

Імітаційна модель тестового контролю знань і умінь

Тестовий контроль все більшою мірою стає невід'ємною частиною навчального процесу для всіх типів і рівнів навчальних закладів України. Здобувши широке поширення в західноєвропейських країнах і США, він поступово завойовує все нові позиції у вітчизняній вищій школі. Не є винятком у цьому відношенні і технічні університети, де тестування починає застосовуватися з самого моменту початку навчання майбутніх студентів і може регулярно використовуватися аж до вручення їм дипломів. Починаючи з 2008 року, випускники середньої школи, які обрали для отримання освіти технічні спеціальності, складають, як і абсолютна більшість інших абітурієнтів, вступні іспити, завдання яких розроблені у формі тестів. Студенти, переходячи з курсу на курс, стикаються з тестуванням впродовж усього терміну навчання і, з огляду на комп'ютеризацію всіх сторін діяльності вузу, цей процес буде не тільки продовжуватися, але і розширюватиметься. Згідно з розпорядженням Міністерства освіти і науки України № 28-р від 05.03.01 введено в дію комплекс нормативних документів з розроблення засобів діагностики якості вищої освіти [1], де одним з основних інструментів об'єктивного контролю кінцевої мети освітньо-професійної підготовки вказано державний іспит, що проводиться в тестовій формі.

Не применшуючи переваг тестового контролю, які очевидні і неодноразово підкреслювалися авторами [2, 3], одночасно слід зазначити, що в зарубіжних теоріях, які, як правило, лежать в основі сучасних методів тестування, багато в чому не враховуються традиції і переваги вітчизняної системи освіти. І якщо формальне значення слів «іспит» (від латинського «examen») і «тест» (від англійського «test») співпадають і означають одне і те саме - випробування, то смисловий зміст, що вкладається в них, у нас істотно відрізняється. Починаючи з Я.А.Коменського і Г.С.Сковороди і до теперішнього часу в українській педагогіці особливе значення на всіх етапах навчання надавалося особі вчителя і неодноразово підкреслювалась його роль та можливості у виявленні дійсних знань під час проведення іспиту.

Обговорюючи значущість участі викладача в іспиті, відзначимо, що в першу чергу мова не повинна йти про навчальну функцію, яка, на нашу думку, тут менш істотна, ніж при інших видах навчальної діяльності. Контакт викладача із студентом на усному іспиті протягом 10-20 хвилин, безумовно, важливий, але непорівнюваний із загальною тривалістю спілкування студента і викладача протягом семестру. Більш того, норми, за якими встановлюються терміни іспиту і частка часу, що виділяється викладачеві на опитування одного студента постійно скорочуються, що часто зводить нанівець навчальний і виховний вплив викладача на студента під час іспиту.

У той же час важко переоцінити діагностичні можливості викладача при проведенні усного іспиту. Ніхто, окрім викладача не зможе за зовні невиразною відповіддю, шляхом бесіди, додаткових, уточнюючих, навідних питань виявити, що ж насправді знає студент і чи не обумовлена така відповідь просто його зайвим хвилюванням. І, навпаки, досвідчений викладач часто здатний вже на початковому етапі контролю встановити рівень знань студента і перейти до наступної теми або перервати іспит, проставивши оцінку за наслідками отриманих відповідей, тоді як при тестуванні, незалежно від очевидності отриманих результатів, буде проходити іспит в повному обсязі.

Значно більше можливостей у викладача і при формулюванні тексту питання не тільки з урахуванням отриманих від студента відповідей, але і залежно від змісту контрольованого навчального матеріалу. Для дисциплін, вивчення яких вимагає нестандартного мислення і неформалізованого підходу, важко скласти адекватні і прості у поданні тестові завдання, і тоді вони будуть у край складні для розуміння студентами. Добре підготований студент може бути погано зосереджений і не розуміти зміст надмірно складних формулювань таких завдань, або завдань, спеціально заплутаних. Не виключено, що невдала відповідь студента тоді буде обумовлена виключно неточністю самих формулювань. В той же час навіть погано підготовленому студенту, крім можливості простого вгадування, часто досить скористатися простими логічними висновками і здоровим глуздом, щоб добре виконати тестове завдання.

При зіставленні усної і тестової форм проведення іспиту часто завищується об'єктивність тестового контролю, що реалізовується за сценарієм більшості загальноприйнятих методик тестування. Підкреслюючи усунення людини, носія суб'єктивності, з процесу безпосереднього проведення іспиту, в той же час поза увагою залишається факт, що розробником тестів, як і раніше, залишається викладач. Тому, відзначаючи можливість реалізації єдиних вимог з конкретної дисципліни для усіх студентів, слід пам'ятати, що вони, можливо, будуть загальними тільки для даної групи тих, хто екзамнується, а в іншій групі або при проведенні тестування іншим викладачем, простою зміною критеріїв виставлення оцінки, можна до невпізнання спотворити всю картину іспиту. Звичайною практикою тестового контролю у вітчизняних вузах є підлаштовування критеріїв під набрані бали для забезпечення заданого рівня складання іспиту. І якщо згідно класичної теорії тестування щоб успішно скласти іспит, студент зобов'язаний набрати не менше 59% від максимально можливої кількості балів (при деяких видах тестів [4]), то у нас для студентів заочної форми навчання критерієм позитивної оцінки за тих же умов може бути 49%, 39% і навіть нижче.

Використання технічних пристроїв при тестовому контролі не обмежує, а навіть збільшує можливість несанкціонованих втручань у хід іспиту, тим самим знижуючи вірогідність його результатів. Недосконалий захист комп'ютерних програм може призвести до проникнення у базу тестових завдань і доступу до еталонних відповідей, що дозволить студентів при мінімальних знаннях успішно скласти іспит, а злам бази результатів іспиту взагалі позбавляє сенсу його проведення. У пошуках потрібних матеріалів студенти можуть скористатися мережею Інтернету або отримати «допомогу» через локальну мережу університету. Списування може стати істотною проблемою через близько розташовані комп'ютери, якщо не виключити синхронне зображення питань на екранах моніторів. Ще однією проблемою є складність ідентифікації студента, що знаходиться перед екраном монітора і проходить тестування. Відзначимо, що особливо важливим є збереження матеріалів на комп'ютері для вітчизняних вузів через їх обмежені можливості у розробленні і створенні захисту власних програм для комп'ютеризованого тестування, а також недостатні фінансування при придбанні якісних програмних продуктів світових виробників.

Тому, при неухильному застосуванні тестових форм контролю, що постійно розширюється, існує настійна необхідність, створити модель тестового контролю, в якій, використовуючи всі переваги тестування, при цьому максимально враховувалися б на вітчизняні традиції активної участі викладача в діагностуванні успішності навчання студентів. Така модель повинна базуватися на сучасних програмних засобах і при цьому передбачати можливість регулювання кількості поставлених питань, залежно від результативності відповідей студентів, мати розширений набір прототипів типових завдань, з яких можна створити тест, максимально відповідний змісту контрольованого матеріалу, встановлювати об'єктивні критерії перекладу набраної суми балів у показники успішності навчання та ін. Поєднання переваг комп'ютеризованого тестування з математичним обґрунтуванням процедур ухвалення рішень екзаменатором дозволить мінімізувати можливі витрати, пов'язані з повсюдним переходом технічних університетів на тестову форму контролю.

Визначаючи об'єми і зміст тестового контролю, в першу чергу слід виходити з цілей, яких необхідно досягти, і завдань які потрібно виконати. Так тестовий тематичний контроль, метою якого є визначення рівня знань і умінь з конкретної теми і на цій основі коригування процесу навчання, за своїм змістом істотно відрізнятиметься, наприклад, від підсумкового контролю, основне завдання якого виявити результативність навчання, оцінити знання і вміння, сформовані в процесі вивчення дисципліни в цілому. Очевидно, що тести тематичного контролю повинні зачіпати невеликий сегмент - розділ або частину, складатися з обмеженої серії тестових завдань, якими всебічно охоплюється тільки одна з областей навчання. Тут завдання повинні бути спрямовані не стільки на встановлення досягнутого рівня, скільки на виявлення недоліків як в роботі викладача, так і, що не менш важливо, самих студентів. На відміну від тематичного, при підсумковому контролі оцінюється широкий діапазон результатів навчання і в кінці тривалого навчального періоду, контролюються знання великого числа питань з предметної галузі всієї дисципліни, включаються питання підвищеного ступеня складності, в порівнянні з іншими видами тестів. Можна відзначити і ряд інших відмінностей, які, не виключено, будуть властиві різного призначення і спрямованості тестам.

Одночасно з цим методики тестового контролю, відрізняючись за деталями, містять загальний набір процедур, які повинні бути витримані при всіх видах тестування і реалізовуватися в обумовленому порядку. Набір таких процедур наведений на рисунку 1, де показана схема, за якою моделюються, згідно з розробленою авторами імітаційною моделлю, дії викладача при традиційному контролі.

Згідно з наведеною схемою (Рис 1) імітаційна модель тестового контролю включає такі процедури.

Розроблення прототипів тестів.

Розроблення тестів починається з добору матеріалу для контролю. В імітаційній моделі тестування припускається, що ця частина роботи, як і в решті всіх моделей тестування, виконується експертами з числа викладачів або методичних працівників. При доборі матеріалу експерти в першу чергу керуються цілями тестування, згідно з якими встановлюється, які види знань і умінь найбільш значущі для поставлених цілей і який при цьому потрібний рівень їх засвоєння.

Після добору змісту тесту переходять до проектування прототипів тестових завдань, які створюються шляхом поділу навчального матеріалу на окремі фрагменти, контроль яких можливий за допомогою типових тестових форм. При цьому поза типовими формами ні завдання, ні його зміст не існують. Прототип завдання є результатом поєднання змісту з найбільш відповідною для нього формою, яка і є способом збереження змісту. Очевидно, що чим більше типових форм у розпорядженні розробника, тим більше у нього можливостей створити прототип, що найбільшою мірою відповідає цілям контролю.

Контрольований навчальний матеріал має різний ступінь складності, як і завдання, призначені для виявлення рівня засвоєння такого матеріалу, теж повинні відрізнятися за складністю. Наприклад, завдання на просте відображення матеріалу на рівні впізнання і відтворення за складністю відрізнятиметься від завдань, виконання яких вимагає творчого застосування надбаних знань і дозволяє виявити уміння застосовувати знання в нестандартних ситуаціях. Тому в імітаційній моделі встановлення міри складності завдання є однією з обов'язкових складових процедур проектування тестів.

Завершується перший етап проектування встановленням кількості тестових завдань і об'єднанням їх прототипів в єдиний тест. Кількість завдань в тесті залежить, по-перше, від змісту матеріалу, знання якого перевіряється: чим значніший обсяг навчального матеріалу, що перевіряється, тим, як правило, більше потрібно завдань, щоб встановити рівень його засвоєння. По-друге, від складності завдань - із зростанням складності повинен збільшуватися час іспиту, який не може розширюватися до нескінченності і єдиним способом регулювання тут може бути зниження числа завдань. При цьому слід зауважити, що немає ніякого зв'язку між формою завдання і його складністю. Наприклад, завдання на вибір відповіді, яке, як правило, має закриту форму і тому традиційно відноситься до простих, може бути для студента набагато складнішим, якщо він погано знає контрольований матеріал, у порівнянні із звичайно складнішим завданням відкритої форми. Таким чином, кількість завдань залежить значною мірою не від форми, а від змісту контрольованого навчального матеріалу. З формою ж більше пов'язані види знань, що перевіряються. У загальному випадку, чим більше завдань, тим двирогідніше можуть виявитися результати тестування. Проте їх надмірне збільшення неприпустиме. Тому достатнім критерієм для формування тесту є можливість зіставлення результатів тестування з різними тем і дисциплінами, за умови, що кількість завдань у тесті достатня для забезпечення прийнятної точності виявлення рівня засвоєння навчального матеріалу.

Контроль знань

Для вірогідного виявлення рівня засвоєння студентами навчального матеріалу необхідно при тестовому контролі перевірити всі знання і уміння, які необхідно сформувати у студентів. Але через низку обставин обсяг, передбачуваний тестовими завданнями, завжди виявляється меншим від обсягу запропонованого до вивчення. Тому при тестуванні постає завдання: за наслідками виконання обмеженої кількості тестових завдань зробити висновок про знання студентами всього матеріалу, а у випадку, якщо цього не вдасться зробити, мати можливість сформулювати додаткові завдання в обсязі, достатньому для прийняття остаточного рішення про досягнутий рівень успішності вивчення всього навчального матеріалу. Істотною при цьому для студента є можливість висловити свої сумніви, якщо надбані ним знання не дозволяють дати однозначну відповідь на поставлене питання.

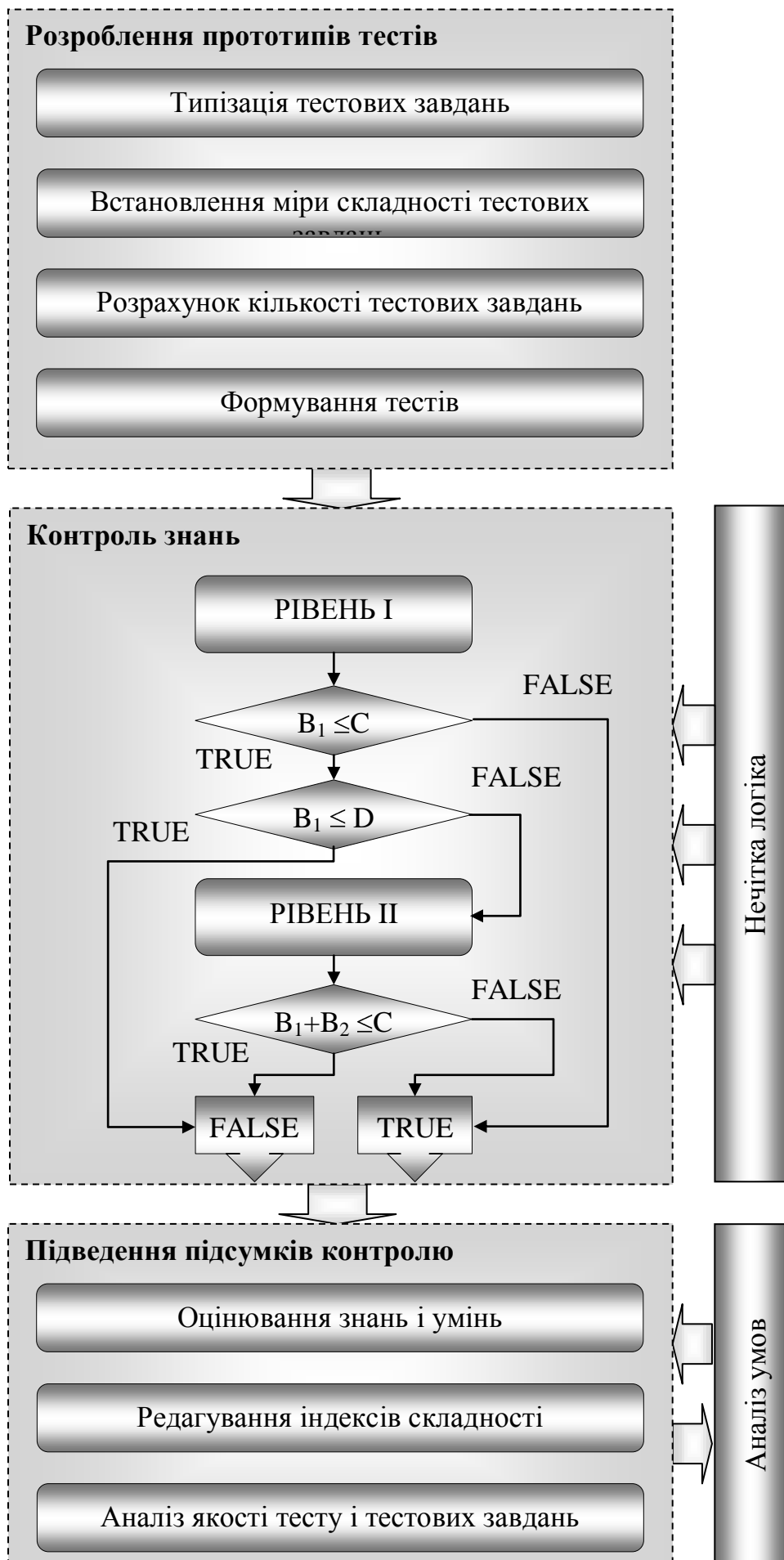


Рис. 1

Підведення підсумків контролю

Підсумки тестування визначаються на основі статистичного опрацювання результатів тестування, яке для всіх студентів повинне проводитися в однакових умовах, що і фіксується на етапі аналізу умов, при яких здійснюється тестовий контроль.

Встановлення рівня і структури знань визначається при аналізі відповідей кожного студента на всі завдання тесту. Чим більше правильних відповідей, тим вище його індивідуальний тестовий бал, який асоціюється з рівнем знань. Одночасно з цим оцінюються характеристики успішності навчання. Подібно до того, як викладач при усному контролі судить про успішність навчання, зіставляючи навчальні досягнення студента з аналогічними результатами групи студентів, так і в імітаційній моделі вона визначається як здатність вийти за однотипних умов навчання на вищий рівень засвоєння навчального матеріалу.

Завершується підведення підсумків аналізом якості тестів і тестових завдань. Його кінцевою метою є інтерпретація отриманих емпіричних результатів і доведення характеристик прототипу тестового завдання до рівня, відповідного визначеним науково обґрунтованим критеріям якості. Аналіз виконується при чергуванні процедур синтезу і верифікації, яке повторюється до того часу, поки всі прототипи тестових завдань не набудуть статусу тестових.

Наявність достатнього числа тестових завдань, отриманих в результаті обов'язкової емпіричної апробації завдань у тестовій формі, дозволяє перейти до складання тесту як системи, що має цілісність, відповідний склад і структуру.

Всі запропоновані в імітаційній моделі нововведення реалізовані в комп'ютерному програмному комплексі SSUquestionnaire, який призначено ж для комп'ютеризованого, так і безмашинного тестового контролю знань.

Програмний комплекс складається з WWW-модуля, серверної і клієнтської частин.

WWW-сервер Microsoft Internet Information Services 5.1 і вище.

Браузером при тестуванні може бути Internet Explorer версії 5.5 і вище.

Середовищем реалізації є платформа Microsoft .NET 3.5.

База даних реалізована на SQL Server 2008 (можливе використання SQL Server Express).

Коректність роботи гарантується на платформі Microsoft Windows 2000/XP/2003/Vista.

Апаратне забезпечення - будь-який комп'ютер, на якому забезпечується функціонування названих програм.

Комп'ютеризований контроль може здійснюватися з використанням мережевих технологій або генеруванням тестових завдань з подальшим контролем на локальній машині.

При безмашинному контролі формуються бланки тестових завдань для студентів і картки відповідей за варіантами завдань.

Основні вікна програми показані на рисунках 2 і 3.

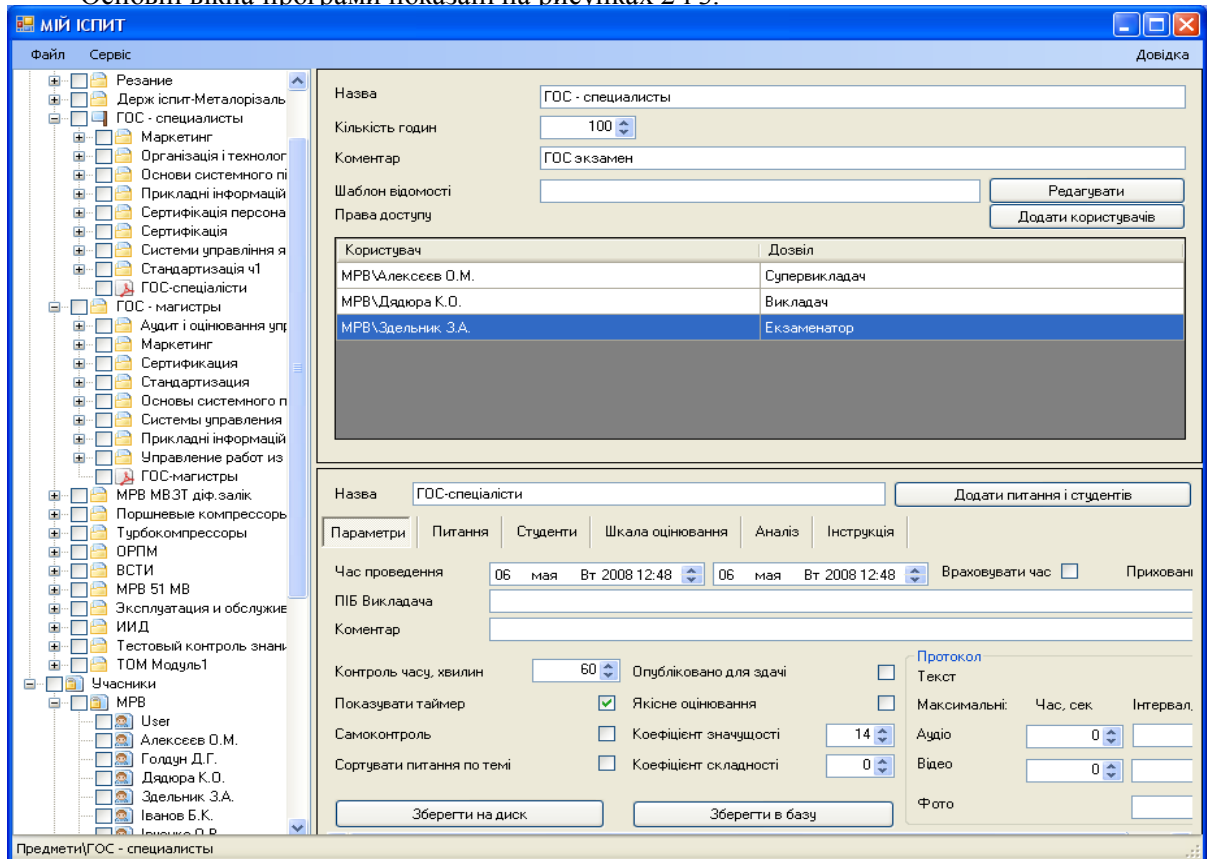


Рис. 2

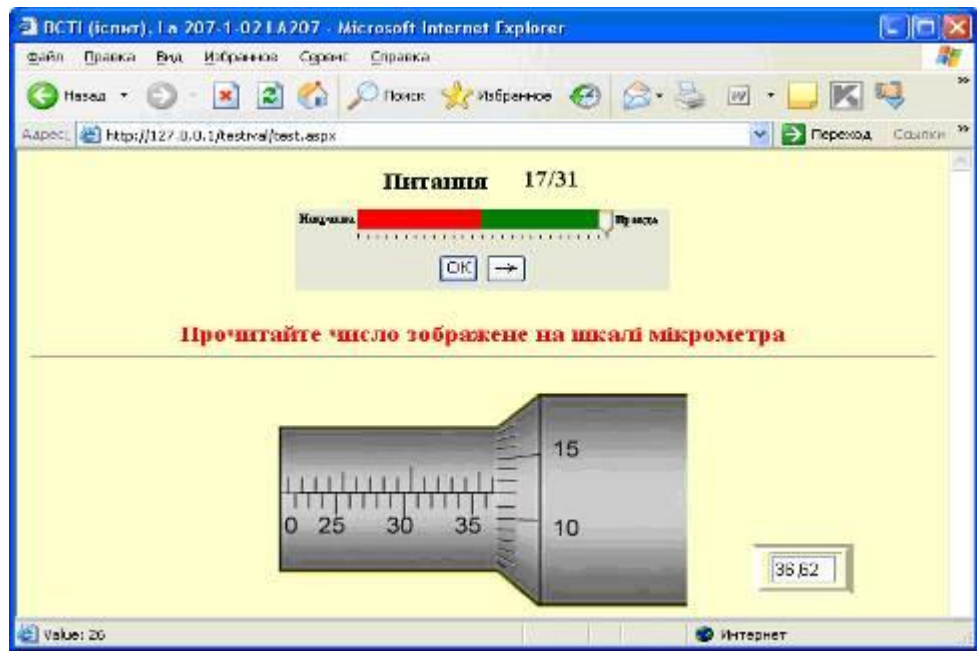


Рис. 3

Для підвищення вірогідності оцінки знань у версії 4.9 реалізовано:

- автоматизоване проектування типових питань (за типами: "Зіставлення", "Класифікація", "Позиціонування", "Введення символів", "Підтвердження", "Вибір відповіді", "Підстановка", "Впорядковування", "За ключовими словами", "Виправлення", "Послідовний вибір", "Послідовність дій");

- алгоритми нечіткої логіки, що дозволяє при відповіді оперувати не тільки класичними значеннями логічних змінних "хибно" та "істинно" (рисунок 3), але і використовувати їх проміжні значення;

- проектний (на етапі розробки тестових питань) і коригування (за наслідками тестування) розрахунки індексів складності, за якими визначається ступінь труднощі питання по відношенню до решти питань;

- розрахунок кількості питань у тестовому завданні;

- оптимізація тестових завдань з використанням генетичних алгоритмів;

- дворівневий контроль і розрахунок числових значень критеріїв, при яких слід організувати додаткові сесії контролю;

- переклад отриманої суми набраних балів за відповідями на всі питання до показників успішності навчання;

- облік результатів тестування і на цій основі розрахунок імовірнісних характеристик, що визначають необхідність внесення змін у формулювання або вилучення некоректно сформульованих питань;

- технологія тонкого клієнта для дистанційного контролю знань, що підвищує захищеність мережесесій тестування від втручання ззовні;

- ведення текстового лог-файла, аудіо, фото і відео протоколів, що дозволяє документувати процес тестування і ідентифікувати студентів.

Програма має російсько-, українсько-, і англійськомовний інтерфейси і файли допомоги з докладним описом методики і технології роботи.

Докладніші відомості про програму SSUquestionnaire наведено на сайті програми за адресою www.test.sumdu.edu.ua

На завершення слід зазначити, що використання теоретико-методичної моделі імітаційного тестового контролю надає можливість ухвалювати математично точні проектні рішення для більшості типових процедур розроблення, проведення і підведення підсумків контролю, які традиційно поклалися на викладача і через його відстороненість при комп'ютеризованому тестуванні, найбільш чутливі для внесення похибок до результатів вимірювань. Пропоноване в рамках моделі математичне обґрунтування рішень, що приймаються екзаменатором, значною мірою пом'якшує недоліки комп'ютеризованого тестування як інструменту педагогічних вимірювань. Одночасно з цим явно недостатньо включених у модель процедур перевірки вимірювальних характеристик тестування, і вони потребують подальшого теоретичного розвитку і методичного оформлення.

ЛІТЕРАТУРА

1. Комплекс нормативних документів для розробки складових системи стандартів вищої освіти. Введено Наказом Міністра освіти України від 31.07.98 р. №285. – К.: 2001. – 126 с.

2. Алексеев А.Н. Дистанционное обучение инженерным специальностям - Сумы: „Университетская книга”, 2005. – 333с.

3. Алексєєв О.М., Алексєєва Г.В. Розподілена система вимірювання якості електронних підручників // Теоретичні питання культури, освіти та виховання: Збірник наукових праць - К.: КНЛУ, 2008. - Випуск 35. – С. 14 – 17.
4. Airasian P. W. Classroom assessment, 3rd ed. - New York: McGraw-Hill, 1997. - 416 p.