

УДК 378:001.891

**ІНТЕЛЕКТУАЛЬНА АВТОМАТИЗОВАНА СИСТЕМА КОНТРОЛЮ ЗНАНЬ
СТУДЕНТІВ ВИЩИХ НАВЧАЛЬНИХ ЗАКЛАДІВ****Катеринчук І.С.¹, Кравчук В.В.², Кулик В.М.², Рачок Р.В.³**¹Національна академія ДПСУ²Хмельницький ЦНТЕІ³Національна академія ДПСУ

Розроблено концепцію побудови інтелектуальної автоматизованої системи оцінювання знань. Така система передбачає наявність наступних підсистем: функціональної підсистеми перетворення інформації із зовнішнього представлення у внутрісистемні дані; підсистеми перетворення із внутрісистемного представлення даних у зовнішню інформацію; підсистеми представлення знань, яка включає блок формування результатів та систему штучного інтелекту.

Ключові слова: *система оцінювання знань, поточний контроль, підсумковий контроль, модульний контроль, тестування, штучний інтелект, автоматизована система оцінювання, математична модель, лінгвістична модель.*

Постановка проблеми. Останнім часом для автоматизації проведення як поточного, так і підсумкового контролю рівня засвоєння знань, використовуються системи тестового контролю. Однак, за своєю функціональністю існуючі системи суттєво обмежують можливості неформалізованої побудови тестових завдань. Зручність програмної реалізації привела до того, що сучасні програмні засоби, які використовуються для тестування, дозволяють будувати питання лише певних типів. Переважно це питання виду «одне питання – декілька варіантів відповідей серед яких одна вірна», «одне питання – декілька варіантів відповідей серед яких декілька вірних», «співставлення варіантів відповідей». Всіх їх об'єднує те, що в них надаються варіанти відповідей, серед яких є вірна або вірні відповіді. В багатьох випадках той хто тестується, не має міцних знань, але має добре розвинену інтуїцію, він може «вгадати» багато вірних відповідей. З іншого боку, наявність декількох близьких варіантів відповідей може розгубити того хто тестується, і він може помилитись. Тобто психологічні особливості особистості, можуть в окремих випадках суттєво вплинути на результат тестування.

Все це зумовлено тим, що жорстка, шаблонна організація сучасних систем тестування, звичайно, визначається не потребою в об'єктивному оцінюванні знань, а зручністю програмної реалізації.

Звичайно, менш формалізованою є надання текстової відповіді на питання в довільній формі. Деякі тестові системи дозволяють формувати такі питання, однак для перевірки відповідей на них залучається експерт – людина. Необхідність залучення людини є недоліком такого комп'ютерного тестування, що суттєво погіршує його ефективність.

Для автоматизації перевірки відповіді, наданої в текстовому форматі природною людською мовою, необхідно розробити методику порівняння такої відповіді зі зразком (зразками) правильної відповіді закладеної при розробці тестових завдань. Звичайно, традиційне порівняння строк для вирішення цієї задачі застосувати неможливо. На нашу думку, первинна причина в необхідності нечіткого порівняння строк, полягає у потребі врахувати (компенсувати) нечіткість, яка властива людському мозку і природній мові. Ця нечіткість, зокрема, проявляється в тому, що одне і те ж висловлювання може бути представлене людиною у текстовому вигляді по різному. І всі ці представлення, з точки зору їх інформаційного наповнення, будуть тотожні.

В Україні, починаючи з 1998 р. йде робота з розробки та впровадження галузевих стандартів вищої освіти. Система державних стандартів вищої освіти України, з одного боку, чітко регламентує вимоги до якості підготовки фахівця, а з іншого, визначає як нормативний обсяг змісту, так і його варіативну частину – вибір навчального закладу і студента.

Галузеві стандарти вищої освіти містять три компоненти: освітньо-кваліфікаційну характеристику – перелік вимог до фахівця у вигляді вмінь, що мають демонструвати випускники навчального закладу, освітньо-професійну програму – зміст навчання, що забезпечує набуття професійних знань та вмінь, та засоби діагностики якості вищої освіти – регламентація вимог до стандартизованого контролю рівня професійної компетентності фахівця. Відповідно до вимог ГСВО стандартизованим засобом діагностики є стандартизований тест. Тому в останні роки з метою оцінювання знань у навчальних закладах широко використовуються системи тестування, у тому числі з використанням спеціального програмного забезпечення.

Аналіз останніх досліджень. Аналіз наукових робіт з питань контролю знань показав, що традиційні методи страждають інформаційною однозначністю, відсутністю об'єктивних вимірювальних показників, якісною однобічністю і суб'єктивним впливом на результат контролю. До того ж проведення контролю знань у традиційній формі вимагає забагато аудиторного часу, якого на сьогодні не мають викладачі вузів. Тому виникає необхідність в нових формах контролю та модифікації вже відомих.

Розробкам інтелектуальних систем оцінювання присвячені дослідження Г.М. Шидло, який пропонує використовувати для їх створення апарат нечіткої логіки. І.Д. Рудинський в своїй статті для побудови таких систем визначає як пріоритетні методи теорії штучного інтелекту. Однак в цих працях лише визначаються в загальних рисах можливі напрями їх побудови і не вказуються, які методи теорії штучного інтелекту та теорії нечітких множин можливо використати для розв'язання цієї проблеми. Окрім цього, кожний автор зосереджується на вузьких, окремих питаннях цього наукового завдання. Відсутнє комплексне бачення проблеми побудови інтелектуальної системи автоматизованого контролю знань. Усе викладене вище спонукало авторів до розроблення інтелектуальної системи оцінювання, яка б частково могла усунути перелічені вище недоліки.

Як вказують автори робіт [1,4,5,9] розробка інтелектуальної автоматизованої системи оцінювання знань знаходиться в стадії формування і не одержала до останнього часу прийняттого для практики закінченого рішення.

У зв'язку з цим виникла необхідність постановки і вирішення актуального **наукового завдання**, що має важливе для адекватного оцінювання навчальної діяльності студентів значення – дослідження і розроблення математичного та програмного забезпечення прикладних процедур комплексного оцінювання знань студентів вищих навчальних закладів на основі застосування системи штучного інтелекту.

Метою даної роботи є розробка алгоритмів нечіткого порівняння текстової інформації для застосування в системах оцінювання знань.

Робота виконується в рамках Державної програми «Інформаційні та комунікаційні технології в освіті і науці» на 2006-2010 роки. На відміну від існуючих систем оцінювання з використанням тестів, розроблювана система надаватиме можливість оцінювати письмові відповіді студентів. Письмова відповідь студента представлятиме нечітку лінгвістичну змінну, яка за розробленим алгоритмом здійснюватиме перевірку граматики, та орфографії, на основі чого, буде сформовано образ відповіді та його порівняння з еталонними лінгвістичними змінними бази знань. База знань представлятиме собою розгалужену структуру інформації предметної області.

Слід відмітити, що для повного вирішення даного завдання необхідно забезпечити порівняння текстової відповіді зі зразком за змістом.

Однак такий підхід потребує створення окремих експертних систем для порівняння фраз різної тематики, що є складною задачею. Більш простою альтернативою є використання спрощених методів нечіткого порівняння строк.

Один з можливих підходів, який може бути використаний для нечіткого порівняння строк передбачає визначення метрики і обчислення відстані між строками [1, 2]. Чим більша відстань, тим більшою є відмінність. Оскільки в комп'ютері текстова інформація кодується числами, кожний текстовий рядок представляє собою вектор в N-мірному просторі, де N – кількість символів в рядку.

Функція $d(x,y)$ для обчислення відстані (метрики) між двома векторами x та y має наступні властивості:

1. невід'ємність: $d(x,y) \geq 0 \quad \forall x,y$;
2. властивість нуля: $d(x,y) = 0 \Leftrightarrow x = y$;
3. симетричність: $d(x,y) = d(y,x) \quad \forall x,y$;
4. нерівність трикутника: $d(x,z) \leq d(x,y) + d(y,z) \quad \forall x,y,z$.

Можливо побудувати багато різних метрик, які б відповідали цим властивостям. Наприклад, може бути використана Евклідова метрика:

$$d(x, y) = \sqrt{\sum_i (x_i - y_i)^2} \quad (1)$$

Однак при обробці текстової інформації така метрика не завжди є зручною. Звичайно, кількість символів, які будуть введені у відповідь на тестове питання не є константою. Тому необхідно мати можливість порівнювати рядки різної довжини і, відповідно, розмірності просторів, в яких вони знаходяться. Тому для нечіткого порівняння текстової інформації доцільно використання метрик, які оцінюють максимальну «вартість» перетворення одного текстового рядка в інший [1, 3].

Однією з найпростіших є відстань (метрика) між рядками за Хеммінгом, яка визначається як число позицій в яких символи не співпадають. Більш складною є метрика Левенштейна, з використанням якої можливим є порівняння рядків різної довжини.

Однак, проведені дослідження застосування метрики Левенштейна показали, що безпосереднє її застосування в задачі нечіткого порівняння відповіді зі зразком в ході тестування є малоефективним. При незначних відхиленнях відповіді від зразка відстань за Левенштейном є невеликою. Проте, якщо в реченні присутній зайвий пропуск, зсув, перестановка слів або інші спотворення, несуттєві з точки зору змісту, інколи отримується значна відстань. У той же час на коротких текстових рядках (одне слово), даний підхід до порівняння показав задовільні результати. Все це свідчить про те, що безпосереднє застосування метрики Левенштейна для перевірки і оцінки відповідей у тестових системах не є ефективним.

Тому для нечіткого порівняння текстової інформації у відповідях в ході тестування, було розроблено алгоритм у якому і зразок і відповідь розбиваються на окремі слова. Після чого проводиться нечіткий пошук співпадінь по словам між зразком і відповіддю, для чого застосовується алгоритм Левенштейна. На основі інформації про відповідність по словам будується оцінка в межах від 0 до 100. Значення 100 відповідає повній збіжності по словам, 0 коли жодного слова з оригіналу немає у відповіді.

Зрозуміло, що порівняння лише за нечітким співпадінням окремих слів не дає об'єктивної оцінки. Тому окрім порівняння за співпадінням слів, проводиться перевірка відповідності структури речення відповіді і зразка. З'ясовується, наскільки порядок слів у відповіді відповідає порядку слів у зразку і оцінюється в межах від 0 до 100.

На основі двох оцінок формується інтегральна зважена оцінка, яка нараховується за відповідь. Вагові коефіцієнти були визначені з залученням експертів. Дослідження розробленого алгоритму показали, що він дозволяє оцінити відповідь на питання в тесті надану у текстовому вигляді. При незначних відхиленнях у відповіді від зразка виставляється оцінка близька до максимальної. Навіть при незначних спотвореннях речення, коли його зміст не втрачається, оцінка відповіді є високою. З іншої сторони, коли надана відповідь не відповідає зразку виставляється низька оцінка, яка по 100 бальній шкалі прямує до 0.

Слід відмітити, що в окремих випадках, коли питання не пов'язане з використанням в відповіді чіткої термінології, надана текстова відповідь в довільній формі за змістом може відповідати зразку, однак використання алгоритму нечіткого порівняння приводить до виставлення низької оцінки.

Для покращення алгоритму потрібно врахувати всі можливі синоніми слів і різні підходи до конструювання речень (можливо, навіть і помилки побудови речень).

Можливим перспективним напрямом подальшого вдосконалення систем тестового контролю, може бути використання методів штучного інтелекту. При використанні природної мови, однакове за змістом висловлювання може бути описане різними способами. Звичайно структура і елементи (окремі слова) текстових представлень може суттєво відрізнятися від зразка. У цьому випадку, використання описаного вище підходу буде некоректним. Тому для порівняння за змістом текстової відповіді зі зразком потрібно виділити цей «зміст» (знання) як з відповіді, так і зі зразка і провести порівняння. Вирішення даної задачі можливе з використанням методів штучного інтелекту.

На жаль, аналіз сучасного програмного забезпечення в цій галузі показує, що програм, які на основі використання методів штучного інтелекту були б здатні повною мірою нетривіально опрацювати (порівнювати) отримані з тексту елементи «знань», сьогодні не існує навіть для англійської мови [4]. Ця ситуація обумовлена двома причинами [4]. Перша, полягає в незначному розповсюдженні систем лінгвістичного аналізу тексту, здатних інтерпретувати відношення між словами і, відповідно, дійсно видобувати знання як певні елементи з внутрішньою структурою і придатні для змістовної обробки «штучним розумом». Подібні системи лише розпочали з'являтися (Net Owl (www.netowl.com), Attensity (www.attensity.com), RCO Fact Extractor (www.rco.ru)) і їх ще не встигли інтегрувати в прикладних застосуваннях. Друга причина полягає в низькій достовірності автоматично отриманих з тексту «знань». Це пов'язано з недосконалістю сучасних алгоритмів інтерпретації тексту і, в окремих випадках, з низькою якістю джерел інформації.

Для «виділення» змісту з текстової інформації, може бути використаний семантичний аналіз. Він дає можливість з довільного тексту на природній мові виділити змістовну структуру (знання). При цьому відбувається виявлення змісту речень, або окремих їх частин. Звичайно, семантичний аналіз є складним завданням. Він важко піддається формалізації. Звичайно, кожна мова має свої особливості, і їх необхідно враховувати при проведенні семантичного аналізу.

Для визначення семантичних відношень між окремими словами, звичайно, використовується тезаурус мови. Природно, що даний тезаурус є специфічним для кожної мови. Він задає набір бінарних відношень на множині слів природної мови. Проблема створення якісного тезаурусу є однією з основних при використанні алгоритмічного підходу. Хоча зараз існують комерційні продукти, в яких використовуються тезауруси для різних мов, вони насправді включають лише їх підмножини (англійської, російської).

У результаті семантичного аналізу будується семантична мережа – структура для представлення знань у вигляді вузлів пов'язаних дугами (зв'язками). Властивості отриманої семантичної мережі:

- вузли мережі представляють собою поняття, предмети, події, стани;
- дуги семантичних зв'язків створюють відношення між вузлами – поняттями (відношення можуть бути різних типів);
- деякі відношення між вузлами є лінгвістичними, просторовими, часовими, логічними та ін.;
- поняття організовані за рівнями відповідно до ступеня узагальненості.

Таким чином, в результаті аналізу тексту з нього автоматично видобувається інформація («знання») у вигляді мережі основних понять і зв'язків з ваговими коефіцієнтами. В якості змістовного «портрету» тексту при подальшому порівнянні розглядається не просто список ключових слів, а мережа понять, яка в певному сенсі є «відбитком» змісту тексту.

Кожне поняття має певну вагу, яка відображає значимість цього поняття в тексті. Зв'язки між поняттями також мають вагу.

Порівняння семантичних мереж двох текстів дозволить провести їх порівняння за змістом. Незалежно від побудови речень, наявності додаткових суджень, несуттєвих якісних характеристик, які можуть бути присутні у відповіді з неї виділяється основний «зміст» у вигляді семантичної мережі. Аналогічна процедура проводиться зі «зразком» вірної відповіді. Порівняння двох семантичних мереж (тексту відповіді і зразка) дозволить достовірно оцінити ступінь їх тотожності і в результаті виставити об'єктивну оцінку.

Для розробки математичного та програмно-алгоритмічного забезпечення автоматизованої системи контролю знань студентів пропонується використання методів як з теорії штучного інтелекту, так і з теорії нечітких множин. На відміну від систем оцінювання з використанням тестів розроблювальна автоматизована система надасть можливість оцінювати письмові відповіді студентів (існуючі системи дозволяють лише зберігати письмову відповідь для подальшої її оцінки викладачем). Письмова відповідь студента представлятиме нечітку лінгвістичну змінну, яка за розробленим алгоритмом здійснюватиме перевірку граматики та орфографії, на основі чого, буде сформовано образ відповіді та його порівняння з еталонними лінгвістичними змінними бази знань. База знань представлятиме собою розгалужену структуру інформації предметної області.

Крім традиційних тестів дана система дозволить оцінювати:

- завдання, що припускають коротку вільну відповідь у вигляді числа або одного слова;
- завдання, що припускають точну відповідь у вигляді правила, визначення, теореми тощо;
- завдання, що припускають логічну відповідь (вибір із множини, впорядкування за ознакою, доведення тощо);
- завдання, відповідь на які повинні бути подані у вигляді довільної розгорнутої відповіді – математичних викладень або тексту, що дозволяє забезпечити належну підготовку фахівців, які б правильно, семантично точно і стилістично доречно «вибирали слово».

Оцінка знань студентів проводитиметься за усіма, прийнятими у вищій школі, системами оцінювання.

Система включатиме модулі: базу даних (предмети, модулі, теми, навчальні групи); базу знань (предмети, модулі, теми); систему навчання; систему оцінювання. Етапи розробки ІАСКЗ показана на рис. 1.

Система штучного інтелекту трактуватиметься як технічна система, аналогічна загальновідомим системам автоматичного регулювання. Така система використовує в своїх методах моделі, апарат і прийоми, запозичені з різних дисциплін: психології, лінгвістики, інформатики, дискретної математики, системного програмування, науки обчислень і ін.



Рис. 1. Етапи розроблення математичного та алгоритмічного забезпечення інтелектуальної автоматизованої системи контролю знань

Робота інтелектуальної лінгвістичної підсистеми показана на рис. 2-4.

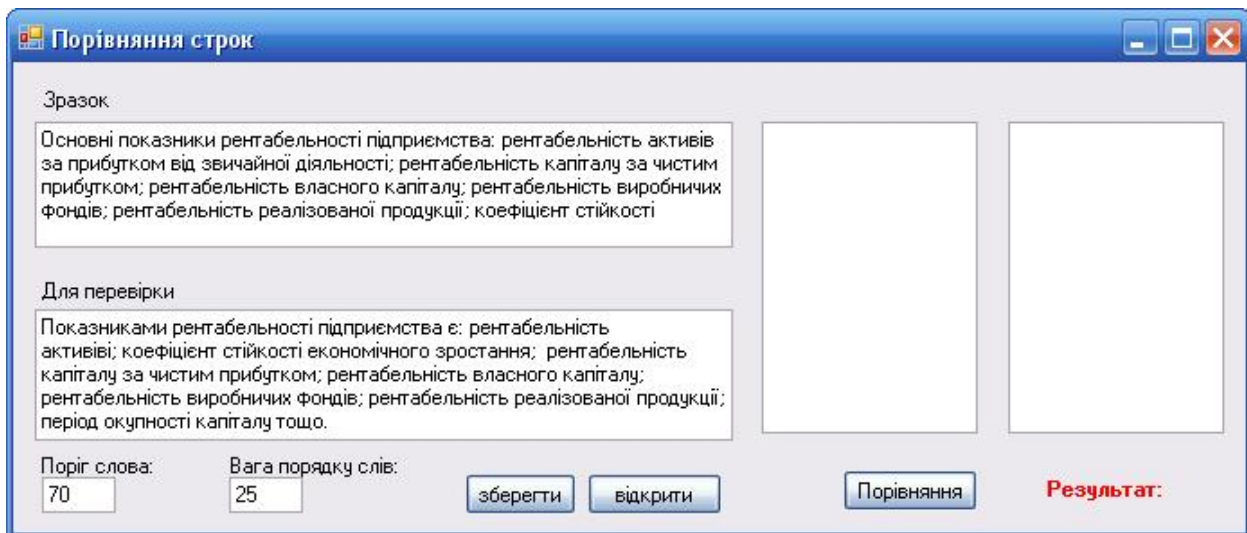


Рис. 2. Робота інтелектуальної лінгвістичної підсистеми

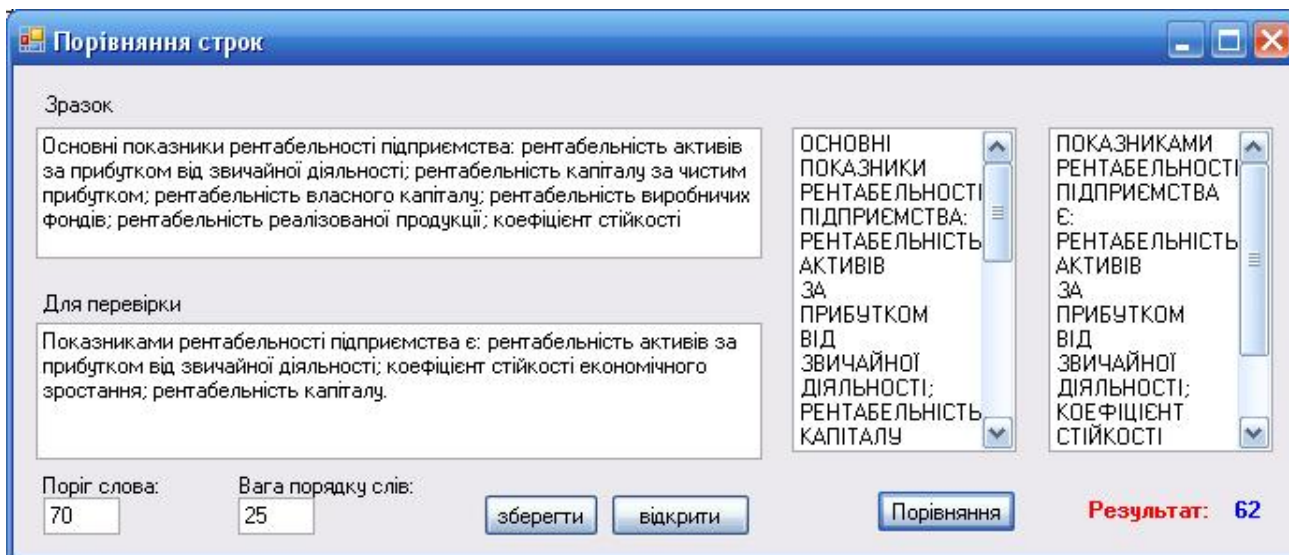


Рис. 3. Робота інтелектуальної лінгвістичної підсистеми

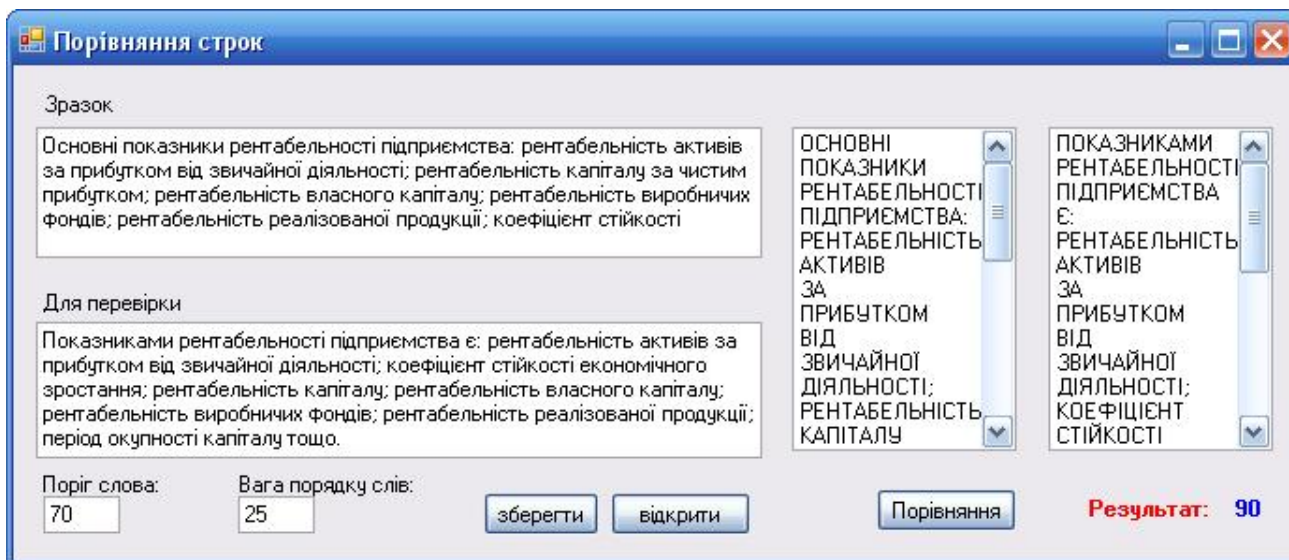


Рис. 4. Робота інтелектуальної лінгвістичної підсистеми

Структура інтелектуальної автоматизованої системи оцінювання знань студентів ВНЗ представлена на рис. 5.

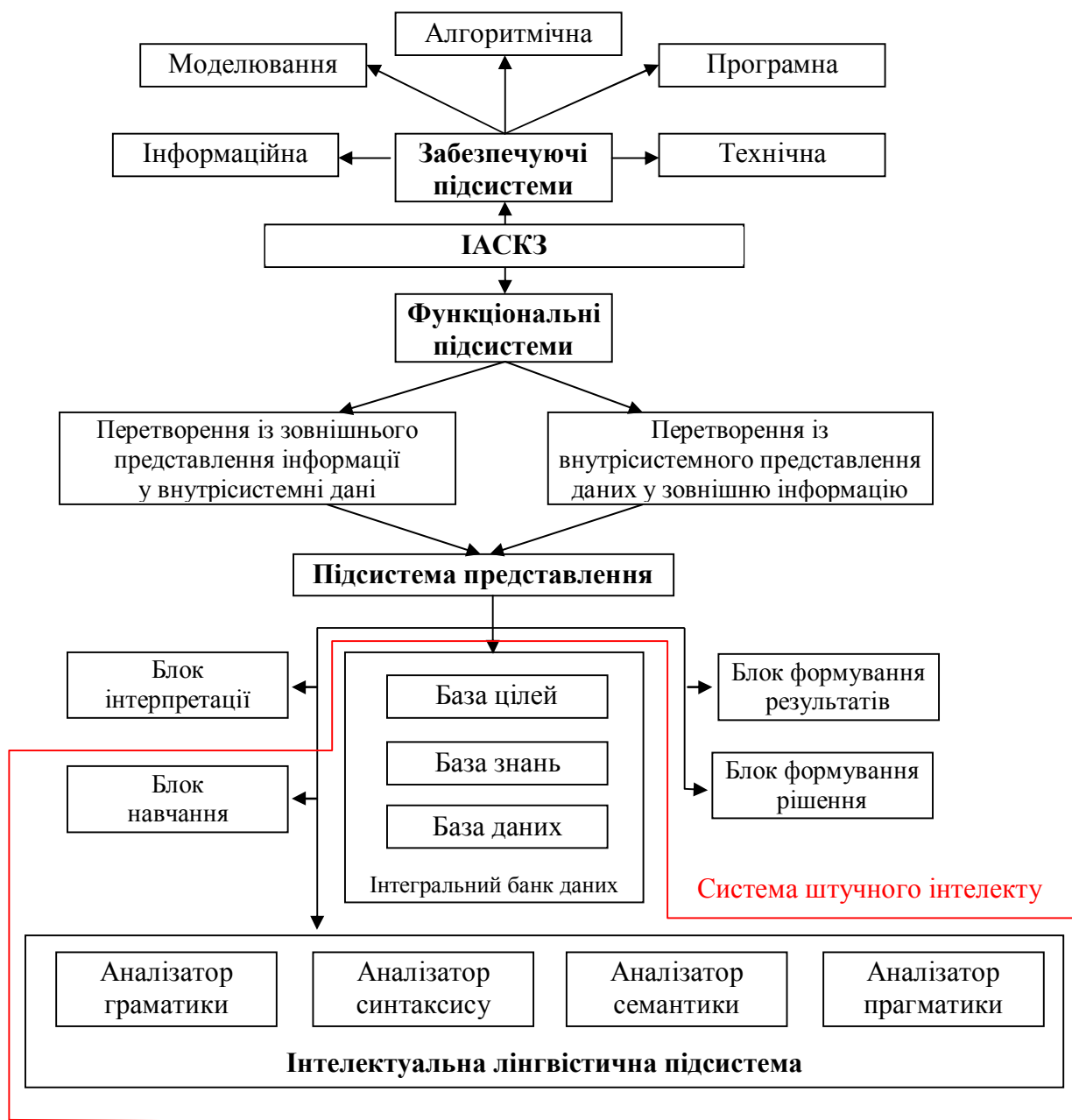


Рис. 5. Структура інтелектуальної автоматизованої системи оцінювання знань

Висновок. Таким чином, в основу інтелектуальної автоматизованої системи оцінювання знань покладено алгоритм перевірки граматики, семантики та прагматики. Для нечіткого порівняння окремих слів у відповіді в алгоритмі використовується метрика Левенштейна, однак для покращення ефективності перевірки додатково проводиться аналіз структури речення. Експериментальна перевірка показала достатню ефективність запропонованого алгоритму при перевірці питань тесту, в яких відповідь потрібно навести у вільному викладенні. Однак для подальшого розвитку пропонується використання семантичного аналізу з подальшим порівнянням семантичних мереж зразка і текстової відповіді.

В роботі алгоритмів задіяні моделі штучного інтелекту. Такі алгоритми, в якому використані елементи «штучного інтелекту», в частині побудови семантичних мереж дозволить суттєво підвищити ефективність роботи системи тестування.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Graham A. Stephen, "String Searching Algorithms," World Scientific Publishing Company, 1994.
2. Hirata M., Yamada H., Nagai H., Takahashi K. (1988) "A versatile data string-search VLSI," IEEE Journal of Solid-State Circuits, Vol. 23, No. 2, p. 329-35, April 1988.
3. А. Д. Жидких, В. И. Костин. Исследование алгоритмов текстового поиска. Международная студенческая конференция "Информатика и компьютерные технологии", ДонНТУ, 2005.
4. Ермаков А. Е. Извлечение знаний из текста и их обработка: состояние и перспективы // "Информационные технологии" – М: Вид. «Новые технологии», 2009. – С. 50 – 55.
5. Заде Л. Понятие лингвистической переменной и его применение к решению проблемных решений / Л. Заде. – М.: Мир, 1976. – 386 с.
6. Кузин Л. Т. Основы кибернетики: В 2-х т. Т. 2. Основы кибернетических моделей / Л.Т. Кузин. – М.: Энергия, 1979. – 584 с.
7. Ермаков А. Е. Извлечение знаний из текста и их обработка: состояние и перспективы // "Информационные технологии" – М: Вид. «Новые технологии», 2009. – С. 50 – 55.
8. Воронина И.Е. Компьютерное моделирование лингвистических объектов : монография / И. Е. Воронина. – Воронеж: Издательско-полиграфический центр ВГУ, 2007 – 177 с.
9. Гинзбург Е.Л. Идеоглоссы: проблемы выявления и изучения контекста / Е.Л. Гинзбург // Семантика языковых единиц: Доклады VI международной конференции. Т. I. М. , 1998. – С. 26-28.
10. Рудинский, И. Д. Создание интегрированной автоматизированной системы контроля знаний – // Информатика и образование. – 2005. – N2. – С. 117-122.
11. Чабан К.О. Прикладна програмна система оцінювання знань для комп'ютерного тестування: дис. канд. наук: 01.05.03 / К.О. Чабан. – 2008.
12. Нечеткие множества в моделях управления и искусственного интеллекта / А.Н. Аверкин, И.З. Батыршин, А.Ф. Блишун, В.Б. Силов, В.Б. Тарасов. – М.: Наука. Гл. ред. физ-мат. лит., 1986. – 312 с.
13. Рыжов. Элементы теории нечетких множеств и измерения нечеткости / Рыжов. – М.: Диалог – МГУ, 1998.–116 с.
14. <http://ru.wikipedia.org/wiki/Предикат>.
15. http://ru.wikipedia.org/wiki/Логика_первого_порядка.
16. Принципы и технологии создания интегрированной автоматизированной системы контроля знаний / [Э.М.Аскеров, М.А.Емелин, И.Д.Рудинский, Н.А.Строилов]. – КГТУ. 2008.- 54 с.