Карпов В.Э., Карпова И.П.

ЯЗЫК ОПИСАНИЯ СИСТЕМЫ КОНТРОЛЯ ЗНАНИЙ [1]

Одной из наиболее сложных задач при создании программных средств учебного назначения является контроль знаний.

В настоящее время существует большое количество систем контроля знаний (СКЗ), как выполненных в виде отдельных программных продуктов, так и встроенных в обучающие системы. Отсутствие стандарта или хотя бы общепринятых норм и правил написания таких систем привело к тому, что практически каждая из них является "вещью в себе" и используется, за редким исключением, только ее собственными разработчиками.

В данной статье представлен язык описания системы контроля знаний. Использование такого языка, в частности, делает создаваемые тесты переносимыми на различные программные и аппаратные платформы (без изменения самих тестов). При создании языка учитывались используемые на практике типы вопросов и ответов и возможность подключения внешних модулей.

Цели и задачи

Создание языка описания СКЗ решает такие задачи, как:

1. стандартизация и унификация СКЗ. Разработка развитого средства описания СКЗ позволяет:
   - унифицировать интерфейс, т.е. использовать одну и ту же оболочку для различных контролирующих и тестовых систем, что приведет к снижению трудозатрат пользователя на изучение системы и позволит ему сосредоточиться на ответах на вопросы;
   - создавать разные оболочки для одного и того же набора контрольных заданий, адаптируя систему для различных операционных систем, аппаратных платформ и требований по организации интерфейса.
2. переход на качественно более высокий уровень общения СКЗ и обучаемого, реализация полноценной обратной связи "обучающая система – обучаемый".

Описание системы контроля знаний

Любая система контроля знаний включает в себя множество предлагаемых вопросов, правильные ответы на них и правила определения корректности ответов, полученных в ходе опроса. Поэтому описание СКЗ состоит из следующих частей:

1. описание вопросов;
2. описание ответов, включающее эталоны ответов, методы определения правильности и оценки за правильные ответы;
3. описание алгоритма контроля знаний (правил формирования списка контрольных заданий на основании характеристик самих заданий).

В настоящем проекте определяются возможные типы вопросов, ответов и оценок полученных ответов, а также синтаксис их описания.

Язык имеет декларативный характер, что позволяет отделить описание теста от его алгоритмической реализации, сделать его переносимым и предоставить некоторую свободу разработчику.

На основании анализа существующих тестовых систем можно выделить:

- типы вопросов (по форме представления):
  1. Текст.
  2. Изображение.
  3. Процесс.
  4. Речевое сообщение (сводимо к типам "текст" и/или "процесс").
- типы ответов (по форме ввода и представления):
  1. Множество элементов (неупорядоченное).
  2. Список элементов (упорядоченный).
     (В качестве элементов в пп.1,2 могут выступать списки и множества.)
  3. Выражение (арифметическое).
  4. Фраза (текст).
  5. Рисунок.
  6. Речевое сообщение (сводимо к фразе).

Оценка ответов

Алгоритм формирования оценки ответа определяется типом ответа.

С практической точки зрения можно выделить следующие типы ответов:

1. Множество элементов (неупорядоченное).
   Правильность ответа оценивается путем сравнения множеств.
2. Список элементов (упорядоченный).
   Правильность ответа проверяется путем определения сходства списков.
3. Выражение (арифметическое).
   Для выражения должны быть определены эталон (правильное значение) и допустимая погрешность ε.
4. Фраза (текст).
   Указывается эталон и способ анализа:
   - полное/частичное совпадение на лексическом уровне;
   - полное/частичное совпадение на синтаксическом уровне;
   - полное/частичное совпадение на семантическом уровне.
5. Рисунок (изображение).
   Сравнение рисунков подразумевает задание эталона и анализ сходства изображений, выявление характерных участков и т.п. Сравнение рисунков как отдельную сложную задачу предлагается вынести за рамки ядра тестовой системы и реализовать в виде внешней утилиты.

Примечание: некоторые методы оценки ответов приведены в разделе "Возможные методы оценки ответов".

Описание теста

Описание СКЗ хранится в текстовом файле и представляет собой последовательность блоков описаний вопросов, ответов и типов контроля. Каждая фраза (блок) начинается с одного из ключевых слов ВОПРОС, ОТВЕТ, КОНТРОЛЬ_ПО_ТЕМЕ или СПИСОК_ТЕМ и заканчивается ключевым словом КОНЕЦ. Внутри каждой фразы (в произвольном порядке) располагаются ключевые слова, определяемые типом фразы, и данные, относящиеся к этим ключевым словам. Необязательные ключевые слова могут быть опущены (в приведенном ниже описании они заключаются в квадратные скобки).

Комментарии. Тип комментария – строчный. В качестве символа начала комментария используется символ ‘*’(звездочка), стоящий в начале строки.

Разделители. Разделителями лексем являются символы пробела, табуляции и перевода строки. Кроме того, символы ‘=’ и ‘:’ (знак равенства и двоеточие) являются специальными разделителями; они рассматриваются как разделители наряду с пробелами и табуляциями, но должны стоять в местах, определенных синтаксисом описания.

Идентификатор является произвольной последовательностью символов длиной не более 255. Если идентификатор содержит пробелы, он должен быть заключен в двойные кавычки.

Система нечувствительна к регистру: идентификаторы и ключевые слова могут быть написаны как прописными, так и строчными буквами. Текст, заключенный в двойные кавычки, воспринимается так, как он записан, т.е. с учетом регистра.

Ключевые слова СКЗ в приведенном ниже описании записаны прописными буквами.

Типы данных

В качестве данных система воспринимает:

- идентификаторы – последовательности символов, не содержащие пробелов;
- строки – произвольные последовательности символов, заключенные в двойные кавычки;
- выражения – вычислимые последовательности символов (констант и идентификаторов, имеющих значение), разделенные знаками арифметических операций.

Блок описания вопроса

ВОПРОС = <имя вопроса>
[ТЕМА = <тема>]
ТИП = <тип вопроса>
[СВОЙСТВО <идентификатор свойства 1> **:** <значение 1>
СВОЙСТВО <идентификатор свойства 2> **:** <значение 2>
. . . . .
СВОЙСТВО <идентификатор свойства n> **:** <значение n>]
СОДЕРЖАНИЕ = <текст> | <командная строка>
КОНЕЦ

<имя вопроса> - идентификатор вопроса (название).

<тема> - идентификатор темы (название).

<тип вопроса> - одно из ключевых слов: ТЕКСТ, КОМАНДА.

СВОЙСТВО - характеристика вопроса, вводимая и определяемая пользователем: <**идентификатор свойства i>** - название i-го свойства, **<значение i>** - значение (произвольно типа) этой характеристики для данного вопроса. Свойства могут быть использованы при формировании списков вопросов для тестирования или при создании модели обучаемого.

ТЕМА - объявление темы может встречаться в любом месте текста. В зависимости от местоположения различаются глобальное и локальное объявление темы. *Глобальным* считается объявление, расположенное непосредственно перед началом фразы, его действие распространяется на

все вопросы до следующего объявления (если в вопросе не введено локальное объявление). *Локальным* считается объявление, расположенное внутри блока описания вопроса, оно действует только в пределах описания вопроса (если объявлений темы в одном блоке вопроса несколько, то действующим считается последнее).

Блок описания ответа

ОТВЕТ = <идентификатор вопроса>
[ТИП = <тип ответа>]
ЭТАЛОН = <список> | <множество> | <текст> | <выражение>
БАЛЛ = min **:** max
ОЦЕНКА = <тип оценки>
КОНЕЦ

<тип ответа> - одно из ключевых слов: СПИСОК, МНОЖЕСТВО, ТЕКСТ, ВЫРАЖЕНИЕ.
ЭТАЛОН определяет один из возможных типов ответов:

- **<список>** - список разделенных пробелами и/или запятыми вариантов, заключенный в квадратные скобки, например [a, b, c].
- **<множество>** - множество разделенных пробелами и/или запятыми вариантов, заключенных в фигурные скобки, например: {a b c}.
- **<текст>** - произвольная последовательность символов, заключенная в двойные кавычки, например: “Физика высоких температур”.
- **<рисунок>** - ссылка на рисунок (например, имя файла).
- **<выражение>** - арифметическое выражение, например: (a+b)/2.

Примечание: тип комбинированного ответа (множество списков или список множеств) определяется внешними скобками. Например:
[{a, b},{c},{d, e}] - СПИСОК
{[a], [b, c, d], [e]} - МНОЖЕСТВО

БАЛЛ - максимальный и минимальный балл, который можно получить за правильный или неправильный ответ на данный вопрос (разделяются двоеточием).
<тип оценки> - описывает способ определения правильности ответа:

- **РАВНО(<величина погрешности>)** - если результат укладывается в интервал (эталон)±(величина погрешности), то ответ считается правильным (оценка равна max), в противном случае ответ неверен, (оценка равна min).
- **СХОДСТВО(<величина погрешности>)** - величина погрешности изменяется в интервале [0,1]. Если степень сходства ответа и эталона не меньше указанной погрешности, то ответ считается правильным (оценка равна max), иначе ответ неверен (не совсем верен) и оценка *O* может быть вычислена так:

$$O = max - (1 - s)(max - min) \qquad (1)$$

где *s* - степень сходства ответа и эталона, *s* принадлежит интервалу [0,1]; max, min - максимальная и минимальная оценки за ответ.

Величина погрешности может задаваться арифметическим выражением.

Блок описания выбора тем

Этот блок дает возможность ограничить список тем, участвующих в контроле, и определить порядок выдачи тем для контроля.

СПИСОК_ТЕМ = <список тем> | ВСЕ
[СПОСОБ_ВЫДАЧИ = ОДИН | ВСЕ]
КОНЕЦ

СПИСОК_ТЕМ - список тем через запятую или все темы (ключевое слово ВСЕ). Темы для опроса выбираются в том порядке, в котором они указаны.

СПОСОБ_ВЫДАЧИ - выдается по одной теме (ОДИН) или все (ВСЕ) темы сразу в виде списка тем. В последнем случае пользователь сам может выбирать последовательность тем для ответов.

Блок описания контроля по теме

Этот блок позволяет указывать методы формирования списка вопросов по теме, выдачи этих вопросов обучаемому и подсчета итоговой оценки. Можно указывать разные методы для разных типов контроля.

КОНТРОЛЬ_ПО_ТЕМЕ = <название темы>
[СПИСОК = {<список вопросов>} | ВСЕ]
ТИП = <тип контроля 1>
[СПОСОБ_ВЫБОРА = СЛУЧАЙНО_БЕЗ_ПОВТОРОВ |
СЛУЧАЙНО | ПОЛЬЗОВАТЕЛЬ]
[СПОСОБ_ВЫДАЧИ = ОДИН | ВСЕ]
[КОЛИЧЕСТВО_ВОПРОСОВ = <кол-во задаваемых вопросов>]
[КОЛИЧЕСТВО_ОТВЕТОВ = <количество полученных ответов>]
ОЦЕНКА = <число> | <процент>
[ТИП_ОЦЕНКИ = СУММА | ПРОЦЕНТ |
ШКАЛА <низший балл> **:** < высший балл >]
[ТИП = <тип контроля 2>
…]
[ТИП = <тип контроля 3>
…]
КОНЕЦ

СПИСОК - для каждой темы может быть указан конкретный список идентификаторов вопросов (через запятую) или ключевое слово ВСЕ. Если данный параметр опущен, то это означает, что вопросы будут выбираться из всего списка вопросов по данной теме.

ТИП - могут быть указаны разные типы контроля. Тип контроля – одно из ключевых слов:

- ЗАВЕРШАЮЩИЙ - итоговый контроль, по результатам которого определяется, сдана ли тема. Например, если тема сдана (полученная оценка не меньше той, что указана в предложении ОЦЕНКА (см. ниже)), то вопросы по этой теме обучаемому больше задаваться не будут.
- ПРОМЕЖУТОЧНЫЙ - контроль знаний с выставлением оценки по теме. Например, если промежуточный контроль проводится несколько раз по

одной и той же теме, то программа может хранить последнюю оценку и среднее арифметическое всех полученных оценок по данной теме.

- ВЫБОРОЧНЫЙ - по результатам контроля выставляется одна общая оценка за все темы, участвующие в выборочном контроле. Может использоваться при итоговом контроле знаний по курсу.

СПОСОБ_ВЫБОРА - способ выбора вопроса из списка:

- СЛУЧАЙНО_БЕЗ_ПОВТОРОВ - вопрос выбирается из списка случайным образом, учитываются ранее заданные вопросы (значение по умолчанию).
- СЛУЧАЙНО - вопрос выбирается случайным образом из указанного (или общего) списка.
- ПОЛЬЗОВАТЕЛЬ - пользователь сам выбирает вопрос из полного списка вопросов. Может использоваться для самоконтроля.

СПОСОБ_ВЫДАЧИ - способ выдачи вопросов из списка:

- ОДИН - пользователю выдается по одному вопросу из списка выбранных вопросов. По умолчанию выдается по одному вопросу.
- ВСЕ - сразу выдается весь список вопросов по теме, пользователь сам выбирает очередной вопрос из списка.

КОЛИЧЕСТВО_ВОПРОСОВ - количество задаваемых вопросов. Если параметр опущен, задаются все вопросы из списка.

КОЛИЧЕСТВО_ОТВЕТОВ - количество полученных ответов, при котором опрос считается состоявшимся. Если параметр опущен, необходимо ответить на все заданные вопросы (опрос нельзя прервать).

ОЦЕНКА - оценка (число или процент правильных ответов), при которой тема считается сданной.

ТИП_ОЦЕНКИ - способ определения общей оценки *O* за тему на основании оценок *Oi*, полученных за каждый ответ:

- СУММА (по умолчанию):

$$O = \sum_{i=1}^{M} O_i \tag{2}$$

где *M* – общее число ответов, i=1,…,M.

- ШКАЛА <min> **:** <max> - оценка нормализуется в соответствии с указанной размерностью балльной шкалы:

$$O = \frac{\sum_{i=1}^{M} O_i * (\max - \min) + \min}{M} \tag{3}$$

где *M* – общее число ответов, i=1,…,M.

- ПРОЦЕНТ:

$$O = 100(m/M) \tag{4}$$

где *M* – общее число ответов, *m* – число правильных ответов.

Возможные методы оценки ответов

Приведем некоторые методы оценки ответов различных типов.

1. Для множества элементов (неупорядоченного) правильность ответа оценивается путем сравнения множеств.

Например, можно оценить ответ обучаемого по степени сходства *s* между ответом обучаемого (множество *Ra* ={a_1, a_2,…, a_i, b_1, b_2,…, b_j}) и эталонным ответом (множество *Re*), которая подсчитывается по формуле

$$s = Na/(Qa + Nb) \quad (5)$$

где *Qa* - мощность множества *Ra*; *Na* - количество элементов a_i, принадлежащих объединению множеств *Ra* и *Re*; *Nb* - количество элементов b_j, не принадлежащих объединению множеств *Ra* и *Re*.

2. Для списка элементов (упорядоченного) правильность ответа проверяется путем определения сходства списков.
   Один из вариантов определения сходства списков базируется на понятии расстояния между списками и происходит в два этапа:
   - Определение степени сходства множеств. Списки *Ra* = {a_1, a_2,…, a_i, b_1, b_2,…, b_j}и *Re* рассматриваются как множества и по формуле (5) определяется степень сходства этих списков s_1.
   - Определение степени сходства списков. Элементы b_j, не принадлежащие объединению списков *Ra* и *Re*, в списке *Ra* заменяются на элементы, заведомо большие, чем любой элемент списка *Re*; получается список *R'a*. Далее путем попарных перестановок список *R'a* упорядочивается в соответствии с отношением порядка, введенным эталонным списком *Re*. Максимальное расстояние между списками принимается равным максимальному числу перестановок *Nj*, возможных для списка длиной *j*, которое равно

     $$Nj = j(j-1)/2 \quad (6)$$

     Степень сходства списков s_2 вычисляется по формуле (3)

     $$s_2 = 1 - Ni / Nj \quad (7)$$

     где *Ni* - количество перестановок (инверсий) списка *R'a*, *Nj* - максимальное количество перестановок для списка длиной *j*.
     Получение окончательной величины степени сходства списков может быть выполнено одним из следующих способов:
     a) $s = (s_1 + s_2)/2$ – среднее арифметическое s_1 и s_2;
     b) $s = \max(s_1 , s_2)$ – максимум из s_1 и s_2;
     c) $s = \min(s_1 , s_2)$ – минимум из s_1 и s_2;
     d) $s = s_1 * s_2$ – произведение s_1 и s_2 .

3. Для арифметического выражения должны быть определены эталон (правильное значение) и допустимая погрешность *e*. Один из вариантов использования погрешности заключается в том, что если ответ находится в *e*-окрестности эталона

   $$(E - e) \le A \le (E + e) \quad (8)$$

   где *E* - эталон, *A* – полученный ответ, то ответ считается правильным и за него выставляется максимальная оценка, иначе ответ неверен и выставляется минимальная оценка.

4. Рассмотрим случай, когда элементами множества *Re* являются множества *Si* или списки *Li* (i = 1,…,N). Тогда оценка ответа *Ra* может быть выполнена в два этапа:
   a) каждый элемент (множество *Si* или список *Li*) сравнивается со всеми элементами эталонного множества *Re* (по правилам сравнения множества или списка соответственно). В качестве степени сходства этого элемента с элементами эталонного множества выбирается

максимальная степень сходства между элементами двух множеств. (Элемент эталонного множества, наиболее близкий к рассматриваемому элементу ответа, исключается из дальнейшего рассмотрения.) Из этих степеней сходства формируется числовой вектор *P*. Длина этого вектора равна мощности *N*. Если элементу *Si* (*Li*) множества *Ra* не соответствует никакой элемент эталонного множества *Re*, то i–й элемент вектора *P* равен 0.

b) степень сходства ответа и эталонного множества вычисляется, например, как среднее арифметическое элементов вектора P.

5. Оценка списка множеств (списков) также может выполняться в два этапа:
   a) на первом этапе выполняется работа, аналогичная первому этапу при сравнении множеств. Степень сходства s_1 множества элементов ответа вычисляется, например, как среднее арифметическое элементов вектора *P*. Если элементу *Si* (*Li*) списка *Ra* не соответствует никакой элемент эталонного списка *Re*, то i–й элемент вектора *P* равен 0. На основании вектора *P* каждому элементу списка *Ra* ставится в соответствие элемент эталонного списка. Таким образом, формируется вектор соответствия *U*, каждый i-й элемент которого равен номеру наиболее близкого элемента списка *Re* или числу, превышающему количество элементов в эталонном списке.
   b) полученный список соответствия *U* упорядочивается путем попарных перестановок. Степень сходства s_2 вычисляется по правилам определения степени сходства списков (7). Окончательно степень сходства *s* ответа и эталонного множества вычисляется, например, как среднее арифметическое, минимум или произведение s_1 и s_2.

---

[1] Статья опубликована в журнале "Компьютеры в учебном процессе", 2000, №4. – с. 147-155.