

Проблеми створення і застосування комп'ютерних тестів для діагностики знань майбутніх учителів початкової школи

В епоху інформатизації суспільства зростає роль освіти. Із усіх соціальних інститутів саме освіта є основою соціально-економічного і духовного розвитку постіндустріального суспільства. З іншого боку лавиноподібне зростання наукової інформації створює певні проблеми для педагогів. Постають запитання: що вчити? Як вчити? Що і як контролювати? Зросла роль контролю за рівнем засвоєння матеріалу та діагностики знань.

Будь-який контроль за рівнем засвоєння знань має завчасно плануватися. Принципи, що забезпечують ефективність контролю, сформульовані в дидактиці. Цією проблемою опікувалися В.С. Аванесов, Г.А. Атанов, І.Є. Булах, І.С. Олексійчук та інші. Згідно з їхніми дослідженнями до основних принципів контролю знань належать:

- повнота охоплення програмного матеріалу;
- достовірність;
- об'єктивність.

Контроль знань є невід'ємною і важливою частиною педагогічного процесу, що виконує ряд функцій: контрольну, навчальну, виховну, розвивальну, організаційну і методичну. Виконуючи ці функції в процесі підготовки висококваліфікованих фахівців, контроль вирішує також завдання здійснення зворотного зв'язку і керування процесом навчання, оскільки, одержуючи своєчасну і достатню інформацію про те, як протікає процес засвоєння студентами знань, викладач може коректувати процес навчання і перебудувати його відповідно до потреб даної конкретної ситуації, що є важливою умовою підвищення якості підготовки фахівців.

Практика показує, що в умовах скорочення навчальних годин з багатьох дисциплін програми і, як наслідку цього, зростання ролі самостійної роботи студентів, контролю знань повинна приділятися все більш істотна роль. Особливо це стосується таких видів контролю, як поточна і періодична перевірка, яким у ВНЗ не завжди приділяється належна увага в порівнянні з таким видом, як підсумкова перевірка знань.

Внаслідок трудомісткості і великої витрати часу, викладачі вважають за доцільне витратити його на інші компоненти процесу навчання: передачу знань, закріплення, формування умінь і навичок і т.п. Внаслідок цього у ВНЗ нерідко використовується "штурмовий метод", при якому вивчення матеріалу проходить за 3-4 дні до іспиту, коли мобілізується з усього апарата мислення лише оперативна пам'ять, і знання забуваються практично відразу після закінчення іспиту або заліку.

Проблема поточної і періодичної перевірки знань студентів вирішується при застосуванні тестового контролю, що дозволяє опитати всіх студентів одночасно і при мінімальних витратах часу одержати об'єктивний результат.

Застосування комп'ютерів для тестування ще більш підвищує ефективність контролю, тому що дозволяє викладачу систематично перевіряти знання всіх студентів у навчальних групах; скоротити загальний час контролю; збільшити кількість контрольних питань, що задаються кожному; автоматизувати перевірку отриманих відповідей; автоматично і негайно повідомляти результати контролю студентам.

Застосування ЕОМ у процесі навчання також дозволяє студентам прилучатися до комп'ютерної письменності, готує їх до життя серед техніки, що розширює можливості майбутнього фахівця в одержанні роботи в сучасному технологічному суспільстві.

Первинна інформація про результати виконання тесту збирається в матриці фіксації результатів тестування. Багатьма науковцями-практиками рекомендується єдина 100-бальна шкала на основі врахування ваги завдань, що входять до тесту. Вага завдання визначається на експертній основі і може варіюватися від 1 до 10 балів. Вага завдання відображає значимість елемента, що перевіряється, в системі знань. Кожний студент за одне завдання може одержати 0 балів, якщо він неправильно виконав або не виконував це завдання, або ту кількість балів, яку експерт приписав даному завданню, якщо студент правильно виконав його.

Результат виконання тесту визначається як процентне відношення суми балів, одержаних студентом за виконані завдання, до максимально можливої суми балів (рівній сумі ваг усіх завдань), що можна одержати за даний тест. Таким чином, здійснюється перехід в єдину 100-бальну шкалу.

Наведений вище алгоритм перекладу первинної інформації в числову дає можливість однозначно поставити у відповідність діяльності будь-якого студента число; тим самим реалізується процедура виміру. Проте одержані числові дані ще не є оцінкою. Для проведення процедури оцінювання необхідний критерій, що дозволить би виразити якісне судження про одержані числові значення. Такий перехід має проводитися за єдиною стандартизованою шкалою. У стандартизованій шкалі можуть бути використані такі критерії оцінювання:

- від 0 до 50 балів - "незадовільно";
- від 51 до 70 балів - "задовільно";
- від 71 до 90 балів - "добре";
- від 91 до 100 балів - "відмінно".

Ці критерії відображають те, що показник успішності більше 70% у педагогіці вважається показником якісного засвоєння знань.

Інколи таку шкалу будують, виходячи із середнього результату для даного тесту і його

середньоквадратичного відхилення. Так, наприклад, в одному із наших досліджень середній результат для певного тесту виявився рівним 72 балам, а середньоквадратичне відхилення було рівне 19. У такому випадку шкала мала вигляд:

- від 0 до 52 балів - "незадовільно";
- від 53 до 72 балів - "задовільно";
- від 73 до 90 балів - "добре";
- від 91 до 100 балів - "відмінно".

Очевидно, що така шкала мало чим відрізняється від попередньої, а при вазі кожного завдання 5-10 балів не викличе зміни оцінки. Тому для тестів, вага завдань яких не є дуже малою, цілком можна користуватися першою із запропонованих шкал. Саме застосування такої універсальної шкали робить зручним комп'ютерне тестування.

Ситуація ускладнюється, якщо завдань багато і їхня вага незначна (1-2 бали). Для створення таких тестів необхідна копійка робота із кореляції тесту, наприклад за допомогою обчислення коефіцієнта кореляції Пірсона, а також оцінки надійності тесту за коефіцієнтом кореляції Спірмана-Брауна [1,232]. Під час такої довготривалої експериментальної роботи із тестів вилучаються завдання, що не мають інформаційної цінності, а також ті, що слабо пов'язані з іншими завданнями.

Таким чином, робимо висновок, що створення універсального тесту, який би давав реальну оцінку знань з предмету, по кафедрі, з спеціальності, є надзвичайно важким, довготривалим і трудомістким процесом. Такі затрати зусиль упорядника тесту не завжди виправдані. Більш доцільними, з нашої точки зору, є "мінітести", що застосовуються для внутрішнього використання в процесі вивчення конкретної теми чи розділу. Звичайно, спрощення процедури створення і кореляції тестів зменшує їхню досконалість, проте за умови їх частого використання загальна картина одержаних результатів буде досить об'єктивною. Розглянемо найпростіші конструкції тестів.

Тестові завдання поділяють на завдання закритого типу, відкритого типу, на встановлення відповідності і на правильну послідовність.

Тестові завдання закритого типу найбільш поширені. Вони формулюються у вигляді висловлення (речення у стверджувальній формі), до якого пропонується 2-5 варіантів відповідей, серед яких лише одна правильна. Залежно від контингенту тих, кого тестують, рівень складності таких тестів варіюється. Найбільш складними, але й ефективними в дидактичному плані, з нашої точки зору, є варіанти відповідей, що майже всі претендують на правильність, але лише одна з них є повною. Такий тест не просто з'ясовує рівень запам'ятовування теоретичного матеріалу, а й глибину його розуміння, примушує аналізувати відповіді, а тому відіграє ще й навчальну роль.

Загалом розрізняють такі способи формулювання відповідей до тестових завдань закритого типу:

- альтернативний;
- класифікаційний;
- сполучний;
- кумуляційний.

У випадку альтернативного способу пропонуються дві (інколи більше) відповідей, що є протилежними за змістом. Досить часто вони реалізуються словами так і ні або двома антонімами. Наприклад:

Для будь-яких двох натуральних чисел виконується комутативний закон ділення

Імовірність угадування Так Ні відповіді Так Ні тестів становить 0,5, тому достовірність результатів контролю знань не висока. Більш ефективними в цьому плані є тестові завдання з подвійною альтернативою. Наприклад:

Якщо швидкість тіла, що рухається рівномірно і прямолінійно, збільшити в 2 рази, то час на подолання тієї ж відстані

Збільшиться в 2 рази	Зменшиться в 2 рази	Збільшиться в 4 рази	Зменшиться в 4 рази
----------------------	---------------------	----------------------	---------------------

Класифікаційний спосіб передбачає кілька можливих відповідей, кожна з яких в принципі відповідає ситуації, але лише одна з них правильна. Наприклад:

Для дії множення натуральних чисел виконуються такі закони

Комутативний	Комутативний і асоціативний	Комутативний, асоціативний і дистрибутивний відносно додавання
--------------	-----------------------------	--

У сполучному способі пропонуються варіанти, що поєднують у собі частково правильні відповіді, дуже близькі до змісту завдання. Вибрати з них повністю правильну – важке завдання для того, хто тестується.

Кумулятивний спосіб формулювання відповідей відрізняється від сполучного тим, що частка правильних відповідей у кожному з варіантів різна.

Тестові завдання закритого типу легко реалізуються в комп'ютерних програмах. Їхнім недоліком вважають можливість вгадування правильного варіанту відповіді, але при достатній

кількості запитань вгадування не впливає на достовірність результатів тестування. На практиці ми зустрілися ще з одним недоліком цього типу тестів – можливість шифрування студентами правильних відповідей. Сильніший студент запам'ятовує або записує номер питання і номер правильної відповіді на нього, а потім передає студентам іншої групи. Цього недоліка можна уникнути, якщо комп'ютерна програма задаватиме питання випадковим чином, тобто кожного разу порядок завдань буде змінюватися.

Ще одним недоліком завдань закритого типу є те, що в ситуації множинного вибору, яка притаманна таким тестам, студент замість формулювання самостійної думки вибирає готову відповідь. Тому, на наш погляд, пріоритетними мають бути завдання відкритого типу.

Завдання відкритого типу базуються на недомовленості. Варіанти відповідей при цьому не пропонуються, а студент має сам доповнити висловлення таким словом або словосполученням, щоб воно стало істинним. Складність для комп'ютерного тестування такого типу полягає в тому, що правильна відповідь має бути однозначною, не допускати можливості вживання синонімів, оскільки навіть близьке за значенням, але різне за написанням слово буде сприйматися комп'ютером як неправильна відповідь.

Конструювання завдань такого типу є складним процесом для викладача. Особливо важким у цьому випадку є створення дескрипторів – описів допустимих відповідей на задане запитання, які зберігаються в базі даних системи тестування. Найдоцільнішими можна вважати запитання, що передбачають коротку однозначну відповідь. При вживанні довгих текстів виникає занадто багато смислових нюансів і різноманітних способів висловлювання, окремі з яких комп'ютером будуть сприйматися як неправильна відповідь.

Побудувати запитання, відповіддю на яке є лише одне слово або конкретне словосполучення, можливо далеко не з кожної дисципліни. Пріоритет тут мають точні науки, а завданнями можуть слугувати означення.

Другою складністю для викладача є визначення ключового слова, яке треба вилучити із запитання. Очевидно, що це слово повинно мати найбільшу інформаційну цінність у семантичних фактах. Для прикладу розглянемо висловлення

Прямокутник, у якого всі сторони рівні, називається квадратом.

З точки зору глибини розуміння матеріалу таким словом у даному випадку є слово *прямокутник*. Саме знання студентами цього слова дозволяє викладачу оцінити уміння встановлювати родові відношення між поняттями. Вилучення слова *сторони* дозволяє контролювати уміння встановлювати видові відмінності понять, але в даному випадку завдання перетворюється в альтернативне, бо передбачає ще лише одну можливу відповідь *кути*, що звужує вибір правильної відповіді. Вилучення ж слова *квадратом* дасть змогу перевірити лише знання термінів, що, в принципі, теж є необхідним, але доцільність таких тестів виправдана хіба що при вивченні тем з великим обсягом нової термінології (наприклад у біології, медицині тощо).

На сьогодні вже розроблені системи комп'ютерного тестування, які методом «ключових слів» розпізнають правильні відповіді. Проте недоліком цих систем є те, що вживання студентом слів у іншому відмінку, синонімів, допущення незначних граматичних чи стилістичних помилок вважається неправильною відповіддю [2,98].

Ситуація кардинально змінюється у випадку використання для ідентифікації текстової відповіді системи Control. В основу цієї системи покладені математичні методи обчислення чисельних мір подібності символічних послідовностей [4;5]. Така система в процесі розпізнавання правильної відповіді ігнорує несуттєві граматичні та синтаксичні помилки і не звертає уваги на стилістику відповіді доти, доки зберігається розпізнаваність її смислу [5]. За допомогою виділення у відповіді окремих смислових блоків система дозволяє оцінювати різноманітними сингулярними правилами семантично різні частини відповіді [2].

Безумовно, такі можливості системи комп'ютерного тестування Control розв'язують цілу низку питань з проблеми оцінювання знань, але вимагають від упорядника тестів глибоких знань не тільки з предмету, а й з програмування і з особливостей семантичної будови речень, або над створенням таких тестів має працювати група фахівців.

Затрати часу, витраченого на створення таких тестів, на нашу думку, не завжди можуть компенсуватися дидактичною ефективністю їхнього застосування. Крім того, намагання студента дати повну відповідь призводить до вживання “зайвих” слів, не врахованих системою, а коротка відповідь може бути оцінена як неповна, бо не всі ключові слова, що наявні в базі системи, використані.

Тому, ми вважаємо, що у випадку, коли відповідь передбачає об'ємний текст, доцільніше користуватися традиційними методами письмового, а ще краще усного опитування студентів. Адже під час діалогу із викладачем студент має змогу уточнити, виправити, доповнити свою відповідь, висловити власну думку.

Тестові завдання на відповідність базуються на теорії графів. Вони легко можуть бути виконані в комп'ютерному варіанті і застосовані для будь-якого контингенту учнів чи студентів. На екрані пропонується дві множини понять і ставиться вимога встановити між ними відповідності. Наприклад:

Встановити відповідність між письменниками та назвами їх творів:

1. Шарль Перро *		* а) Мауглі
2. Джозеф Ред'ярд Кіплінг *		* б) Попелюшка
3. Уолт Дісней *		* в) Дикі лебеді
4. Астрід Ліндгрєн *		* г) Кіт у чоботях

5. Ганс Крістіан Андерсен *		* д) Білосніжка та семеро гнонів
6. Жюль Верн *		* е) Пеппі Довгапанчоха

Часто вимагається відповідь вводити за допомогою пари (число, буква). У нас, наприклад, (2,а), (1,б) і т.ін. Якщо завдання даються на індивідуальних картках, то відповідності зручніше зображати стрілками (див.таблицю). У комп'ютерному варіанті можливі обидва способи введення відповіді, але кожний з них має свої незручності. При введенні пари студент має чітко знати, що число і буква набираються у строгому порядку, без пропусків, через ризик, через кому і т.п., що вимагає значної уважності.

Для школярів більш наочним і цікавим є використання стрілок, але це вимагає достатнього володіння комп'ютерною мишкою. Проте, такий спосіб введення відповідей зустрічається в багатьох комп'ютерних дидактичних іграх і не викликає у дітей значних труднощів.

Тестувальні завдання на встановлення відповідностей можуть бути застосовані для контролю знань з будь-якої дисципліни, але в порівнянні з іншими типами тестів часто мають низький рівень складності. Тому вони більш доцільні для застосування в початкових та середніх класах загальноосвітніх шкіл.

Великий інтерес і бажання працювати, як показує досвід, викликають у молодших школярів тести, що містять малюнки. Наприклад:

Покажи, з якого дерева зірвано плід.

За умови правильної відповіді вказаний плід і відповідне дерево зникають з екрана, або плід "приростає" до дерева і т. ін. На жаль, створення тестів з малюнками вимагає додаткових зусиль, але їхній естетичний вигляд і емоційний вплив того варті. Багато таких тестових завдань у казковій формі використано в комп'ютерних дидактичних іграх (див.рис.1).

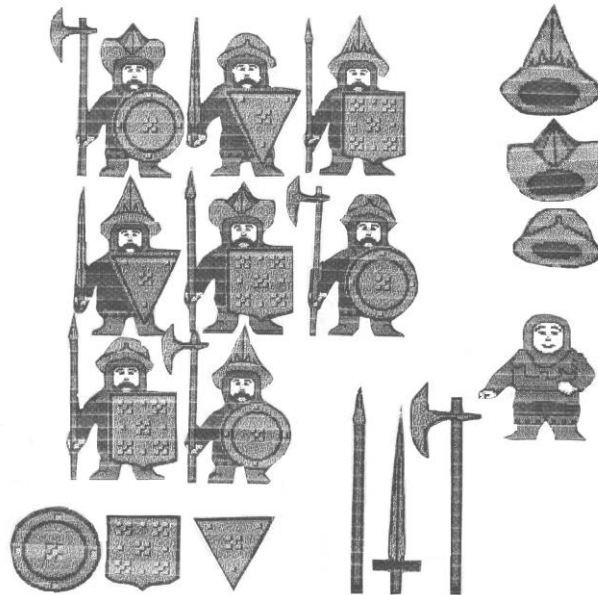


Рис.1. Варіант комп'ютерного тесту на встановлення відповідностей між множинами.

Із запропонованих щитів, шоломів і видів зброї, враховуючи підмічену закономірність, дитині потрібно вибрати все необхідне для лицаря, щоб він міг "стати в ряд". Тут відповідь вводиться наведенням курсору і натисканням клавіші мишки.

Тестові завдання на встановлення правильної послідовності передбачають, що студент хаотично розміщені об'єкти чи явища розмістить (чи пронумерує) в певному порядку (за хронологією, важливістю, зростанням чи спаданням тощо). Слова в завданні такого тесту здебільшого нумерують, а тому відповідь матиме вигляд числової послідовності, яка дуже добре розпізнається комп'ютером. Інколи відповідь подають у вигляді пар чисел, з яких першим є номер об'єкта, а другим його порядковий номер у правильній послідовності. Наприклад:

Розмістити величини в порядку зростання

1) 1 см; 2) 2 мм; 3) 5 м; 4) 7 дм; 5) 3 км.

Форми відповідей: 21435 або 1-2, 2-1, 3-4, 4-3, 5-5. Очевидно, що в другому випадку є ризик випадково переплутати числа, що створює певні незручності. У випадку ж, коли об'єкти запитання не пронумеровані, відповідь може передбачати їхнє пронумерування. Якщо такий тест подається в комп'ютерному варіанті, то відповідь може передбачати також звичайне натискання кнопок клавіатури або мишки після наведення курсора на об'єкти питання в правильній послідовності.

З усього сказаного випливає, що застосування комп'ютерної діагностики знань студентів за умови дотримання принципів достовірності й об'єктивності є ефективним засобом оцінювання. Крім того, у вищих педагогічних навчальних закладах її застосування має ще й суттєве методичне значення. Майбутні вчителі таким чином практично знайомляться з різноманітністю тестів і вивчають у процесі навчання їхні особливості. Це дасть змогу студентам у майбутньому самостійно створювати тести і впроваджувати їх у навчальний процес. На завершення варто зазначити, що

використовувати комп'ютерні системи тестування варто лише тоді, коли це доцільно і дидактично виправдано.

ЛІТЕРАТУРА

1. Атанов Г.А., Пустынникова И.Н. Обучение и искусственный интеллект, или Основы современной дидактики высшей школы. – Донецк: Изд-во ДОУ, 2002. – С.211-247.
2. Баранов В.Ю. Методи оцінювання семантично різномірної відповіді у системі комп'ютерного тестування знань //Комп'ютерно-орієнтовані системи навчання: Зб. наук. праць.– К.: НПУ ім.М.П.Драгоманова.– Випуск 5.– 2002.– С.97-110.
3. Булах І.Є. Тестовий контроль та проблеми його стандартизації // Вища освіта в Україні: реалії, тенденції, перспективи розвитку. – К.: ІСДО, 1996.– Ч.4.– С.112-114.
4. Леоненко Л.Л., Поддубный Г.В. Новые принципы идентификации ответов и их реализация в автоматизированной системе тестирования знаний Control // Материалы науч.-метод.конф. «Программно-технические средства информатизации образования».– К., 1995.– С.91-92.
5. Леоненко Л.Л., Поддубный Г.В. Теория подобия конечных последовательностей и ее приложение к распознаванию образов // Автоматика и телемеханика.– М., 1996.– №8.– С.119-131.
6. Сілкова О.В. Застосування автоматизованих атестаційних систем та класифікація тестових завдань при використанні комп'ютерних технологій // Комп'ютерно-орієнтовані системи навчання: Зб.наук.праць.– К.: НПУ ім. М. П. Драгоманова. – Випуск 3.– 2001.– С.173-178.