

УДК 004.04 (045)

Шелевицький І.В., д.т.н., професор,
Дубан Р.М., аспірант

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ПІДТРИМКИ ТЕСТОВОГО КОНТРОЛЮ ЗІ СПЛАЙН-МОДЕЛЯМИ IRT

В статті наведена інформаційна технологія підтримки тестового контролю знань, що дозволяє отримати тест з визначеними характеристиками якості. Особливістю запропонованої технології є підтримка тесту на всіх стадіях його життєвого циклу, як інформаційного продукту. Для автоматизованого розрахунку профілів IRT запропоновано використовувати сплайн-моделі.

Информационная технология поддержки тестового контроля со сплайн моделями IRT. В статье приведена информационная технология поддержки тестового контроля знаний, которая позволяет получить тест с определенными характеристиками качества. Особенностью предложенной технологии является поддержка теста на всех стадиях его жизненного цикла, как информационного продукта. Для автоматизированного расчета профилей IRT предложено использовать сплайн-модели.

Information technology for support the test control with spline-models IRT. In the article proposed the information technology support for knowledge testing that allows test the defined characteristics of quality. The peculiarity of the proposed technology is to support the test at all stages of its life cycle, as an information product. For automatic calculation of profiles IRT proposed use spline model.

Вступ. Тестовий контроль знань широко застосовується в сучасній освіті. Він є чи не найефективнішим способом перевірки знань. Тестування має ряд переваг, основними з яких є економія часу та об'єктивність результатів. Системи тестового контролю є важливим елементом освітніх інформаційних комп'ютерних технологій. Питання розвитку таких технологій входять до проекту "Відкритий світ" пріоритетного напрямку соціально-економічного та культурного розвитку держави "Нова якість життя" [1]. Головною проблемою використання тестів в якості інструменту оцінки рівня знань є складний процес розробки та оцінки рівня якості тесту.

Аналіз досліджень і публікацій. Основні вимоги до тестів та процедура їх створення і перевірки викладені в комплексі нормативних документів МОН [2], монографіях та навчальних посібниках [3]. Аналіз результатів

проводиться на основі класичної теорії тестів або на основі Item Response Theory (IRT) [3]. В умовах достатнього розвитку комп'ютерних технологій процес складання тестів і тестових завдань здебільшого виконується з використанням комп'ютерної техніки та баз даних. Найбільш відоме та вживане вільне програмне забезпечення є Moodle, iTest, та OpenTEST2, що дозволяють створювати бази тестів та проводити тестування з використанням комп'ютерної техніки, а в OpenTEST2 і проведення статистичного аналізу результатів. Але сам підхід до створення тестів в цих та інших програмних засобах має закритий характер, автор чи група авторів розробляють тест і використовують його для вимірювання знань без попереднього рецензування та випробування. Тому, на жаль, в переважній більшості випадків, при тестуванні застосовуються тести, що не пройшли належної перевірки і мають невизначені властивості, як інструменту вимірювання знань. Ймовірною причиною такого положення є відсутність практично зорієнтованих методів, алгоритмів і програмного забезпечення для оцінювання тестів.

Постановка завдання. Забезпечити належний рівень якості тестів, що використовуються для оцінки рівня знань, можливо за умови дотримання певної технології їх розробки та супроводу. У зв'язку із широким використанням комп'ютерної техніки в освітній діяльності та її високою доступністю технологія має бути реалізована в інформаційній системі. Застосування спеціальних математичних моделей має забезпечити можливість автоматичного розрахунку необхідних параметрів (рівня складності завдань тесту, рівня диференційної здатності завдань, рівня знань учасника тестування та ін.). В якості таких моделей пропонується використовувати сплайни.

Життєвий цикл тесту та інформаційна технологія його підтримки. За результатами проведених досліджень існуючого програмного забезпечення підтримки тестового контролю знань була розроблена інформаційна технологія, що забезпечує підтримку тестового контролю на різних етапах життєвого циклу (по аналогії із життєвим циклом інформаційних систем). Схематично послідовність етапів процесу розробки тесту зображено у вигляді Process Flow Description Diagrams (PFDD) за стандартом IDEF3 на рисунку 1.

Перший етап полягає у створенні тесту за певним призначенням та напрямком знань. Тесту призначається певна назва, код за УДК, та опис цільової аудиторії, для якої його планується застосовувати. Таким чином на першому етапі закладаються початкові властивості об'єкту "Тест". Перехрестя *J1* показує, що перед початком наступних етапів попередній, перший етап, має бути завершений. Другий і третій етап починаються одночасно. На другому етапі відбувається наповнення тесту розділами, та завданнями, а також редагування вмісту тесту. Розділи та завдання тесту є посиланнями на окремі об'єкти, що включені в тест але мають свої властивості та методи. Разом з наповненням тесту завданнями на другому

етапі відбувається процес рецензування вмісту тесту. Рецензування дозволяє виявити помилки та невідповідності вмісту тесту ще на перших етапах життєвого циклу.

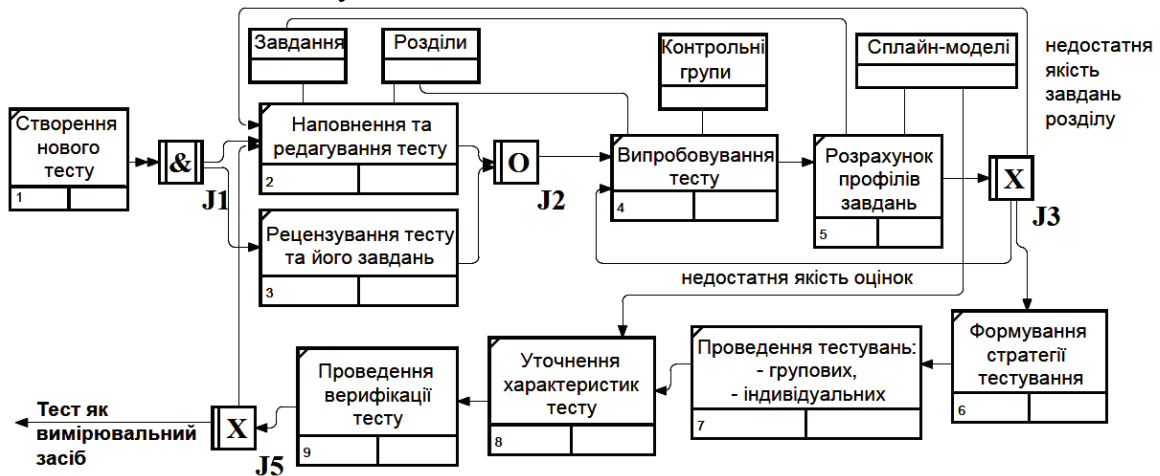


Рис. 1. Діаграма PFDD за стандартом IDEF3 технології розробки тесту

Після наповнення тесту розділами та завданнями в достатній мірі, на перехресті J_2 , процес наповнення може бути припинений. В подальшому другий етап може застосовуватись для внесення змін в тест та його складові при виявленні серйозних порушень. Етап рецензування може продовжувати виконуватись і на наступних стадіях.

На четвертому етапі життєвого циклу тесту виконується його випробовування. На цьому етапі проводиться тестування на контрольних групах респондентів за розділами тесту. Тому існують посилання на об'єкти "Розділи" та "Контрольні групи". Контрольна група респондентів має відповідати цільовій аудиторії для якої ведеться розробка тесту, тоді отримані результати тестування матимуть більшу цінність. Для зниження вартості цієї процедури доцільно залучити до цього процесу широке коло тестувальників. Процедура аналогічна бета тестуванню програмного забезпечення. Тести передаються бета тестувальникам на умовах, які передбачають їхню зацікавленість і відповідальність. Оперативний зв'язок доцільно здійснювати через електронну пошту або інтерактивні сторінки. До розроблювачів тесту повинна надходити інформація про характеристики тестових груп і результати тестування.

На п'ятому етапі за результатами тестування проводиться оцінка характеристик тесту. Оцінка профілів тестових питань вимагає значного числа респондентів із різним рівнем знань, тому практично доцільною є ітераційна процедура оцінки характеристик тесту. За умов використання сплайн-моделей профілів IRT можлива повна автоматизація розрахунків профілів завдань тесту і отримання характеристик якості. Застосування таких засобів суттєво спрощує процедуру оцінку якості і виводить розробку тестів на якісно новий рівень. На перехресті J_3 визначається один із наступних напрямків переходу. За умови, що завдання тесту мають

недостатню якість оцінок виконується повернення до четвертого етапу і проводяться додаткові випробовування. За умови, що були виявлені недоліки самих завдань тесту, здійснюється перехід до другого етапу та вносяться зміни в тест. Якщо недоліки не виявлені і якість тесту є достатньою, то на шостому етапі відбувається подальше формування стратегії тестування.

Протягом визначеного періоду часу, згідно стратегії тестування сформованій на шостому етапі, виконується процес тестування на групах та індивідуальний (сьомий етап). За результатами проведених тестувань вже можна визначати рівні знань з врахуванням рівня складності завдань. Згодом необхідно провести уточнення характеристик тесту (восьмий етап). Уточнення характеристик дозволяє виявити приховані недоліки, що могли проявитись лише з часом. При досягненні необхідного рівня вірогідності оцінок профілів тестових питань виконується оптимізація тесту (наприклад, вилучення малоінформативних питань і питань зі статистично близькими профілями). На дев'ятому етапі верифікація тесту. За необхідності внесення змін в тест на перехресті *J4* можна здійснити перехід на другий етап.

Якщо якість тесту задовольняє вимогам то зміни не вносяться. Формується паспорт тесту, що містить профілі питань. Наявність цих характеристик дає можливість більш об'єктивно оцінювати рівень знань респондентів. Із цього моменту тест можна вважати завершеним інформаційним продуктом із нормованими характеристиками. Тобто одержуємо своєрідний вимірювальний прилад, що вимірює знання в певній області.

Таким чином, виходячи з процесів, описаних на діаграмі PFDD технології розробки тесту, побудуємо діаграму стану та трансформації в процесі основного об'єкта – тесу. Діаграма Object State Transition Network (OSTN) будується за стандартом IDEF3 показана на рисунку 2.

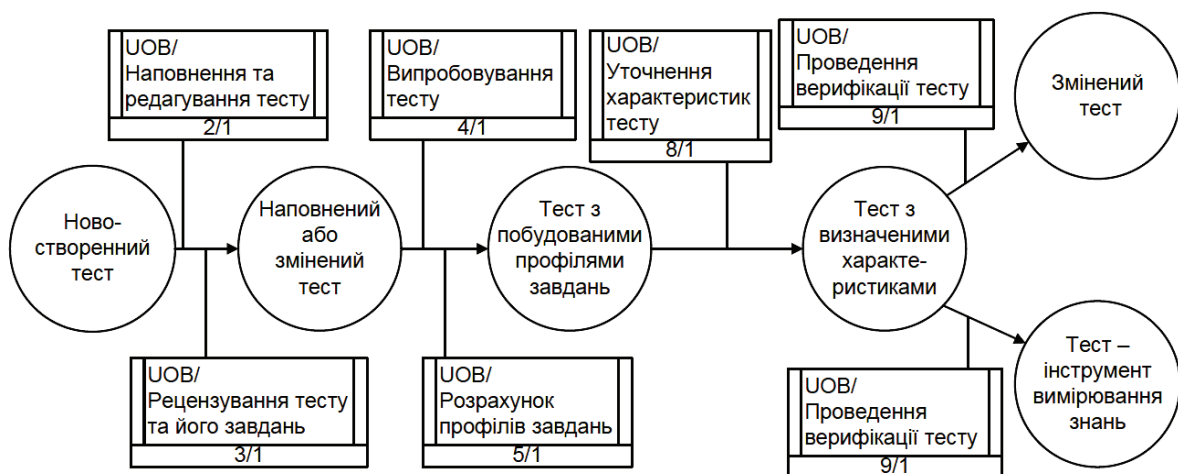


Рис. 2. Діаграма OSTN за стандартом IDEF3 класу "Тест"

На діаграмі стани об'єкту «Тест» відображено колами, а їх зміни направленими лініями. Кожна лінія має посилання на відповідний функціональний блок, в результаті виконання якого відбулась зміна основного об'єкта.

Інформаційна технологія дозволяє в результаті виконання всіх етапів отримати тест з визначеними характеристиками якості. Такий підхід до формування, випробуванню й застосуванню тестів здатний гарантувати необхідний (або заданий) рівень вірогідності тестування, перетворивши тест на вимірювальний інструмент із нормованими характеристиками щодо похибок. Звичайно, бажано встановити певні стандарти на такі вимірювальні інструменти.

Об'єктна модель. В процесах розробки та підтримки тесту за представленою технологією окрім основного об'єкту приймають участь і додаткові об'єкти, що пов'язані між собою функціональними зв'язками. Укрупнена діаграма класів в нотатії UML, що відображає основні класи системи та взаємозв'язки між цими класами, зображено на рисунку 3.

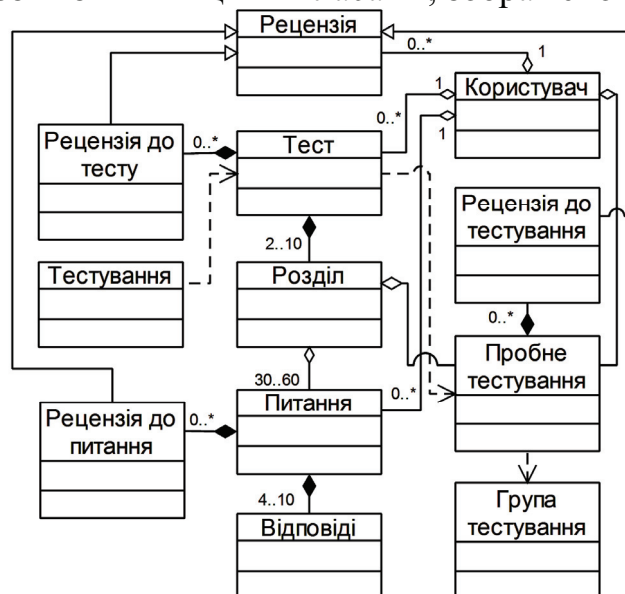


Рис. 3. Діаграма класів інформаційної технології

Така схема дозволяє зосередити увагу на основних об'єктах системи та їх властивостях на перших етапах проектування інформаційної системи та написання програмного коду з використанням об'єктно-орієнтованої моделі. Кожний клас інформаційної системи має характерні йому властивості та методи, що забезпечать функціонування системи. «Тест» є центральним об'єктом системи. Властивості характерні класу «тест»: область знань (наприклад, за УДК), назва тесту, дисципліна, анотація, автор тесту, адміністратор тесту, стадія, дата створення, дата редагування, якість, статус. Клас «Розділ» знаходиться із класом «Тест» у відношеннях композиції, має наступні властивості: тест з яким пов'язаний та назва розділу. Розділів в тесті має бути від двох до десяти. Між класами

«Питання» та «Розділ» установлене відношення агрегації, бо питання можна переміщати в інші розділи. Властивості об'єкта класу «Питання» наступні: тест та розділ до якого він належить, дата створення, дата редагування, профіль, складність, диференційна здатність, текст питання, статус, автор питання. Питань по кожному розділу має бути близько тридцяти-п'ятдесяти. У відношенні композиції до класу «Питання» знаходиться клас «Відповіді» і має наступні властивості: питання до якого належить, текст відповіді, правильність відповіді. Відповідей по кожному питанню має бути від чотирьох до десяти.

Головними суб'єктами системи є користувачі, їх властивості описані в класі «Користувач»: назва користувача, пароль, прізвище, ім'я, по батькові, e-mail адреса, icq номер, skype, рівень доступу, статус тестувальника, статус автора, статус рецензента, науковий ступінь, вчене звання, вид діяльності, анкета, дата реєстрації, дата останніх змін у профілі, дата останнього візиту, статус користувача, ідентифікатор сесії, кількість спроб входу з неправильним паролем. Усі дії в системі виконуються користувачами, тому клас «Користувач» майже з усіма класами знаходиться у відношенні агрегації. Користувачі можуть мати різні набори статусів та рівнів доступу.

На всіх етапах формування тесту відбувається процес рецензування що забезпечується класами: «Рецензія до тесту», «Рецензія до питання», «Рецензія до тестування», похідними від класу «Рецензія». Властивості цих класів: об'єкт рецензування, автор рецензії, текст рецензії, дата рецензії. Рецензії мають вигляд обговорення-коментування об'єкта з яким вона пов'язана відношенням композиції.

Підчас випробування тесту тестування відбувається з використанням класу «Пробне тестування». Результати пробних тестувань дають змогу побудувати профілі питань та визначити характеристики якості тесту. Властивості класу «Пробне тестування» включають: тест та розділ за яким проводиться пробне тестування, дата проведення, тестувальник що проводив тестування, групи респондентів, результати тестування.

На результати пробних тестувань впливають характеристики груп респондентів на яких проводилось тестування. Властивості класу «Група тестування» наступні: назва групи, характеристика групи. Після закінчення вдосконалення тесту, характеристики його якості впливатимуть на процес тестування.

Для побудова профілів IRT використовуються авторські сплайн моделі.

Сплайн-моделі IRT. Суть IRT полягає в тому, що тестові питання та респонденти характеризуються деякою функціональною залежністю від складності питання та рівня знань респондента. Ці складові є латентними, прихованими, а теорія встановлює зв'язок між множинами цих характеристик. Теорія базується на математичних моделях, що дозволяють

будувати профілі складності завдань тесту та респондентів. Кількість математичних моделей IRT постійно збільшується. Причиною цього є, насамперед, значний інтерес до цих питань в освіті, бажання отримати більш точну, надійну та просту у використанні модель. Різні автори в IRT пропонують нові параметри та їх комбінації, обґрунтовуючи необхідність їх врахування і застосування. Незважаючи на велику кількість моделей в IRT значного поширення і застосування набула лише однопараметрична модель Г.Раша, проте і її застосування є досить складним процесом. Нелінійна залежність моделі від параметрів значно ускладнює оцінювання функцій за емпіричними даними. Так, якщо модель не може адекватно описати емпіричні данні, то емпіричні данні вибраковуються і не використовуються. Такий підхід порушує цілісність емпіричних даних і є малопродуктивним.

Тому серед моделей профілів особливої уваги заслуговують сплайн-моделі, які дозволяють описувати функціональні залежності достатньо складної форми, у поєднанні із простотою розрахунків та лінійною залежністю від параметрів. Не зважаючи на це, застосування сплайн-моделей в IRT не набуло широкого вжитку. Розроблені сплайн-моделі IRT [4] використовують в якості базисів лінійні поліноми, так звані В-сплайни, що не враховують особливостей профілів IRT. Адекватність сплайн-моделі профілям IRT досягається правильним підбором базису сплайна, який забезпечить виконання апріорних умов профілів.

В інформаційній технології підтримки тестового контролю знань, передбачено використання розроблених сплайн-моделей профілів IRT:

- ермітів кубічний сплайн з фіксованими краями (SHF);
- інтеграл від лінійного В-сплайна (SLI).

Детальний опис цих моделей наведено в роботах [5, 6], особливістю яких є виконання апріорних умов профілів IRT. Застосування адекватних сплайн-моделей профілів IRT дозволяє будувати профілі завдань тесту в автоматичному режимі без участі оператора-експерта (рис. 4).

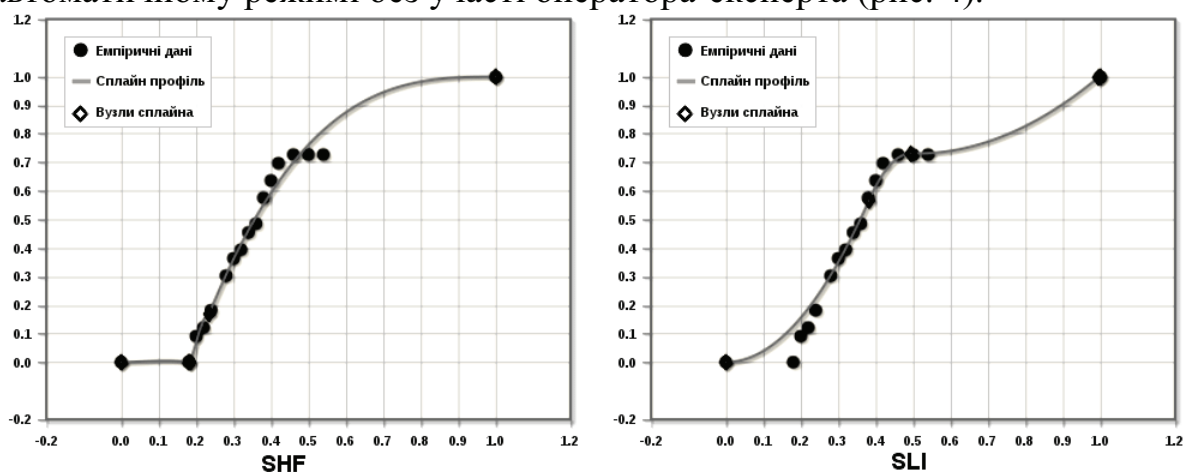


Рис. 4. Автоматично побудовані профілі завдань тесту

Завдяки лінійній залежності сплайн-моделей від параметрів можливе використання комбінованого критерію оптимізації розміщення вузлів, що дозволяє досягти мінімум відхилення сплайн-профілю від емпіричного профілю. Відомі параметричні моделі IRT не можуть забезпечити побудову профілів в автоматичному режимі.

Висновки. Запропонована інформаційна технологія підтримки тестового контролю знань і сплайн-моделі IRT знайшли своє втілення в інформаційній системі “Logit”. Інформаційна система дозволяє спростити процедуру розробки якісних тестів і в результаті виконання отримати готовий інформаційний продукт – тест як інструмент вимірювання знань. Система побудована виключно на вільному програмному забезпеченні, є web-орієнтованою і знаходиться у вільному доступі за адресою [http://logit.kdpu.edu.ua]. Система була впроваджена у Криворізькому педагогічному інституті ДВНЗ “Криворізького національного університету”. За результатами впровадження було розроблено тест з фізики для 11 класів за назвою “Атомна та ядерна фізика” та за результатами тестувань на контрольних групах розраховані сплайн-профілі завдань тесту.

Використані джерела інформації:

1. Указ Президента України від 08.09.2010 року № 895 “Про заходи щодо визначення і реалізації проектів із пріоритетних напрямів соціально-економічного та культурного розвитку”. // Офіційний вісник Президента України. - 2010. - № 26. - С.15.
2. Комплекс нормативних документів для розробки складових системи вищої освіти. Додаток 1 до Наказу Міністерства № 285 від 31 липня 1998 р. / Г.Я. Антоненко, І.С.Булах, В.Л. Петренко та ін. – К. : Інститут змісту і методів навчання, 1998. – 124 с.
3. Ким В.С. Тестирование учебных достижений. Монография. Уссурийск: УГПИ, 2007. – 214 с.
4. Michal Abrahamowicz, James O. Ramsay. Multicategorical spline model for item response theory / Psychometrika , 1992, vol. 57, no. 1, pp. 5-27.
5. Дубан Р.М., Шелевицький І.В. Сплайн-моделі профілів складності питань та знань респондентів в тестовому контролі знань.// Автоматизированные системы управления и приборы автоматики: всеукр. межведомств. науч.-техн. зб. – Харьков, 2011. – В.156. - С. 71-77.
6. Дубан Р.М., Бойко І.Ф. Застосування інтегралу від лінійного В-сплайна в якості моделі IRT. // Електроніка та системи управління. – Київ, 2012. - №1(31). – С 131-138.