої ІТ. Додатне значення («+» в таблиці) показує на скільки символів отриманий результат більше (менше, якщо «-») за LCS, обчислену на основі редакційної відстані. Коефіцієнт узгодженості відповіді та еталону за технологією LSA обчислено з використанням онлайн ресурсу <http://lsa.colorado.edu/> (Sentence Comparison).

За результатами порівняння отриманих результатів на основі запропонованої ІТ та відстані Левенштейна можна зробити висновки, щодо більшої ефективності ІТ у разі зміни порядку слів та застосування синонімів у відповіді. Щодо відстані Левенштейна, то висновки такі:

- 1) не враховується порядок слів: під час перестановки місцями слів відстань порівняно велика;
- 2) відстань між короткими, але геть різними словами є невеликою, але для схожих довгих слів є незначною.

В порівнянні з коефіцієнтом узгодженості, отриманим на основі методу латентно-семантичного аналізу (LSA), ІТ дає кращий результат під час перевірки відповідей, в яких вживаються слова-синоніми.

Для оцінки достовірності застосовано критерій узгодженості Пірсона  $\chi^2$ . За нульову гіпотезу береться твердження, що немає істотних відмінностей між оцінками відповідей, отриманих з використанням системи, та оцінками викладача. Обсяг вибірки — 50 відповідей, два класи — правильна, неправильна відповідь. Розподіл відповідей подано в табл. 2.

Таблиця 2

Розподіл результатів перевірки відповідей системою та експертом

Відповідь	Оцінки, отримані з використанням системи	Оцінки викладачів
Правильна	34	38
Неправильна	16	12

Отримано значення  $\chi^2 = 1,75$ . Беручи рівень значущості рівним 0,05 (5%) та враховуючи значення ступенів вільності  $n^2 = 1$ , отримаємо критичне значення  $\chi_{кр}^2 = 3,84$ . Оскільки  $\chi^2 < \chi_{кр}^2$ , то нульову гіпотезу приймаємо на рівні значущості 0,05.

Найвищу точність та найменше число помилок отримано, якщо під час порівняння довжина рядка відповіді та еталону була меншою за 200 символів.

Отже, за результатами досліджень розроблено та протестовано інформаційну технологію перевірки відповідей, поданих у довільній текстовій формі, перевагою якої є:

1) застосування методу Хіршберга, а не відстані Левенштейна, забезпечило зменшення вимог відносно витрат пам'яті. Тому цей підхід ефективніший для перевірки великих речень, текстів. Порівнюючи метод Хіршберга з іншими методами обчислення LCS (Хатна–Шиманського, Укко-нена–Майерса [15]), констатуємо, що перевагою метода Хіршберга є простота його реалізації.

2) врахування слів-синонімів під час перевірки відповіді, що дозволяє підвищити результативність оцінювання. В існуючих підходах, зокрема в роботі [5], для врахування слів-синонімів передбачається одночасне зберігання декількох можливих подібних варіантів відповідей. Запропонований підхід щодо зазначення слів-синонімів у дужках після основного слова в одному варіанті відповіді дозволяє скоротити час наповнення бази правильних відповідей та, відповідно, час перевірки наданої відповіді.

3) врахування можливої зміни порядку слів у відповіді на запитання та еталоні. Пропонується перевіряти кожне слово відповіді з усіма словами еталону. В результаті порівняння повертається максимальне значення довжини найбільшої загальної підпоследовності одного слова відповіді та усіх слів еталону.

### Висновки

Запропоновано інформаційну технологію перевірки відповідей поданих у довільній текстовій формі інтелектуальної автоматизованої системи контролю рівня знань. Структуровано етапи обробки вхідної відповіді, представлено обчислювальні алгоритми їх виконання мовою C#. Реалізовано інформаційну технологію обробки «відкритих» відповідей, яка базується на рекурсивному алгоритмі визначення довжини найбільшої загальної підпоследовності, наданої користувачем відповіді, та зразка правильної відповіді, що зберігається в базі даних. Перевагою запропонованого підходу є циклічне порівняння кожного слова вхідної відповіді з усіма словами еталону, що дозволяє знайти довжину LCS навіть за незбігу порядку слів у відповіді та еталоні. Проведено тестування ІТ, в результаті якого визначено рекомендовану довжину відповіді та еталону ( $\leq 200$  символів).

Майбутні дослідження будуть спрямовані на тестування та удосконалення запропонованої інформаційної технології: підвищення ефективності алгоритму шляхом введення додаткового етапу обробки, визначення «загальної» міри подібності відповіді та еталону на основі коефіцієнта Жаккара; реалізацію автоматизованої інтелектуальної системи контролю рівня знань на базі технології клієнт-сервер; приведення результату перевірки у відносну шкалу оцінювання рівня знань.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] *Modular Object Oriented Dynamic Learning Environment* [Online]. Available: <https://moodle.org>.
- [2] SunRav TestOfficePro. Програма для создания тестов [Электронный ресурс]. Режим доступа: <http://sunrav.ru/testofficepro.html>.
- [3] WriteToLearn is a web-based tool for building writing skills and reading comprehension in grades 4-12. [Online]. Available: <https://www.pearsonassessments.com/professional-assessments/products/programs/write-to-learn.html>.
- [4] Versant. Automated language tests for recruitment, training and student placement [Online]. Available: <https://www.pearson.com/english/versant.html>.
- [5] О. І. Комарницька, «Метод семантичного порівняння нечіткої текстової інформації», *Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка*, вип. 47, с. 127-132, 2014. [Електронний ресурс]. Режим доступу: [http://nbuv.gov.ua/UJRN/Znrviknu\\_2014\\_47\\_22](http://nbuv.gov.ua/UJRN/Znrviknu_2014_47_22). Дата звернення: Квіт. 21, 2020.
- [6] Ю. Ю. Черепанова, «Контроль знаний с ответами на естественном языке», *Восточно-европейский журнал передовых технологий. Информационные технологии*, № 4/2(40)2009, с. 32-36, 2009. [Электронный ресурс]. Режим доступа: <http://journals.urau.ru/eejet/article/view/20825/18401>. Дата обращения: Апрель 21, 2020.
- [7] Ю. П. Кондратенко, и С. О. Волкова, «Програмный комплекс для автоматизованого тестування знань студентів», *Технічні вісті*, № 1(22), 2(23), с. 32-36, 2006.
- [8] И. Д. Рудинский, *Основы формально-структурного моделирования систем обучения и автоматизации тестирования знаний (монография)*. Москва, РФ: Горячая линия-Телеком, 2004, 204 с.
- [9] A. Barlybayev, and A. Sharipbay, "An intelligent system for learning, controlling and assessment knowledge", *Information*. vol. 18, no 5(A), pp.1817-1827, 2015.
- [10] Heinz Dreher, "Interactive On-line Formative Evaluation of Student Assignments," *Issues in Informing Science and Information Technology*, vol. 3, pp. 189-197, 2006. <https://doi.org/10.28945/882>.
- [11] R. Williams, and H. Dreher, "Formative assessment visual feedback in computer graded essays," *Issues in Informing Science and Information Technology*, vol. 2, pp. 23-32, 2005. <https://doi.org/10.28945/808>.
- [12] Olga C. Santosa, Jesus G. Boticario, and Diana Pérez-Marín, "Extending web-based educational systems with personalised support through User Centred Designed recommendations along the e-learning life cycle," *Science of Computer Programming*, vol. 88, pp. 92-109, 2014. <https://doi.org/10.1016/j.scico.2013.12.004>.

[13] К. Т. Кузьма, «Аналіз методів перевірки відповіді в системах тестування, поданої в текстовій формі,» *Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки*, т. 29 (68) № 1, ч. 1, с. 163-167, 2019. [Електронний ресурс]. Режим доступу: [http://www.tech.vernadskyjournals.in.ua/journals/2018/1\\_2018/part\\_1/30.pdf](http://www.tech.vernadskyjournals.in.ua/journals/2018/1_2018/part_1/30.pdf). Дата звернення: Квіт. 21, 2020.

[14] К. Т. Кузьма, і О. В. Мельник, «Обчислювальна технологія перевірки відповідей у системах тестування,» *Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки*, т. 31 (70) № 1, ч. 1, с. 85-88. 2020, <https://doi.org/10.32838/2663-5941/2020.1-1/15>.

[15] Б. Смит, *Методы и алгоритмы вычислений на строках: пер. с англ.* Москва, Россия, РФ: ООО «И.Д. Вильямс», 2006, 496 с.

[16] D. S. Hirschberg, "A linear space algorithm for computing maximal common subsequences," *Communications of the ACM*, vol. 18, no. 6, pp. 341-343, June, 1975. <https://doi.org/10.1145/360825.360861>.

[17] Orgad Keller, Tsvi Kopelowitz, and Moshe Lewenstein, "On the longest common parameterized subsequence," *Theoretical Computer Science*, vol. 410, Issue 51, pp.5347-5353, November, 2009, <https://doi.org/10.1016/j.tcs.2009.09.011>

Рекомендована кафедрою автоматизації та інтелектуальних інформаційних технологій ВНТУ

Стаття надійшла до редакції 5.05.2020

**Кузьма Катерина Теодозіївна** — канд. техн. наук, доцент кафедри інформаційних технологій, e-mail: [katushke2017@gmail.com](mailto:katushke2017@gmail.com).

Миколаївський національний університет імені В.О. Сухомлинського, Миколаїв

**К. Т. Kuzma<sup>1</sup>**

## Information Technology for Verification Answers in the Intellectual Automated Knowledge Control System

<sup>1</sup>Mykolaiv V. O. Sukhomlynskyi National University

*The process of verification the answers to the "open" type questions (short text answer, essay) in the testing systems was researched. The analysis of automated systems in which this process is implemented has been carried out. The limitations of their application for control the level of knowledge on technical disciplines were determined. There has been justified the relevance of research on the "fuzzy" string comparison methods in the problems of verification of answers, submitted in an arbitrary text form. The functional structure of the module for verification such answers in intelligent automated system for control the level of knowledge in the form of a structural step-by-step diagram of the process of testing the input answer was proposed. The computational algorithms of each stage of processing are given; their implementation in C# was executed. The first step is to "normalize" the words of the answer and the etalon (the correct answer to the question stored in the database). The result of the first stage is a string arrays: the first is a set of response words, the second is a set of etalon words (words less than four characters in length are not included). The second step involves the cyclic calling the function of finding the length of the longest common subsequence (LCS) of the words of arrays of the response and the etalon. A block diagram of the LCS calculation procedure based on the recursive algorithm proposed by Hirschberg was presented. Comparison of each word of the input answer with all the words of the etalon, including synonyms, makes it possible to find the length of the LCS, even if the order of the words in the answer and the etalon does not match, which is an advantage of the proposed approach. In the third step, the total indicator of similarity of response and etalon is calculated by finding the sum of LCS of individual words. The last step is to formulate the result of validation of the answer based on the value of the similarity indicator (set depending on requirements: high level of coincidence — 50 %, sufficient — 30 %, low — 10 %). The proposed information technology has been tested for the processing of the answers in an arbitrary text form. A comparison of the obtained results with Levenshtein distance and latent-semantic analysis was performed. The proposed IT gives the best result when checking the answers in which synonymous words are used.*

*The number of false results for the 50 variants of responses of different lengths (from 10 to 200 characters) was 4 %. As a result of the test, the recommended length of the answer and the etalon (maximum 200 characters) was set. Such length provides the higher accuracy. The directions of future research are formed: increasing the efficiency of the algorithm by introducing an additional stage of processing, — determining the "general" degree of similarity of the answer and the etalon based on the Jacquard coefficient; implementation of an automated intelligent knowledge control system based on client-server technology; forming the result of the test to the relative scale of assessment of the level of knowledge.*

**Keywords:** intellectual system, test control of knowledge, answers in arbitrary text form, "fuzzy" line comparison, length of the longest common subsequence, Hirschberg algorithm.

**Kuzma Kateryna T.** — Cand. Sc. (Eng.), Associate Professor of the Chair of Information Technologies, e-mail: [katushke2017@gmail.com](mailto:katushke2017@gmail.com)



## Информационная технология проверки ответов в интеллектуальной автоматизированной системе контроля знаний

<sup>1</sup>Николаевский национальный университет имени В. А. Сухомлинского

*Исследован процесс проверки ответов в системах тестирования на вопросы «открытого» типа (короткий текстовый ответ, эссе). Проведен анализ автоматизированных систем, в которых реализован этот процесс. Определены ограничения их применения для контроля уровня знаний по техническим дисциплинам. Аргументирована актуальность исследований по применению методов «нечеткого» сравнения строк в задачах проверки ответов, представленных в произвольной текстовой форме. Предложена информационная технология проверки таких ответов интеллектуальной автоматизированной системы контроля уровня знаний в виде структурной поэтапной схемы процесса обработки входящего ответа. Приведены вычислительные алгоритмы каждого этапа обработки, выполнена их реализация на языке C#. На первом этапе выполняется «нормализация» слов ответа и эталона (правильного ответа на вопрос, хранящегося в базе данных). Результатом первого этапа является строчные массивы: первый — набор слов ответа, второй — слова эталона (не учитываются слова длиной меньше четырех символов). Вторым этапом предусматривается циклический вызов функции нахождения длины наибольшей общей подпоследовательности слов (LCS) массивов ответов и эталона. Представлена блок-схема процедуры вычисления LCS, которая базируется на рекурсивном алгоритме, предложенном Хиршбергом. Сравнение каждого слова входящего ответа со всеми словами эталона, включая синонимы, позволяет найти длину LCS даже при несовпадении порядка слов в ответе и эталоне, что является преимуществом предложенного подхода. На третьем этапе происходит вычисление общего показателя сходства ответа и эталона путем нахождения суммы LCS отдельных слов. Последний этап — формирование результата проверки ответа на основе значения показателя сходства (устанавливается в зависимости от требований: высокий уровень совпадения — 50 %, достаточный — 30 %, низкий — 10 %). Проведены испытания предложенной информационной технологии обработки ответов, представленных в произвольной текстовой форме. Выполнено сравнение полученных результатов с расстоянием Левенштейна и латентно-семантическим анализом. Предложенная ИТ дает лучший результат при проверке ответов, в которых употребляются слова-синонимы.*

*Количество неверных срабатываний для 50 вариантов ответов различной длины (от 10 до 200 символов) составило 4 %. В результате тестирования установлена рекомендованная длина ответа и эталона (не более 200 символов), при которой достигается наибольшая точность. Сформированы направления будущих исследований: повышение эффективности алгоритма путем введения дополнительного этапа обработки — определение «общей» меры сходства ответа и эталона на основе коэффициента Жаккара; реализация автоматизированной интеллектуальной системы контроля уровня знаний на базе технологии клиент-сервер; приведение результата проверки в относительную шкалу оценки уровня знаний.*

**Ключевые слова:** интеллектуальная система, тестовый контроль знаний, ответы в произвольной текстовой форме, «нечеткое» сравнения строк, длина наибольшей общей подпоследовательности, алгоритм Хиршберга.

**Кузьма Екатерина Теодозиевна** — канд. техн. наук, доцент кафедры информационных технологий, e-mail: katushke2017@gmail.com