

ТЕХНОЛОГИИ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

УДК 159.9+681.3

ОСОБЕННОСТИ ПРОГРАММНОЙ РЕАЛИЗАЦИИ АДАПТИВНОГО ТЕСТА ОЦЕНКИ ИНТЕЛЛЕКТА БЕРТЛИНГА-ХОЛЛИНГА

В.К. Войтов

Рассмотрены особенности программной реализации адаптивного теста общего интеллекта.

The peculiarities of software implementation of an adaptive general intelligence test are considered.

1. ВВЕДЕНИЕ

В настоящее время адаптивные тесты активно используются в психологии и педагогике для решения задач диагностики и контроля качества обучения. Порядок заданий в тестах данного типа определяется результатами испытуемых, полученными при выполнении предшествующих заданий.

Программная реализация содержательных адаптивных тестов является весьма трудоемкой процедурой и на практике может быть выполнена только с использованием современных компьютерных технологий. Необходимые программные средства часто приходится разрабатывать заново, причём методы и технические решения, используемые для этого программистами, представляют самостоятельный научный и практический интерес.

В этой работе рассмотрены технические особенности программной реализации адаптивного теста оценки общего интеллекта, предложенного в 2009 году Дж. Бертлингом и Х. Холлингом из Университета города Мюнстера (Германия) и построенного на основе технологии IRT (*Item Response Theory*) [2, 3]. Реализована вторая программная версия теста, которая функционирует в сети Интернет.

Более подробно математические аспекты рассматриваемого здесь теста изложены в статье «Расчет значений сложностей заданий для адаптивного теста интеллекта» в этом же сборнике.

2. ОПИСАНИЕ ТЕСТА

При вызове теста появляется окно с фрагментом, изображенным на рис. 1. Тест запускается при нажатии кнопки «Start», но до запуска пользователю следует ввести свой идентификатор и пароль. Когда пользователь вызывает тест первый раз, его имя и пароль автоматически заносятся в базу данных, то есть происходит регистрация. Пользователь также может ввести номер группы, но это необязательный параметр.

При нажатии кнопки «information» без имени пользователя выдается информация о нескольких лучших результатах пользователей. При нажатии кнопки «information» с ука-

занными именем пользователя и паролем выдается информация о всех результатах этого пользователя.

Input your name:

Input password:

Input group:

Рис. 1. Вызов теста.

Использовавшийся в работе тест принадлежит к группе невербальных тестов индуктивного мышления. В тестах индуктивного мышления испытуемый должен вначале обнаружить правила, на основании которых организовано тестовое задание, а затем применить это правило для выбора надлежащего ответа. Наиболее известными тестами такого рода являются тесты Равена, тестовое задание одного из которых (АРМ) приведено на рис. 2.

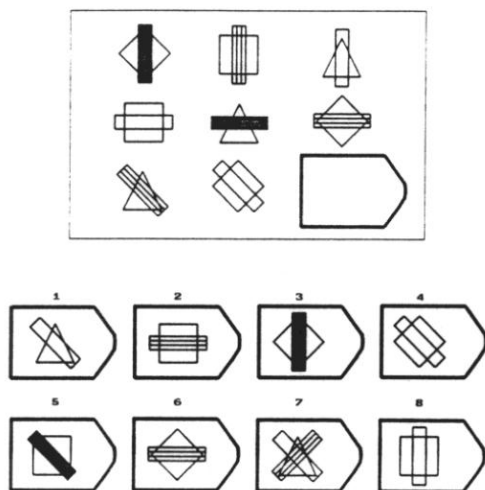


Рис. 2. Задание теста АРМ Равена.

Чтобы правильно решить представленное на рис. 2 задание, испытуемому необходимо обнаружить 3 правила.

Правило А: Каждый ряд содержит три геометрические фигуры (ромб, квадрат и треугольник), распределенные между тремя колонками.

Правило Б: Каждый ряд содержит три линии с разными текстурами (темная, штрихованная и светлая), распределенные между тремя колонками.

Правило В: Ориентация линий (вертикальная, горизонтальная или наклонная) одинакова внутри каждого ряда, но различается для разных рядов.

На основе этих правил легко найти верный ответ для пропущенной в нижнем правом углу фигуры, это ответ №5.

Тесты индуктивного мышления, как показали исследования, являются очень эффективным методом оценки общего интеллекта [1], потому широко применяются как исследователями,

так и практиками. Однако порой возникает задача предъявлять тест испытуемым много раз, при этом повторять одни и те же задания нежелательно ввиду научения. Более того, тест, включающий фиксированный набор заданий, быстро становится общеизвестным, что делает возможным предварительный доступ к нему тестируемых и приводит к потенциальному искажению результатов.

Выходом из описанной ситуации было бы создание теста, который мог бы по ходу тестирования создавать новые задания из фиксированного набора элементов. При этом необходимо, чтобы сложность задания была бы заранее известна, т.е. набор элементов определял эту сложность. Дополнительным преимуществом подобного теста была бы возможность адаптивного тестирования: испытуемому предъявляются задания того или иного уровня сложности в зависимости от того, насколько успешно он решал предшествующие задания. Именно такой тест, разработанный Холлингом и Бертлингом, применялся в настоящей работе.

Тест предполагает нахождение правил, определяющих соотношение фигур А и В, и выбор из различных вариантов ответов такой фигуры D, которая находится в таком же отношении к данной фигуре С, как В к А.

После нажатия кнопки «Start» запускается тест и испытуемому предлагаются три сложные геометрические фигуры А, В и С. Первые две фигуры одинаковы. Вторая фигура должна отличаться от первой положением в пространстве. Кроме того, в фигурах присутствуют дополнительные элементы (например, точки, прямоугольники, кружки и т.д.), которые могут быть перемещены во второй фигуре в другие места.

Третья фигура отличается по форме от первых двух фигур. Она имеет те же детали, что и первые фигуры, но в других местах.

Пользователь должен определить закономерность в изменении второй фигуры (фигура В) по отношению к первой фигуре (фигура А) и выбрать из предъявленных восьми фигур ту фигуру, которая получается из третьей фигуры (фигура С) при учёте указанных выше закономерностей. Варианты выбора обозначены маленькими латинскими буквами от «а» до «i». Буква «i» соответствует тому случаю, когда среди восьми предложенных решений (от «а» до «h») нет подходящего.

3. МАТЕМАТИЧЕСКИЕ АСПЕКТЫ ТЕСТА

На первом этапе тестирования по результатам выполнения короткой последовательности заданий разной степени трудности формируется начальная оценка интеллектуальных способностей испытуемого. Второй этап тестирования представляет собой итерационную процедуру, на каждом шаге которой респонденту в соответствии с концепцией Г. Раша [2,3] предлагается очередное задание с трудностью, соответствующей текущей оценке его способностей. Эта оценка уточняется по результатам выполнения каждого нового задания.

Пусть вектор $\mathbf{x} = (x_1, x_2, \dots, x_n)$ представляет результаты испытуемого, полученные в n тестах. Величина x_i принимает значение 1 в случае правильного ответа и 0 – в случае неправильного ответа. Вероятность P получения последовательности ответов (x_1, x_2, \dots, x_n) при этом выражается следующим произведением:

$$P(x_1, x_2, \dots, x_n | \theta, \sigma_1, \sigma_2, \dots, \sigma_n) = \prod_{i=1}^n P(x_i | \theta, \sigma_i),$$

где θ – текущая оценка способностей, σ_i – трудность теста i , а вероятность $P(x_i | \theta, \sigma_i)$ согласно модели Раша выражается логистической функцией

$$P(x_i | \theta, \sigma_i) = \frac{e^{x_i(\theta - \sigma_i)}}{1 + e^{x_i(\theta - \sigma_i)}}$$

Текущая оценка величины θ вычисляется методом максимального правдоподобия, при численной реализации которого используется итерационная процедура Ньютона-Рафсона.

Применяемый подход позволяет получать стандартные отклонения и границы доверительных интервалов для характеристики θ , что даёт возможность останавливать процесс тестирования по достижении заданных значений этих параметров.

При построении заданий теста Бертлинга-Холлинга используются несколько различных геометрических, половина из которых являются выпуклыми, а половина – нет. Каждая фигура может занимать восемь различных положений. Прилагаемые к фигурам детали создают новые различия. Будем называть эти различия характеристиками. Многообразие вариантов фигур позволяет генерировать тесты различной трудности. Трудность каждого сгенерированного теста рассчитывается по следующей формуле:

$$\sigma_i = \sum_{j=1}^k q_{ij} \beta_j,$$

где $i=1, \dots, n$ – номер задания теста, β_j – трудность, вносимая j -й характеристикой теста, q_{ij} – параметры, принимающие значения 1 или 0 в зависимости от использования или неиспользования соответствующих деталей фигур. Всего в тесте присутствует 14 характеристик, задающих различную степень трудности.

4. ОСОБЕННОСТИ ПРОГРАММНОЙ РЕАЛИЗАЦИИ

Для реализации основных компонентов теста рассматривались три возможных варианта системы программирования, а именно:

- *ActionScript* [4-7];
- *JavaScript* [8-10];
- *PHP* [11-12] в комбинации с *JavaScript*.

Каждый из этих вариантов имеет свои недостатки:

- из программ, написанных на языке *ActionScript*, неудобно передавать данные в программы, написанные на языке *PHP*;
- на языке *JavaScript* неудобно рисовать наклонные линии;
- генерация фигур на *PHP* может привести к временным задержкам их отображения при передаче данных по сети *Internet*.

В результате сопоставления характеристик указанных систем для программной реализации содержательной части теста был выбран язык *ActionScript*. Системы *JavaScript* и *PHP* использовались для регистрации входа пользователей в систему и поддержки их взаимодействия с базой данных.

Система состоит из следующих функциональных блоков и модулей:

- административного модуля;
- модуля регистрации пользователя;
- модуля интерактивного взаимодействия с пользователем;
- модуля генерации теста;
- модуля отображения;
- модуля вычисления текущей оценки способностей;
- модуля взаимодействия с базой данных.

Тест поддерживает три уровня пользователей:

- администратор базы данных, который имеет полный доступ ко всем таблицам и обладает правами для их создания и удаления;
- администратор процесса работы теста;
- обычный пользователь.

Информация о пользователях, включающая имена и пароли, хранится в базе данных.

Число заданий в тесте задается администратором при настройке теста.

В основном окне теста имеется кнопка «следущ» (*next*). При её нажатии модуль интерактивного взаимодействия обращается к модулю генерации теста для создания следующего задания, а затем – к модулю отображения, чтобы вывести на экран составленное задание. При нажатии кнопки «правила» (*rules*) появляется окно с описанием правил пользования тестом.

Справа расположены восемь вариантов ответа. Пользователь должен учесть закономерности в изменении второй фигуры относительно первой и выбрать из предъявленных восьми фигур ту, которая получается из третьей фигуры с учётом заданных выше закономерностей. Если такой фигуры нет, то пользователю следует выбрать вариант «No correct alternative». Выбор осуществляется щелчком левой кнопки мыши. Получив ответ, модуль интерактивного взаимодействия обращается к модулю расчёта способностей, после чего в окне появляется кнопка «следущ». До выбора нужного варианта мышью этой кнопки не видно.

Обозначим левые четыре фигуры теста буквами А, В, С и D. Для обеспечения выбора сочетаний вида фигуры и её деталей эти сочетания были пронумерованы, после чего был составлен соответствующий массив, отсортированный в порядке возрастания значений трудностей.

При генерации фигур элементы массива сочетаний выбираются случайным образом. Для фигур А и В выбираются:

- вид фигуры;
- положение фигуры;
- вид деталей;
- расположение деталей;

Для фигуры С выбираются:

- вид фигуры (из оставшихся – за исключением фигуры, выбранной для А и В);
- положение фигуры;
- расположения деталей.

Вид деталей фигуры А должен совпадать с видом деталей фигуры С. Фигура D, которую должен найти испытуемый, при отображении на экране заменяется вопросительным знаком. В случае отсутствия фигуры D среди предложенных вариантов ответа испытуемому следует выбрать «No correct alternative».

После сочетаний вида фигуры и её деталей генерируются альтернативы a, b, \dots, i . Событие «Нет корректного варианта» («No correct alternative») генерируется с вероятностью $1/10$. Остальные восемь вариантов ответа случайным образом делятся на три группы равной вероятности ($3+3+2$, $3+2+3$, $2+3+3$). Расположение фигур первой группы совпадает с расположением фигуры правильного ответа, который, если он есть, содержится в этой же группе. Другие элементы этой группы имеют одно отличие от правильного ответа. Вторая группа содержит элементы с двумя отличиями от правильного ответа, а третья группа – с тремя.

Модуль отображения используется для вывода на экран фигур А, В, С и D, девяти вариантов ответа, результатов прохождения теста и параметров теста в режиме администратора.

В левом верхнем углу окна отображается общее время выполнения теста и время выполнения текущего задания. Строкой ниже отображаются номер задания и число выполненных заданий в процентах (например, Item 2 40%). После выполнения теста в третьей строке может выдаваться достигнутый процент от наивысшего результата (например, *result*: – 74%).

При входе в систему в режиме администратора можно просмотреть и задать параметры теста в ручном режиме.

5. ОСОБЕННОСТИ РАБОТЫ ТЕСТА И НАПРАВЛЕНИЯ ЕГО РАЗВИТИЯ

Для хранения результатов теста поддерживается база данных со списком идентификаторов испытуемых и информацией о результатах прохождения теста.

Тест может функционировать как в сети Internet, так и в локальном режиме. Для вызова теста через Internet на компьютере должен быть установлен проигрыватель Flash. При использовании теста в локальном режиме пользователю следует создать на своём компьютере Web-сервер и установить систему управления базами данных MySQL, для чего необходима свободно распространяемая программа Денвер.

В новой версии программной реализации планируется:

- создать более гибкую структуру базы данных, расширив возможности хранения информации о прохождении теста;
- обеспечить перенос фрагментов, написанных на языке ActionScript, на язык JavaScript;
- использовать систему FLEX [13] вместе с Adobe Flash CS3.

6. ЗАКЛЮЧЕНИЕ

Выполнена компьютерная реализация адаптивного теста оценки интеллекта, предложенного профессором Хейнцем Холлингом из университета города Мюнстера (Германия). Тест предполагает генерацию большого числа различных заданий, суть которых состоит в выявлении закономерностей в строении геометрических фигур с дополнительными элементами. Тест показал стабильное функционирование.

При вычислении способностей использовалась IRT (Item Response Theory) технология. Полученные данные имеют приближенный характер, но по мере эксплуатации теста, вновь полученные результаты можно использовать для уточнения значений трудностей заданий, так как они сохраняются в базе данных.

Тест доступен по адресу: <http://it-fat.mgppu.ru>.

ЛИТЕРАТУРА

1. Snow R.E., Kellonen P.C., Marshalek B. The topography of abilities and learning correlations // R.J. Sternberg (Ed.) Advances in psychology of human intelligence. Hillsdale, NJ: Erlbaum. P. 47-103.
2. V. Natarajan. Basic Principles of IRT and Application to Practical Testing & Assessment: <http://www.docstoc.com/docs/8937076/Basic-Principles-of-Item-Response-Theory>.
3. В.С. Аванесов. Основные понятия и положения математической теории измерений (Item Response Theory): <http://testolog.narod.ru/Theory60.html>.
4. Adobe Flash CS3. Официальный учебный курс. Разработан экспертами компании Adobe. Триумф 2008.
5. Мук К. ActionScript 3.0 для Flash. Подробное руководство. СПб: «Питер», 2010.
6. Панкратова Т. Flash MX 2004. Учебный курс. СПб: «Питер», 2004.
7. Дунаев В. Flash MX 2004. Самоучитель. СПб: «Питер», 2004.
8. Айзек С. Dynamic HTML. СПб.: «БХВ – Петербург», 2001.
9. Шапошников И. Самоучитель HTML 4. Санкт-Петербург “БХВ - Петербург”, 2003.
10. Монкур М. Освой самостоятельно JavaScript за 24 ч. М. и др. Издательский дом “Вильямс” 2001.
11. Коретов Д. Самоучитель PHP 4. Санкт-Петербург, «БХВ – Петербург», 2003.
12. Мазуркевич А., Еловой Д. PHP: настольная книга программиста. Минск, ООО «Новое знание», 2003.
13. Коул А. Изучаем Flex 3. СПб: Символ-плюс, 2009.

Работа поступила 23.03.2011