

Література

1. Ромапов А. В. Методика подготовки и проведения тестового контроля в учебном процессе. — Чобоксары: Клио, 1998. — 145 с.
2. Палій Л. Тестування в навчальному процесі // Шлях освіти. — 2000. — № 2. — С. 37–39.
3. Информационные технологии обучения в профессиональном образовании // Информатика и образование. — 1996. — № 1. — С. 13–19.
4. Пардюжев В. И., Пардюжев И. В. Модели и алгоритмы информационной системы компьютерного тестирования: Монография. — М.: Прометей, 2000. — 412 с.
5. Авапсова В. С. Научные основы тестового контроля знаний. — М.: Исслед. центр, 1994. — 135 с.

УДК 371.26

В. В. Гаращук**ЗАСТОСУВАННЯ ТЕСТОВОЇ СИСТЕМИ
ДЛЯ ОЦІНКИ ЯКОСТІ ПІДГОТОВЛЕНОСТІ СТУДЕНТІВ**

У загальному випадку роботу викладача можна розбити на два етапи — етап навчання і етап перевірки знань. Етап навчання, при всьому його різноманітті, вимагає окрім вільного володіння висловлюваним матеріалом, постійного застосування творчого підходу, адаптуючи інформаційний потік під конкретну групу слухачів або персонально до кожного з них. При цьому доводиться враховувати безліч індивідуальних особливостей студентів. На цьому етапі для підвищення ефективності навчання можна керуватися лише деякими загальними рекомендаціями, шукати аналогію в типових ситуаціях і прикладах.

По-іншому йде справа з етапом оцінки одержаних знань. Тут існують напрацьовані і затверджені методики, покликані зробити цей етап якомога більш прозорим і об'єктивним. І, хоча неприпустимо повністю виключати з освітнього процесу живе спілкування учня і наставника, багато проміжних етапів можна автоматизувати, при цьому перетворивши їх з простої і даремної пробіжки очима по переліку контрольних запитань, в процес свідомий і невід'ємний. Відповідаючи на грамотно складені запитання тесту і одержуючи в результаті залік або незалік, студент упевнений в об'єктивності своєї оцінки, неупередженості екзаменатора і не тішить себе думкою, що наступного разу він потрапить під більш прихильний настрій або йому пощастить більше. Крім того, сьогодні на одного викладача припадає в середньому від 100 до 200 студентів, а обсяги матеріалу збільшуються з року в рік, тому навіть чисто технічно, в «ручному режимі», складно ефективно контролювати засвоєння нового матеріалу. Практика контролю знань в навчальних закладах і аналіз останніх публікацій показує, що в навчальному процесі все частіше використовується педагогічне тестування.

Педагогічний тест прийнято визначати як систему завдань певного змісту, специфічної форми, що дозволяє якісно і ефективно виміряти рівень і оцінити

структуру підготовленості учнів, контролювати результат засвоєння ними в процесі навчання знань і умінь [1]. Найпоширенішим є тестування, засноване на вживанні персональних комп'ютерів. Це дає можливість автоматизувати процес перевірки відповідей, значно скоротити час, необхідний на обробку результатів контролю.

Сучасний етап розвитку теорії і досвід вживання тестування в навчальному процесі припускають вживання двох видів комп'ютерних тестів: лінійного (статистичного) і адаптивного (динамічного) [2; 3].

Лінійний — це традиційний тест з випадковим вибором завдань (питань) з кожної групи завдань тесту. Завдання є елементарною одиницею будь-якого тесту. Такий тест пропонує різноманітні завдання — від найпростіших до найскладніших. Результуюча оцінка виставляється залежно від кількості набраних балів при відповіді на завдання тесту.

Не завжди повна і достовірна інформація про реальний рівень знань студентів і існуючі відмінності в їх здібностях стає головною причиною появи адаптивних тестів, заснованих на принципі індивідуалізації навчання. В адаптивних тестах кожний, хто тестується, одержує завдання по рівню складності у відповідність з його знаннями по даній дисципліні.

Адаптивний тест — це тест, при якому складність подальших завдань залежить від результатів відповідей на попередні завдання.

Алгоритми адаптивного тестування реалізуються таким чином: на кожному кроці тестування випробовуваному пред'являється група завдань певного рівня складності. Після відповіді на задану групу завдань аналізуються відповіді і формується оцінка. Залежно від оцінки, одержаної за попереднє завдання, вибирається наступне завдання або група завдань. Рівень складності чергового завдання повністю залежить від якості виконання попереднього. Тестування звичайно починається із завдань середньої складності. В процесі виконання завдань адаптивного тесту більш точно визначається оцінка рівня знань випробовуваного. Головними параметрами оцінки якості тестів є надійність і валідність.

Надійність — стійкість результатів тесту, одержуваних при його вживанні. Надійність результатів адаптивного тестування дуже висока, оскільки здійснюється пристосування під рівень знань конкретного слухача, що забезпечує більш високу точність оцінки знань.

Валідність — придатність тесту, тобто здатність якісно оцінювати те, для чого він створений.

Під довжиною тесту розуміється кількість завдань, що входять в тест. Теорія затверджує: чим довше тест, тим він надійніше. Практика тестування показує, що дуже довгий тест погіршує мотивацію і увагу. Оптимальна довжина тесту знаходиться в межах 30–60 завдань. Співвідношення довжини тесту до числа тестових завдань в банку повинне наближатися до співвідношення 1:10.

Порівняння статистичних і адаптивних тестів дозволяє наголосити на такому:

- Статистичний тест дає не достатньо об'єктивну оцінку знань, оскільки результати тестування завжди залежать від конкретного тесту. На відміну від

нього адаптивний алгоритм дозволяє визначити рівень підготовленості слухача залежно від його знань.

- З психологічної точки зору адаптивне тестування краще і ефективніше статистичного у тому випадку, коли статистичний тест виявляється дуже складним або навпаки дуже легким. При адаптивному тестуванні студенту пред'являються завдання, рівень складності яких відповідає його істинним знанням. Таким чином, слухач знаходиться в постійній розумовій напрузі, і завдання для нього є стимулюючими.

- При неодноразовому проходженні адаптивного тестування тестованому пропонуються завдання, відповідні рівню його знань, одержаному в ході попереднього тестування. Таким чином, формується деяка методика закріплення добре вивчених розділів і методика поступового вивчення слабо засвоєних складніших розділів.

Проте при всіх своїх перевагах адаптивні тести дуже складні для створення, отже — дорогі. Крім того, така система є недостатньо відкритою для розвитку, вдосконалення і модернізації через велику кількість зв'язків і переходів. Ці недоліки можуть бути визначаючими для освітніх напрямів, що динамічно розвиваються. Наприклад, в юридичних і економічних науках, а також напрямках, пов'язаних з сучасними інформаційними системами, системами телекомунікацій та іншими. В таких випадках більше підходять лінійні тести. А для того, щоб підвищити їх якість, потрібна незначна модифікація, доступна практично в будь-якій вже функціонуючій програмі комп'ютерного тестування. Для підвищення надійності необхідно скласти завдання декількох рівнів складності в залежності від шкали оцінювання. Частіше за все студент наперед знає, на яку оцінку він може претендувати і вибрати для себе бажаний рівень складності. Або по черзі проходити тести від простих до більш складних. Позитивною стороною такого методу є те, що студент прагнутиме підвищити рівень своїх знань, щоб пройти тест більш високого рівня складності. Цей метод вже декілька десятиріч успішно застосовується в індустрії комп'ютерних ігор, і той, хто тестується, з легкістю приймає такі, для багатьох звичні, правила.

На ринку програм комп'ютерного тестування лише деякі з них можуть проводити такий аналіз і видавати відповідні рекомендації. Але при цьому майже всі тестові системи дозволяють встановлювати вагові коефіцієнти для кожного із завдань тесту. Використовуючи цю особливість, можна вдосконалити будь-який існуючий лінійний тест з метою підвищення його якості та надійності, а також видачі кількості правильних відповідей по кожній темі тесту. Така доробка особливо буде корисна для таких тестів, що об'єднують завдання по великій кількості тем, наприклад заліковий тест по всьому курсу чи по окремому його розділу. Для такої доробки (або створення нового лінійного тесту з поліпшеними характеристиками) необхідно:

1. Розподілити всі завдання по темах курсу.

2. Розподілити завдання по кожній темі по групах складності, так, щоб в кожній групі були завдання в кількості, приблизно в 10 разів перевищуючій кількість завдань, що пред'являються випробовуваному по кожній темі.

3. Завдання з однаковим рівнем складності об'єднати в один тест, налаштувавши тестову систему так, щоб з кожної теми випадково вибиралася рівна кількість завдань.

4. Призначити вагові коефіцієнти таким чином: всім завданням з першої теми, у разі правильного рішення, призначаються вагові коефіцієнти 1, завданням другої теми — 0,1, третьої — 0,01 і т. д.

Таким чином, після проходження тесту, по набраній кількості балів можна судити про те, на скільки питань по кожній темі студент відповів правильно, не уточнюючи, які саме це були питання. Наприклад, якщо було набрано 0,10635 балів, це означає, що по першій темі не було правильних рішень, по другій — 1, по третій — 0, по четвертій — 6 і т. д. Вагові коефіцієнти, в даному застосуванні, позбавлені звичного для них змісту — надавати кількісну оцінку завданням тесту. Тут їх призначення виявити структуру підготовленості студентів під час проходження лише одного тесту. Кратність вагових коефіцієнтів $\times 0,1$ була вибрана саме такою не випадково. По-перше — і викладач і учень мають можливість досить легко проаналізувати результати, тому що кожен розряд, при відображенні результуючої кількості балів, відповідає за свою тему. По-друге, результат можливо легко розшифрувати навіть у тому випадку, коли дана правильна відповідь лише на завдання однієї теми, проводячи декодування звично зліва направо. Так, наприклад, коли результуюча кількість балів дорівнює 0,3, це означає, що по другій темі — 3 правильних рішення, а по іншим — їх немає, при цьому не потрібно тримати в голові загальну кількість тем для кожного тесту. Для впровадження такої модернізації необхідно, щоб тестуюча програма мала можливість вводити вагові коефіцієнти для кожного завдання, бажано у вигляді десяткового дробу.

Дане вдосконалення дає можливість таким тестам оцінити структуру підготовленості слухачів, а по своїх характеристиках дозволяє їм наблизитися до адаптивних тестів, при тому що простота побудови і подальшої модернізації залишається на рівні лінійних. Крім того такий алгоритм досить легко реалізувати при розробці нових тестових програм.

Література

1. Палій Л. Тестування в навчальному процесі // Шлях освіти. — 2000. — № 2. — С. 37–39.
2. Майоров А. П. Теория и практика создания тестов для системы образования. — М.: Интеллект-центр, 2001. — 296 с.
3. Мельникова М. Б. Теория и практика конструирования педагогических тестов. — М.: Прометей, 2002. — 186 с.
4. Аванесов В. С. Композиция тестовых заданий. — 2-е изд., испр. и доп. — М.: Адепт, 1998. — 217 с.
5. Казарипов А. С., Култышева А. Ю., Мирошниценко А. А. Технология адаптивной валидности тестовых заданий. — 1999. — 62 с.