

Ю. Ю. ЧЕРЕПАНОВА, М. С. ШИРОКОПЕТЛЕВА

Харьковский национальный университет радиоэлектроники

МЕТОДИКА ФОРМИРОВАНИЯ ТЕСТОВЫХ ЗАДАНИЙ И ОЦЕНКИ РЕЗУЛЬТАТОВ ДЛЯ ПРОГРАММНОЙ СИСТЕМЫ ТЕСТИРОВАНИЯ ЗНАНИЙ ЯЗЫКА SQL

Компьютерный контроль знаний студентов сейчас является основой для получения объективной оценки уровня их знаний. Большой интерес представляет собой проблема тестирования знаний с применением тестов в открытой форме, особенно в связи с возрастанием роли дистанционного и самостоятельного обучения. В большинстве технических учебных заведений сферы IT изучают основы баз данных, в том числе язык, без этого в наше время невозможно разрабатывать большинство программных продуктов. Система тестирования знаний языка запросов к базе данных может использоваться в учебных заведениях для упрощения задач проверки и обучения студентов. Целью исследования является разработка методики формирования тестовых заданий и оценки результатов для программной системы тестирования знаний языка SQL. Речь идет о тестировании с помощью тестов в открытой форме, а именно способ, при котором студенту предлагается решить задачу - написать текст запроса к базе данных. Система должна оценить правильность запроса.

В существующих исследованиях рассмотрены вопросы формирования тестового набора, но не учтены особенности проверки знаний по языкам программирования, в частности, языка запросов.

Для методики формирования тестовых заданий, кроме стандартных критериев для тестов (критерий надежности, валидности, сложности теста, дискриминативности) [1], предложим критерии, которым должны удовлетворять тестовые наборы:

1. Критерий полноты - вопросы в тестовом наборе должны покрывать все элементы темы, по которой проводится тестирование;
2. Критерий отсутствия пересечения тестовых вопросов: проверке подлежат только фрагменты задания, по которому сформулирован вопрос.

Подробнее опишем критерий полноты. Для этого приведем термины «содержание понятия» и «объем понятия»: содержанием понятия является совокупность существенных признаков, характеризующих данный предмет, а объем понятия - совокупность или множество предметов, которые входят в состав понятия. Проведя аналогию между понятием (основным объектом предметной области) и дисциплиной, по которой проводится тестирование, можно заметить, что в содержание дисциплины входит весь перечень тематических единиц (тем, разделов, подразделов и т.д.), а объем составляет множество тестовых вопросов, которые покрывают весь смысл каждой темы. Формализуем:

$$S = \cup R^i,$$

где S – объем всей дисциплины; R^i – i -я тематическая единица, $i \in [1; n]$, n – количество тематических единиц в дисциплине.

Например, для проверки знаний по языку структурированных запросов, выделим разделы: DML, DDL, DCL. В каждом из них определим тематические единицы, взаимосвязь которых для раздела DML показана в табл.1. Приведенный уровень сложности может применяться для автоматизации формирования тестовых наборов для устойчивой группы пользователей или дискриминационного тестирования (для получения максимальной оценки «удовлетворительно» за успешное прохождение теста первого уровня сложности).

Таблица 1 – Тематические единицы подраздела DML

| Подраздел | Тема | Тематическая единица | Уровень |
|-----------|--------------------------------|---|---------|
| SELECT | Простые запросы | Формирование вывода | 1 |
| | | Ограничения на данные | 1 |
| | | Специальные операторы, упорядочение | 1 |
| | | Групповые операции | 1 |
| | | Ограничения на группы | 2 |
| | | Оператор IS NULL | 2 |
| | Запросы на нескольких таблицах | Объединение таблиц | 2 |
| | | Объединение таблицы с собой | 2 |
| | | Ограничения на данные и группы из нескольких таблиц | 3 |
| | Подзапросы | Простые подзапросы | 2 |
| | | Коррелированные подзапросы | 3 |
| | | Предикаты ANY, ALL, EXISTS | 3 |

| | | |
|--------|--|---|
| INSERT | Добавление единичной записи | 1 |
| | Множественное добавление | 2 |
| UPDATE | Простая модификация данных | 1 |
| | Модификация данных с использованием данных из другой таблицы | 2 |
| DELETE | Простое извлечение данных | 1 |
| | Извлечение данных с использованием данных из другой таблицы | 2 |

Для предотвращения «обмена мнениями» во время проведения тестирования тест должен содержать не менее 10 вопросов по каждой теме. То есть для теста из этого подраздела необходимо составить не менее 180 вопросов.

На основании критерия отсутствия пересечения тестовых вопросов, предлагается ограничить проверку ответов только тематикой вопроса. Например, при проверке знаний по тематическому разделу «Ограничения на данные» необходимо проверить результирующую выборку на количество строк и на их соответствие, и не учитывать количество и порядок полей вывода и упорядочение строк в результате запроса. Формализуем:

$$\forall i, j: R^i \cap R^j = \emptyset,$$

где R^i, R^j – i (j) -а тематическая единица, $i, j \in [1; n]$, n – количество тематических единиц в дисциплине.

Таким образом, все тестовое множество может быть представлено:

$$S = \cup Q_l^i,$$

где Q_l^i – l -й вопрос тестового набора, который входит в состав $R^i, l \geq 10$.

Так, для формирования тестового набора из одного тематического раздела, следует привести однотипные вопросы, например, для тематического раздела «Ограничения на данные» приведем таблицу Students (id, Name, DOB, Group, Address) и запросы:

- Выведите студентов, обучающихся в группе ИТ-21;
- Выведите студентов, проживающих в Харькове и т.д.

Для определения корректного ответа следует для каждого запроса сформировать эталонный запрос, с помощью которого можно получить результирующую выборку. Также необходимо разработать показатели оценки

результатирующих запросов для каждой из тематических единиц и ввести их в базу тестовых вопросов с указанием тематической единицы, а все задачи должны относиться к определенной единице для осуществления проверки.

Тестовый набор для итогового контроля знаний должен содержать вопросы по каждой тематической единице, но при достаточно большом количестве тем и ограничении времени предлагается проводить тестирование, предназначенное для отдельных групп пользователей, а в тестовые наборы включать вопрос одного уровня сложности. Например, в первую очередь можно сформировать тест по уровню сложности 2 (оценка «хорошо»), после сдачи которого можно сформировать 2 набора: с уровнем сложности 1 (для проверки владения базовыми знаниями для тех, кто не сдал первую часть) и уровнем 3 - для пользователей, которые могут претендовать на отличную оценку.

Также можно проводить текущий контроль путем автоматического формирования тестовых наборов на основании вопросов с выбранного раздела / подраздела без учета сложности задач. Количество вопросов в задании преподаватель может самостоятельно задавать, но это количество не должно быть меньше, чем количество тематических единиц. То есть при иерархической структуре организации задач облегчается процесс формирования тестовых наборов с учетом критерия валидности по содержанию.

Для проверки знаний могут использоваться различные критерии формирования оценки [1]. Методы на основе количественных критериев предполагают использование количественной шкалы, то есть оценка в этом случае задается числом. К примеру, коэффициент усвоения $K = A / P$, где A - количество правильных ответов в тесте, P - общее количество ответов. Также при формировании оценки можно учитывать типы и характеристики тестовых заданий, уровень подготовленности тестируемого, принадлежность пользователя к стойким классам и т.д. Перечисленные подходы могут применяться для тестов, которые не предусматривают несколько правильных семантически различных вариантов ответов. Для проверки знаний и навыков в области программирования следует предусмотреть возможность автоматизированной проверки

предложенного решения, то есть создание системы тестирования, которая позволит выполнить фрагмент кода программы и проверить результаты по темам.

Для программной системы проверки знаний по языку SQL необходимо создать набор требований по каждой тематической единице и алгоритм сравнения результирующих выборок (эталонной и предложенной пользователем).

Например, для темы «Ограничение данных» результаты полученных выборок предлагается сравнивать по строкам (для данной тематической единицы) по следующим показателям: количество строк должна совпадать, идентификаторы строк (ROWID) должны совпадать при выборки из базы данных, частично неизвестных пользователю (для предотвращения некорректного формирования запроса). При выполнении этих правил задачу можно считать выполненной правильно.

Для проверки запросов на модификацию следует предусмотреть выполнение двух изолированных транзакций (для эталонного и предложенного запросов) для предотвращения произвольного изменения исходных данных, а результаты выполнения задания проверять с помощью сверкой данных таблиц по такому критерию: полное совпадение данных таблицы при выполнении запроса на выборку всех данных с составлением по ключевому полю.

При проведении текущего контроля и формировании задач с равным уровнем сложности для оценки можно применять коэффициенты, соответствующие уровням сложности. Предлагается общую оценку рассчитывать по формуле: $M = k * A / P$, где k - коэффициент сложности. Для среднего уровня сложности (2) – $k = 1$, для низкого (1) - $k = 0.7$, для заданий повышенной сложности (3) - $k = 1.3$. При большем количестве уровней коэффициент сложности задавать самостоятельно.

Литература:

1. *Карпенко, А. П.* Тестовый контроль качества обучения и критерии качества образовательных тестов. Обзор [Электронный ресурс] / *А. П. Карпенко, А. С. Домников, В. В. Белоус* //Наука и образование: электронное научно-техническое издание. – 2011.– № 04. – Режим доступа: <http://technomag.edu.ru/doc/184741.html>. – 15.06.2016г. – Загол. с экрана.