

Рис. 2. Редагування параметрів тестових завдань шляхом використання команди Find/Replace.

Висновок. Знання структури побудови тестових завдань в форматі Moodle XML дозволяє виконувати редагування параметрів безпосередньо в Moodle XML файлі без проміжних перетворень в більш «читабельний» вигляд.

Запропонована методологія редагування параметрів тестових завдань дозволяє виконувати одночасні зміни параметрів великої кількості тестових питань зменшуючи час для внесення правок.

Література

1. Moodle XML format [Електронний ресурс] – режим доступу http://www.qedoc.org/en/index.php?title=Moodle_XML_format.
2. Moodle XML format [Електронний ресурс] – режим доступу http://docs.moodle.org/23/en/Moodle_XML_format.
3. Using Moodle, 2nd Edition. Chapter 5. Quizzes [Електронний ресурс] – режим доступу http://download.moodle.org/docs/en/using_moodle/ch5_quizzes.pdf.
4. XML [Електронний ресурс] – режим доступу <http://en.wikipedia.org/wiki/XML>

УДК 378.146

Сергій Подласов

Національний технічний університет України «Київський політехнічний інститут»

СТАТИСТИЧНИЙ АНАЛІЗ СИСТЕМИ ТЕСТОВИХ ЗАВДАННЯ ДЛЯ ПОТОЧНОГО КОНТРОЛЮ НАВЧАЛЬНОЇ ДІЯЛЬНОСТІ СТУДЕНТІВ ПРИ ВИВЧЕННІ КУРСУ ФІЗИКИ

© Сергій Подласов, 2012

В роботі представлені результати статистичного аналізу результатів поточного контролю навчальних досягнень студентів на основі Item Response Theory. Одержані результати використані для удосконалення системи тестових завдань.

Ключові слова: тестування, тестові завдання, рівень складності завдань.

The Item Response Theory was used for statistical analysis of tests results of current control of the students knowledge and for improvement of the system of test tasks.

Keywords: testing, test questions, level of task difficulty.

Вступ. Згідно з сучасною тенденцією розвитку освіти значна частина навчального матеріалу переноситься на самостійне опрацювання студентів. За таких умов важливого значення набуває моніторинг результатів їх навчальної діяльності. В останній час для цього все ширше застосовуються тестові форми контролю, що забезпечує його технологічність, об'єктивність та оперативність. Незважаючи на те, що у чималій кількості науково-методичних робіт, присвячених використанню тестових методик, описуються різноманітні системи тестових завдань та результати їх застосування, тільки у незначній частці представлений аналіз якості як тестових завдань, так і тестів у цілому. Саме тому актуальним постає завдання не тільки розробки, апробації, але й визначення якості складеної сукупності тестових завдань та пошуку засобів для їх покращення.

Тести як засіб контролю навчальної діяльності. Моніторинг результатів навчальної діяльності має на меті встановлення рівня оволодіння студентами запланованим на даний період часу навчальним матеріалом, при цьому не виникає потреби у встановленні статистичного розподілу студентів за рівнем їхніх навчальних досягнень. Це дозволяє тести для такого контролю вважати критеріально орієнтованими (КОрТ).

При створенні будь-яких тестів, і зокрема КОрТ, повинна бути вирішена низка важливих завдань, серед яких чільне місце посідають два: забезпечення змістової валідності при оптимальній кількості завдань та забезпечення об'єктивності контролю.

Для забезпечення змістової валідності, тобто повноти представлення завдань для перевірки набутих студентами знань та умінь, повинен бути складений детальний перелік (специфікація) знань та умінь у даній предметній області, які підлягають контролю. Наприклад, у розділі «Кінематика» треба перевірити засвоєння студентами основних понять (шлях, переміщення, середня шляхова швидкість, середній вектор швидкості, нормальне і тангенціальне прискорення, вектор кутової швидкості та кутового прискорення, перетворення швидкостей та прискорень при переході з однієї системи відліку в іншу). Крім цього, треба також перевіряти володіння студентами елементами математики: векторною алгеброю, диференціюванням та інтегруванням. Останнє є дуже важливим, оскільки саме необхідність

застосовувати новий для студентів першого курсу математичний апарат може зумовлювати нерозуміння ними усього навчального матеріалу, який має бути засвоєним.

Питання про оптимальну кількість завдань повинно вирішуватися, враховуючи обсяг матеріалу, який підлягає перевірці, та реальний бюджет часу студентів. За результатами нашого опитування студенти мають можливість виділити на тестування 1,5 – 2 години на тиждень. З досвіду тестування учнів та студентів відомо, що на виконання окремого завдання вони витрачають у середньому 3 – 5 хвилин. Таким чином, оптимальним можна вважати тест, який складається з 20 – 25 завдань.

Об'єктивність контролю передбачає виконання декількох умов. Перш за все, повинні бути забезпечені однакові умови організації проходження студентами тестування, однакова «ступінь новизни» завдань та рівень їхньої складності.

Що стосується однакових умов проходження тестування, то нами була вибрана дистанційна форма, коли всі студенти виконують завдання в мережі Інтернет у зручній для них час і у зручному місці.

Однакова ступінь новизни і однаковий рівень складності завдань точно можуть бути забезпечені якщо всі студенти одночасно виконують одні й ті ж завдання. При дистанційному тестуванні, коли завдання використовуються неодноразово і студенти виконують їх неодноразово, зміст завдань і правильні відповіді доволі швидко стають відомими широкому загалу, що порушує однаковість умов, відтак і об'єктивність результатів контролю. Для запобігання цьому по кожному із запланованих для перевірки пунктів треба сформувати групи з достатньо великої кількості завдань приблизно однакового рівня складності, з яких випадковим чином вибирати завдання при компонуванні тесту для кожного окремого студента.

Рівень (інакше – індекс) складності (легкості) завдання є поняттям доволі складним, яке визначається як об'єктивними, так і суб'єктивними факторами. Об'єктивними факторами є кількість розумових дій, які повинен здійснити студент для виконання завдання, та трудність, інакше – «рівень фізичної прихованості» правильної відповіді, а суб'єктивним – рівень підготовленості самого студента.

Дійсний рівень складності кожного завдання можна визначити за результатами його апробації. При цьому рівень складності (легкості) визначають як частку неправильних (правильних) відповідей від їх загальної кількості. Для одержання надійних результатів за цією методикою треба мати репрезентативну вибірку відповідей на кожне із завдань тесту, оскільки інакше можуть виникати ситуації, коли дане завдання виконують переважно або

добре, або ж слабо підготовлені студенти, що призводить до неадекватного оцінювання рівня його складності. Таким чином, за умови випадкового вибору завдань з великої бази даних постає проблема застосування такої методики аналізу результатів тестування, яка б мінімально залежала від рівня підготовленості студентів. Таку методику, основу на логістичній моделі, запропонував Г. Раш [1] і її називають “Item Response Theory” (IRT). У цій теорії в основу покладено припущення, що рівень підготовленості студента та рівень складності завдання є латентними параметрами. Обидва ці параметри оцінюється в одній і тій самій шкалі логитів¹, що визначає відносну інваріантність параметрів складності завдань, їх незалежність від властивостей вибірки студентів, відтак, більш об’єктивне оцінювання дійсного рівня складності завдань.

Для поточного контролю навчальної діяльності студентів при вивченні курсу загальної фізики нами були складені завдання по окремих темах за попередньо розробленою специфікацією. З метою встановлення глибини розуміння студентом певних питань і рівня оволодіння необхідними вміннями завдання були поділені на три рівні складності. Завдання першого рівня повинні перевіряти знання законів, закономірностей, постулатів, принципів, понять, означень, моделей, фактів. Завдання другого рівня повинні перевіряти вміння студентів застосовувати знання на практиці в знайомих ситуація, або ж у подібних до них, а завдання третього рівня – вміння застосовувати знання в незнайомих ситуаціях, або ж використовувати знання з різних розділів курсу фізики. У тест включалося 40 – 50 % завдань першого рівня, та по 20 – 30 % завдань другого та третього рівнів. Складені нами завдання розміщені на сайті Українського інституту інформаційних технологій в освіті (<http://uiite.org>), де для їх підготовки, зберігання, пред’явлення студентам та зберігання результатів тестування використовується програмна платформа підтримки навчального процесу Moodle.

Корекція первинного поділу за рівнями складності завдань проводилася в два етапи. На першому для цього використовувалися дані, які автоматично обчислюються математичним пакетом платформи Moodle: індекс легкості тестового завдання (частка правильних відповідей), середнє квадратичне відхилення (дисперсія), індекс диференціації та коефіцієнт диференціації

¹ "Логит рівня знань" – натуральний логарифм відношення частки правильних відповідей випробовуваного на всі завдання тесту, до частки неправильних відповідей; "логит рівня трудності завдання" – натуральний логарифм відношення частки неправильних відповідей на завдання тесту до частки правильних відповідей на це завдання по безлічі випробовуваних.

(коефіцієнт кореляції між множиною значень відповідей, одержаних студентами при виконанні конкретного завдання [2]. На другому етапі аналіз результатів апробації тестових завдань проводився згідно з теорією ІРТ. Такий аналіз полягає у побудові характеристичних кривих – ймовірності правильної відповіді студента на певне тестове завдання залежно від рівня складності цього завдання (виражений у логітах). Для побудови характеристичних кривих була складена спеціальна програма відповідно до алгоритму, описаному М. Б. Челишковою [3].

Сумісне застосування даних, одержаних в Moodle та за теорією Г. Раша дозволило сформувати групи завдань приблизно однакового рівня складності. На рис. 1 показані характеристичні криві тестових завдань рівнів 1, 2 та 3, які були одержані в результаті аналізу для завдань з розділу «Кінематика». Аналогічні результати були одержані для розділів «Динаміка матеріальної точки» та «Закони збереження в механіці».

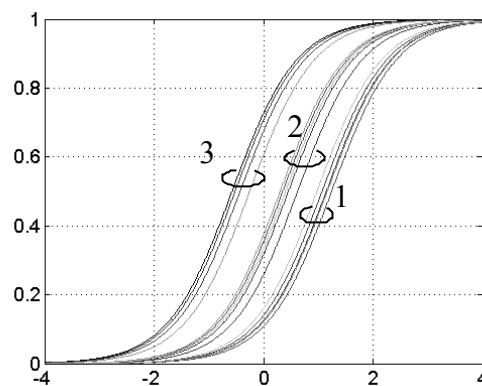


Рис. 1

Для перевірки можливості використання результатів дистанційного тестування для визначення рівнів складності завдань проводилося порівняння результатів такого тестування та результатів контрольних робіт, які проводилися під контролем викладача тими самими студентами. Виявилось, що коефіцієнт кореляції між результатами цих робіт лежить у діапазоні 0,3 – 0,65 і тим вищий, чим вище середній бал групи студентів.

Висновки. Проведений статистичний аналіз результатів тестування студентів дозволив добирати групи завдань приблизно однакового рівня складності для формування КОрТ. Показано, що результати дистанційного тестування можна застосовувати для коректного аналізу якості тестових завдань.

Література

1. Rasch G. *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research. – 1960. – 126 p.
2. Подласов С.О. Використання можливостей системи підтримки навчального процесу Moodle для удосконалення тестових завдань з курсу загальної фізики / Інноваційні комп'ютерні технології у вищій школі: Матеріали 3-ї науково-практичної конференції. М. Львів, 18-20 жовтня 2011 року / Відп. за випуск Л.Д. Озірковський – Львів: видавництво Львівської політехніки, 2011. 168 с. 3.

Чельшкова М.Б. Теория и практика конструирования педагогических тестов: Учебное пособие. - М: Логос, 2002. - 432 с: ил.

УДК 378.14

Олександр Диховичний, Анна Дудко, Галина Філіппова
Національний технічний університет України «КПІ»

АВТОМАТИЗОВАНА СИСТЕМА АНАЛІЗУ РЕЗУЛЬТАТІВ ЕЛЕКТРОННОГО ТЕСТУВАННЯ

© Олександр Диховичний, Анна Дудко, Галина Філіппова, 2012

У праці проінформовано про розробку в НТУУ «КПІ» автоматизованої системи статистичного аналізу результатів електронного тестування на основі як класичних статистичних методів, так і сучасних математичних моделей тестів.

Ключові слова: автоматизований аналіз тестових завдань, латентні параметри, IRT- моделі .

The present paper informs about development of computer-based statistical analysis system of results of online testing. This system is based on classic statistical methods and modern mathematical model of tests.

Keywords: computer-based statistical analysis system, latent parameters, IRT-models.

Вступ. Кафедрою математичного аналізу та теорії ймовірностей НТУУ «КПІ» було створено комплект дистанційної освіти «Вища математика» [1]. Розвиток існуючих електронних курсів та створення нових потребує вдосконалення та підвищення ефективності їх тестової бази. Водночас зі збільшенням обсягів тестування застосування сучасних математичних методів до аналізу тестів стає можливим тільки на підставі застосування сучасних інформаційних технологій. У зв'язку з цим останнім часом значна увага приділяється автоматизації аналізу якості тестових завдань комплекту «Вища математика», для чого й створюється відповідна автоматизована система.

Функціональні можливості системи. В основу системи покладено як класичні статистичні методи, так і сучасні математичні моделі аналізу тестових завдань.

Класичні методи передбачають:

- первинний аналіз вибірки;
- розрахунки основних вибірковоїх статистичних характеристик;
- аналіз розподілу вибірки ;
- порівняння вибірок;
- елементарний кореляційний і дисперсійний аналіз.