

Федюков А.Ю.,

Магистрант

2 курс, факультет «Космический»

МФ МГТУ им. Н. Э. Баумана

Россия, г. Москва

Научный руководитель: Афанасьева В.В.,

кандидат технических наук

ОПРЕДЕЛЕНИЕ СХОЖЕСТИ ТЕКСТОВ, НАПИСАННЫХ НА ЕСТЕСТВЕННО-НАУЧНОМ ЯЗЫКЕ

***Аннотация:** Результатом данной работы является разработка математической модели обработки входящих потоков строк в алгоритм *MgtyMf*, позволяющий реализовывать поставленные цели.*

***Ключевые слова:** Проблемы прикладной математики, построение математических моделей, математическая модель обработки входящих потоков строк, уравнение взвешенного расстояния Левенштейна.*

***Annotation:** The result of this work is the development of a mathematical model for processing incoming streams of strings in the *MgtyMf* algorithm, which allows implementing the set goals.*

***Keywords:** Problems of applied mathematics, construction of mathematical models, mathematical model of processing incoming streams of strings, Levenshtein weighted distance equation.*

При разработке интеллектуальной информационной системы использовались несколько сред компилирования и языков программирования. Это применяется, когда разрабатывается кроссплатформенное приложение, и для гибкости написания, когда можно применять PHP, HTML, CSS и

JavaScript, которые в свою очередь имеют возможность взаимодействовать друг с другом.

В стационарной версии для персональных компьютеров интерфейсной части программы выбран язык программирования Borland C++ Builder 6. Основными аргументами в пользу выбора данной среды разработки были:

- высокая производительность создаваемых модулей ПО промежуточного уровня для web-служб.
- обеспеченность полной поддержкой стандартных протоколов SOAP, XML, WSDL и XSL.
- возможность разработки кроссплатформенных приложений для Windows и Linux.
- поддержка баз данных IBM DB2, IBM Informix, Oracle, Sybase, MySQL, dBASE, Paradox и Borland InterBase. Драйверы баз данных dbGo for ADO, dbExpress и BDE, входящие в состав C++Builder, обеспечивают высокопроизводительную работу приложений с такими СУБД как DB2, Informix, Oracle, Sybase, Microsoft SQL Server, MySQL, Access, Paradox и InterBase. SQL Monitor и другие отладочные инструменты позволяют достичь высокой производительности, масштабируемости и уменьшения времени отклика создаваемых приложений баз данных.

Для хранения базы данных вопросов и результатов тестирования была выбрана СУБД MySQL. База данных MySQL используется как для стационарной, так и для сетевой версии программы. MySQL относится к продуктам класса Open Source (открытые исходные тексты), которые можно получить бесплатно (с некоторыми ограничениями для платформы Windows). Как правило, сервер и клиент MySQL входят в любой дистрибутив операционных систем семейства BSD (FreeBSD, NetBSD, OpenBSD) и Linux, которые используются на большинстве веб-серверов, но при необходимости последнюю версию MySQL для всех поддерживаемых систем (в том числе и для Windows) также можно найти на сайте компании MySQL AB –

разработчиков MySQL (<http://www.mysql.com>). Второй немаловажной причиной, является надежность и защищенность. Например, если в других СУБД опираются на концепцию транзакций, которая обеспечивает целостность и согласованность таблиц, фиксируя только законченные изменения, легко откатывая базу данных к последнему согласованному состоянию в случае неудачи, то в MySQL применяется концепция «запираания» таблицы перед проведением операций, потенциально опасных для целостности базы данных, не позволяя в этот момент другим соединениям изменять и читать данные с этой таблицы и ставя запросы в очередь. Существует множество различных СУБД. На данный момент, на рынке малых, средних и промышленных систем, лидирует корпорация – Oracle. Ей принадлежит две популярных СУБД: ORACLE и MySQL. Сравнить две такие разные СУБД довольно тяжело, так как они принадлежат к разным областям применения. По мнению автора, MySQL больше подходит для проектов, в которых будут обрабатываться малое количество информации и не будет, относительно, большого количества пользователей. ORACLE же больше подходит для больших проектов.

Интеллектуальная система контроля знаний «MgtyMf» (рис. 1) предназначена для автоматизированного контроля знаний учащихся по дисциплинам ответы, на которые необходимо давать на естественнонаучном языке. Применение новых методов и алгоритмов позволяет ей, в большей степени, заменить живое общение тестируемого и преподавателя. Разработанные методы и алгоритм обработки ответов, позволяют проводить анализ текста и объективно оценивать знания тестируемого.

Разработанная структура системы тестирования представляет собой пополняемую (изменяемую) систему динамически подключаемых программных модулей:

- 1) функциональные модули (реализуют отдельные подсистемы и функции, к примеру, работу модуля, отвечающего за закрытые типы вопросов, модуля дополнительных вопросов или модуля вопросов открытого типа);
- 2) модуль оценки (отвечает за итоговый результат тестирования, выводит результат в виде процентного соотношения правильных и неправильных ответов);
- 3) модуль пользовательского интерфейса (обеспечивают интерактивное взаимодействие пользователя с системой тестирования);
- 4) база данных (организация хранения результатов тестирования и служебной др. информации).
- 5) справка (отвечает за информационное обеспечение пользователя в ходе работы с программой);

Взаимодействие пользователя с системой тестирования MgtyMf проходит через графический интерфейс (рис. 21). В ходе проведения тестирования, полученные результаты сохраняются в базе данных.

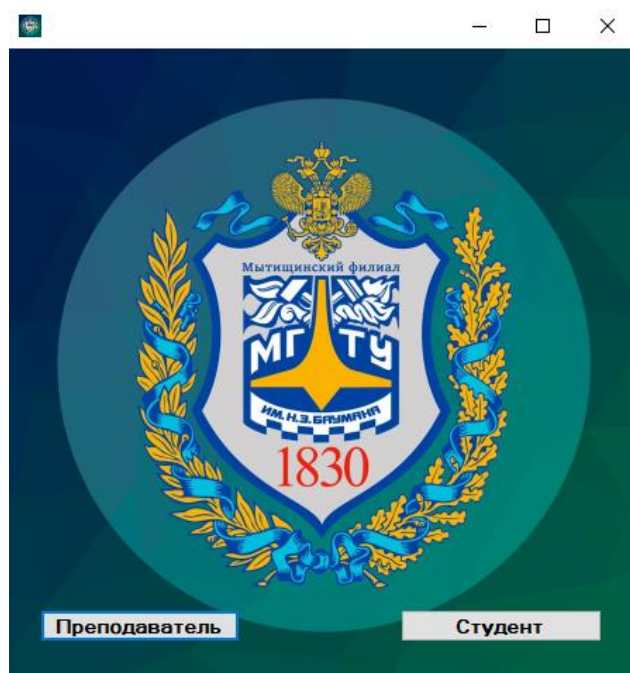


Рисунок 1. Интерфейс программы

Разработанная программа состоит из двух модулей: модуль «студент» и модуль «преподаватель». Модуль «студент» позволяет проходить

тестирование, а также посмотреть свои результаты. Модуль «преподаватель» позволяет править базу данных вопросов, т.е. вносить изменения, добавлять новые вопросы или варианты ответов, а также просматривать полную статистику результатов тестирования.

Модуль студент представляет собой отдельный исполняемый файл с набором библиотек и служит только для проведения тестирования и просмотра своих результатов. Внести какие-либо изменения в базу данных не получится, т.к. база данных располагается отдельно и устанавливается только на сервер, где устанавливается модуль «преподаватель». Для корректной работы модуля «студент» на ПК требуется установка двух бесплатных приложений, которые входят в состав комплектации: MySQL Connector c++ и MySQL Connector ODBC. В редких случаях, может потребоваться установка на локальный компьютер MySQL 5.5.11. После настройки ODBC на сервер, где установлена база данных вопросов и модуль администратор, программа готова к работе.

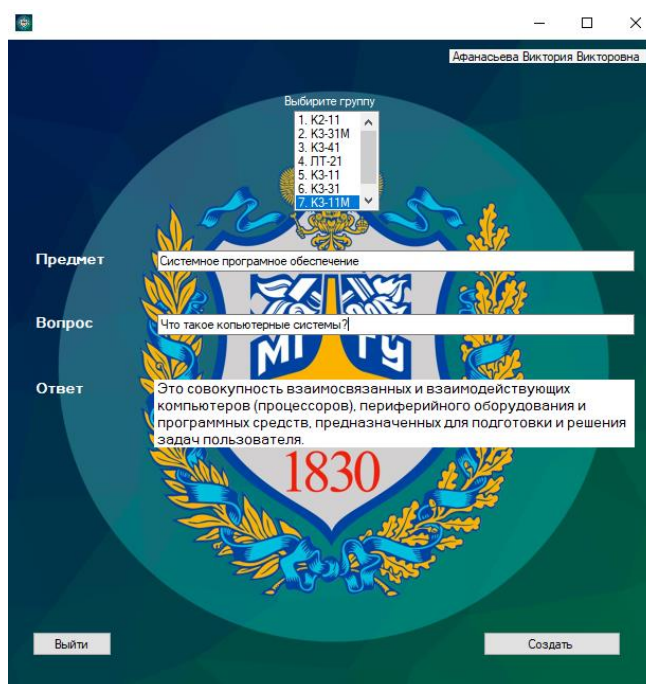


Рисунок 2. Пример создания вопроса

Модуль «преподаватель» позволяет вносить изменения в базу данных, т.е. править существующие вопросы и ответы, добавлять новые вопросы или варианты ответов, а также просматривать полную статистику результатов тестирования. Среди прочих инструментов для удобства преподавателей планируется внедрить функцию импорта результатов в Microsoft Excel.

При добавлении вопросов в базу данных, от организаторов тестирования требуется указать, к какой конкретно теме относится тот или иной вопрос. Далее заполняются поля «Ответ», а также отмечается, какой ответ будет являться правильным. Это требуется для того, чтобы в результатах тестирования отобразить общее количество вопросов, на которые был дан верный ответ. После того как нажать на «создать» появится поле для ввода обучающего множества и коэффициента истинности ответа.

Заключение

Интеллектуальная система контроля знаний предназначена для тестирования по дисциплинам ответы, на которые необходимо давать на в развернутом виде на естественнонаучном языке.

Программа может использоваться в ВУЗах, средне специальных учебных заведениях и организациях, где требуется проводить анализ текста.

В настоящее время ведется разработка клиент-серверного приложения, которое позволит устранить существующие ограничения, а также использовать систему тестирования MgtuMf на любой операционной системе через Браузер.

Использованные источники:

1. Analytical survey Distance Education for the Information Society: Policies, Pedagogy and Professional Development. Moscow 2000, 86 pp., UNESCO Institute for Information Technologies in Education, p.3-6.
2. Мельчук И.А. Опыт теории лингвистических моделей «Смысл-Текст» - Москва: «Языки русской культуры», 1999. – 346 с.