

ФОРМИРОВАНИЕ ТЕСТОВ ЗАДАННОЙ СЛОЖНОСТИ

к.т.н. В.Е. Климнюк, к.т.н. В.А. Кирвас, к.т.н. С.И. Козыренко
(представил д.т.н., проф. Е.И. Бобыр)

Предложен метод автоматического формирования тестов с заданным уровнем сложности.

Постановка проблемы. Педагогический контроль является неотъемлемой частью процесса образования и профессиональной подготовки специалистов. Мировая история подготовки высококачественных специалистов свидетельствует о том, что наиболее корректным средством измерения характеристик личности субъекта обучения являются тесты – психодиагностические методики измерения и оценивания достигнутого уровня развития способностей, умений и знаний.

Среди средств объективного контроля, наиболее научно обоснованным является метод тестирования с привлечением технических средств. Использование автоматизированных тестовых методик в системе образования позволяет сократить финансовые затраты и время при одновременном повышении качества и объективности процедуры педагогического контроля.

Создание системы объективной тестовой оценки знаний и умений, которую можно было бы использовать при всех видах педагогического контроля и, которая удовлетворяла бы его целям и функциям, должно базироваться на современных технологиях педагогических измерений. Это требует от разработчиков не только высокой квалификации в педагогике и в предметной области, не только понимание целей подготовки специалистов определенного профиля, но и глубоких знаний из теории тестирования. Можно утверждать, что исследование вопросов теории построения подсистемы формирования педагогических тестов является одной из важнейших задач современной высшей школы.

Анализ литературы. В педагогической и психологической литературе широко обсуждаются вопросы оценки знаний методом объективного автоматического тестирования [1 – 3]. Его преимущества и особенности наиболее полно раскрываются при диагностике уровня подготовки обучающихся по дистанционной форме [4 – 6]. Одной из обязательных подсистем в системе дистанционного обучения является подсистема контроля знаний. Данная подсистема диагностики должна обеспечивать

возможность как самотестирования обучающегося, так и промежуточный и итоговый контроль со стороны преподавателя [1, 3]. При этом научность педагогического контроля предполагает соответствие тестов определенным критериям оценки качества методов измерения, важнейшими из них являются: объективность, надежность, валидность и точность [6 – 8]. Для обеспечения функциональной валидности теста, при контроле определенного уровня знаний, тестовые вопросы должны отвечать заданному уровню сложности.

Целью данной статьи является разработка метода формирования тестов с заданным уровнем сложности.

Анализ требований к подсистеме контроля знаний. При ее разработке необходимо учесть тот факт, что преподаватели, создающие тесты по своим дисциплинам, могут устанавливать разные системы, определяющие уровень сложности каждого вопроса. Например, можно установить 5-и, 10-и или N-бальную систему, устанавливать относительный уровень сложности от 0 до 1, выражать сложность в виде процентов и т.д. При этом подсистема должна иметь удобный интерфейс управления системой установки сложности тестов.

Уровень тестовых задач должен соответствовать уровню подготовки студента. Тестовые задания должны иметь разные уровни сложности, определяемые преподавателем или программой в зависимости от ответов студента, т.е. должна обеспечиваться адаптивная сложность теста.

Для поддержки мотивации большое значение имеет произвольный порядок следования вопросов по уровням сложности и преподаватель должен иметь возможность их корректировать. Большое значение для мотивации также имеет время ожидания студентом очередного вопроса или теста в целом. Это время должно быть как можно меньшим.

В подсистеме контроля знаний необходимо предусмотреть возможность управления уровнем сложности контролируемых тестов от самого низкого до самого высокого уровня. Управление уровнем сложности может осуществляться как для группы в целом (равные условия для всех) – например, при итоговом тестировании, так и для отдельных студентов в соответствии с их уровнем подготовки – на этапе обучения. На этом этапе целесообразно дать возможность самому обучаемому устанавливать тот или иной уровень сложности. При успешном освоении учебного материала уровень сложности должен возрастать. Только в этом случае поддерживается необходимая степень мотивации обучения.

Немаловажную роль играет необходимость исключения возможных «консультаций» со стороны соседей при проведении итогового тестирования. Это может быть достигнуто случайным выбором порядка следования

одних и тех же вопросов для разных студентов. Более эффективно эта задача может быть решена, если при равных суммарных уровнях сложности тестов тесты для разных студентов будут иметь разный набор вопросов, которые в свою очередь будут вызываться также по случайному закону.

Таким образом, к подсистеме контроля знаний можно предъявить следующие общие требования:

- простота использования приложения (удобный интерфейс);
- хранение, обработка и выдача необходимой информации (персональные данные, перечень вопросов по различным темам, результаты тестирования);
- возможность внесения изменений в список вопросов по темам и самим тем для тестирования;
- обеспечение проведения объективного автоматизированного контроля, в том числе и дистанционного;
- обеспечение проведения текущего и итогового контроля;
- обеспечение гибкости в подборе количества уровней сложности тестов;
- возможность формирования тестов с заданной степенью сложности;
- приемлемое время формирования и выдачи контролируемых тестов.

Данные требования должны быть учтены при разработке подсистемы контроля знаний. Ниже рассматривается метод формирования тестов с заданной степенью сложности.

Разработка метода формирования тестов заданной сложности. Пусть в базе данных находится N вопросов по определенной теме. Сложность каждого вопроса – Q_i . Из этих вопросов необходимо сформировать тест, представляющий собой набор из M вопросов, относящихся к данной теме. Необходимо обеспечить суммарную сложность каждого теста на уровне $Q_{\text{зад}}$, заданном преподавателем при формировании вопросов по теме. Предусмотреть также возможность оперативного изменения преподавателем, при необходимости, значения M и $Q_{\text{зад}}$ для определенной темы. Одним из методов решения данной задачи может быть метод прямого перебора всех допустимых вариантов. Его суть заключается в формировании множества всех возможных вариантов тестов, вычислении суммарной сложности каждого из них и выборе случайным образом одного из тестов, удовлетворяющих заданной сложности. Однако этот метод неприемлем, так как при его практической реализации возникают трудности, связанные с хранением и обработкой большого объема информации и недопустимо большим временем формирования тестов. Например, при $N = 50$ и $M = 10$ $C_{50}^{10} = 10272278170$, иными словами необходимо выбирать тесты из более чем 10 млрд. вариантов, что неприемлемо ни с временных, ни с материальных соображений.

Предлагается использовать другой метод решения данной задачи, представляющий собой метод управляемого выбора очередного вопроса при формировании теста.

Для формирования неповторяющихся вопросов в тестах выбор очередного вопроса осуществляется случайным образом из множества всех вопросов, составляющих изучаемую тему. При этом используются датчики случайных чисел, распределенных по равномерному закону. В обычных методах все вопросы выбираются с одинаковой частотой. В предлагаемой методике вначале все вопросы изучаемой темы распределяются по k группам в соответствии со своими уровнями сложности, где k – количество уровней сложности вопросов в данной теме. На очередном шаге формирования теста выбор вопроса для включения его в тест осуществляется в два этапа. Вначале определяется группа сложности, из которой нужно выбрать очередной вопрос, а затем из этой группы с помощью датчика случайных чисел осуществляется выбор вопроса.

Суть предлагаемого метода заключается в управлении выбором очередной группы вопросов. Управление осуществляется за счет коррекции значений вероятности обращения к той или иной группе.

Коррекция состоит в увеличении вероятности P_n выбора более «легкой» группы вопросов, если при очередном выборе вопроса текущее значение сложности теста превышает заданное значение и наоборот.

Так как события выбора i -й группы ($i = 1, \dots, k$) составляют полную группу событий, то по сути метод заключается в изменении размеров диапазонов для каждой группы сложности на интервале $[0, 1]$ (рис. 1). Если необходимо увеличить долю сложных вопросов в тесте, то соответствующий диапазон увеличивается и наоборот.

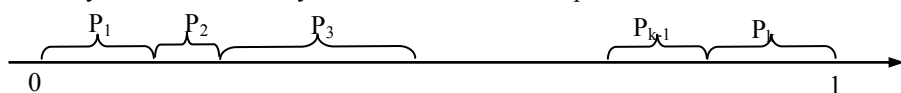


Рис. 1. Распределение откорректированных величин P_i

Рассмотрим более детально предлагаемую методику формирования тестов. Пусть имеется N вопросов. Сложность i -го вопроса – Q_i . Из этих вопросов необходимо выбрать M вопросов так, чтобы обеспечить заданную суммарную сложность набора на уровне $Q_{\text{зад}}$ ($Q_{\text{min}} \leq Q_{\text{зад}} \leq Q_{\text{max}}$). Алгоритм формирования тестов с заданным уровнем сложности приведен на рис. 2.

1. Формирование k групп вопросов в зависимости от их сложности;
2. Определение плотности распределения вероятностей каждой группы вопросов в общем множестве всех вопросов

$$P_i = \gamma_i / N,$$

где γ_i – количество вопросов в i -й группе, причем:

$$\sum_{i=1}^k r_i = N; \quad \sum_{i=1}^k P_i = 1. \quad (1)$$

3. Расчет заданной средней сложности вопросов в тесте

$$\bar{Q} = Q_{\text{зад}}/M.$$

4. Определение средней сложности вопросов по данной теме

$$m_0 = \sum_{i=1}^k q_i \cdot P_i,$$

где q_i – сложность i -й группы.

5. Коррекция P_i в зависимости от соотношения \bar{Q} и m_0 . Если $\bar{Q} > m_0$, то необходимо увеличить плотность распределения вероятностей для групп вопросов с большей сложностью и уменьшить для групп вопросов с меньшей сложностью, при этом:

$$\sum_{i=1}^k P'_i = 1; \quad |m'_0 - \bar{Q}| \leq E_0, \quad (2)$$

где P'_i – откорректированное значение плотности распределения вероятностей i -й группы; m'_0 – новое значение средней сложности вопросов после коррекции P_i ; E_0 – заданная точность формирования теста с заданной сложностью

$$P'_i = \begin{cases} x_0 \cdot P_i & \forall i: q_i < \bar{Q}; \\ P_i + x_i \cdot P_i & \forall i: q_i > \bar{Q}. \end{cases} \quad (3)$$

Если $\bar{Q} < m_0$, то необходимо уменьшить плотность распределения групп вопросов с большей сложностью и увеличить – для групп вопросов с меньшей сложностью. При этом должны выполняться соотношения (2):

$$P'_i = \begin{cases} x_0 \cdot P_i & \forall i: q_i > \bar{Q}; \\ P_i + x_i \cdot P_i & \forall i: q_i < \bar{Q}. \end{cases} \quad (4)$$

Если $\bar{Q} = m_0$, то коррекция не требуется, т.е. $P'_i = P_i$.

6. Для поиска неизвестных величин (x_0 и x_i) в выражениях (3) и (4) целесообразно применить один из численных методов, например метод деления пополам

$$\Delta P_i = \sum (P'_i - P_i) \quad \forall i: q_i < \bar{Q}; \quad x_i = (P_i / \sum P_i) \cdot \Delta P_i \quad \forall i: q_i > \bar{Q}.$$

7. Затем используется датчик случайных чисел от $[0; 1]$ и определяется группа вопросов, из которой будет выбран очередной вопрос.

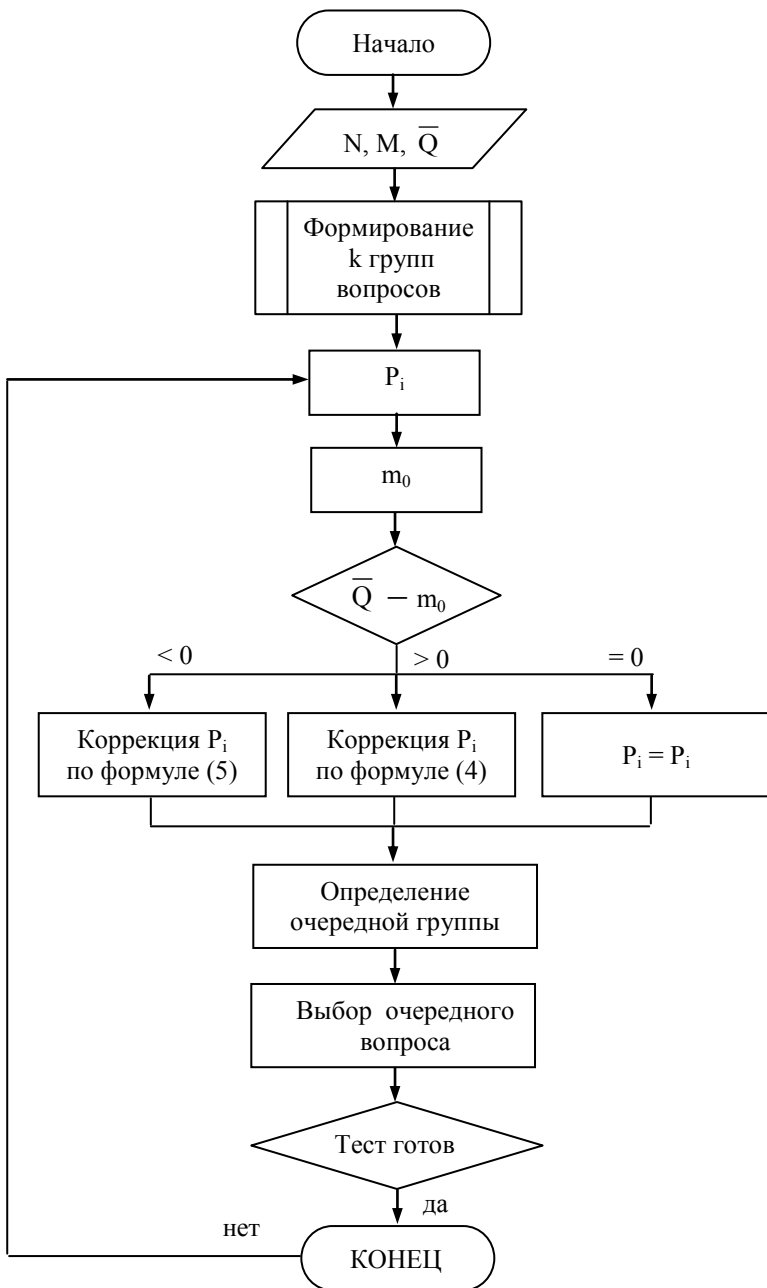


Рис. 2. Блок-схема алгоритма формирования тестов с заданным уровнем сложности

8. В выбранной группе вопросов j с помощью датчика случайных целых чисел в диапазоне $[1; r_j]$ выбирается номер очередного вопроса и помещается в тест. Если после этого в тесте оказалось M вопросов, то КОНЕЦ, иначе повторяется шаг 2.

Выводы. Анализ показал, что одним из основных требований к подсистеме контроля знаний системы дистанционного обучения является возможность формирования тестов с заданной степенью сложности. Разработанный метод позволяет управлять выбором очередной группы вопросов за счет коррекции значений вероятности обращения к той или иной группе. Приведенный алгоритм обеспечивает формирование тестов заданной сложности. Полученные результаты могут быть использованы при разработке подсистемы автоматизированной тестовой диагностики уровня знаний обучающихся.

ЛИТЕРАТУРА

1. Аузіна М.О., Голуб Г.Г., Возна А.М. Система комплексної діагностики знань студентів: Навчальний посібник. – Львів: ЛБІ НБУ, 2002. – 38 с.
2. Вроейстийн А.И. Оценка качества высшего образования. Рекомендации по внешней оценке качества в ВВУЗах. – М.: МНЭПУ, 2000. – 180 с.
3. Селевко Г. Современные образовательные технологии. – М.: Наука, 1998. – 256 с.
4. Основы дистанционного обучения. Дистанционный курс / Под ред. проф. В.Н. Кухаренко – Х.: ХГПУ, 1999. – 182 с.
5. Кирвас В.А, Кирвас В.В. Задача постепенного перехода от заочного к дистанционному обучению. Экспертные оценки элементов учебного процесса // Материалы V МНТК, Харьков, 24 октября 2003 г. – ХГУ “НУА”. – 2003. – С. 33 – 36.
6. Кирвас В.А., Кирвас В.В. Классификация учебно-методических материалов для дистанционного обучения // Проблемы информатики и моделирования. Материалы второй МНТК. – Х.: НТУ «ХПИ». – 2002. – С. 11 – 12.
7. Беспалько В.П. Системы педагогической технологии. – М.: Педагогика, 1989. – 302 с.
8. Рудинский И.Д., Клеанорова А.А. Концепция количественного оценивания объективности педагогического тестирования знаний // Информатика и образование. – № 12. – 2003. – С. 100 – 104.

Поступила 21.04.2004

КЛИМНИЮК Виктор Евгеньевич, к.т.н., доцент, доцент кафедры вычислительной техники и программирования ХГЭУ. В 1970 г. окончил ВИРТА ПВО. Область научных интересов – информационные технологии в образовании.

КИРВАС Виктор Андреевич, к.т.н., доцент, профессор кафедры информационных технологий ХГУ “НУА”. В 1976 г. окончил ХВВКУ. Область научных интересов – информационные технологии в образовании.

КОЗЫРЕНКО Светлана Ивановна, к.т.н., доцент, доцент кафедры прикладной математики ХНУРЭ. В 1974 г. окончила ХИРЭ. Область научных интересов – информационные технологии в образовании.