

Krassimir Markov, Vitalii Velychko, Oleksy Voloshin
(editors)

**Information Models
of
Knowledge**

**ITHEA[®]
KIEV – SOFIA
2010**

Krassimir Markov, Vitalii Velychko, Oleksy Voloshin (ed.)

Information Models of Knowledge

ITHEA®

Kiev, Ukraine – Sofia, Bulgaria, 2010

ISBN 978-954-16-0048-1

First edition

Recommended for publication by The Scientific Council of the Institute of Information Theories and Applications FOI ITHEA
ITHEA IBS ISC: 19.

This book maintains articles on actual problems of research and application of information technologies, especially the new approaches, models, algorithms and methods for information modeling of knowledge in: Intelligence metasynthesis and knowledge processing in intelligent systems; Formalisms and methods of knowledge representation; Connectionism and neural nets; System analysis and synthesis; Modelling of the complex artificial systems; Image Processing and Computer Vision; Computer virtual reality; Virtual laboratories for computer-aided design; Decision support systems; Information models of knowledge of and for education; Open social info-educational platforms; Web-based educational information systems; Semantic Web Technologies; Mathematical foundations for information modeling of knowledge; Discrete mathematics; Mathematical methods for research of complex systems.

It is represented that book articles will be interesting for experts in the field of information technologies as well as for practical users.

General Sponsor: Consortium FOI Bulgaria (www.foibg.com).

Printed in Ukraine

Copyright © 2010 All rights reserved

© 2010 ITHEA® – Publisher; Sofia, 1000, P.O.B. 775, Bulgaria. www.ithea.org ; e-mail: info@foibg.com

© 2010 Krassimir Markov, Vitalii Velychko, Oleksy Voloshin – Editors

© 2010 Ina Markova – Technical editor

© 2010 For all authors in the book.

® ITHEA is a registered trade mark of FOI-COMMERCE Co., Bulgaria

ISBN 978-954-16-0048-1

C/o Jusautor, Sofia, 2010

«ОТКРЫТЫЕ» ОТВЕТЫ В КОМПЬЮТЕРНОМ ТЕСТИРОВАНИИ ЗНАНИЙ

Леонид Леоненко

Аннотация: Обосновывается возможность адекватной оценки компьютером ответов “свободной” формы, в том числе на натуральном языке. Методы сравнения ответов, позволяющие оценивать “открытые” тестовые задания, основаны на выводах по аналогии. Для разных типов заданий применяются разные математические меры оценки «аналогичности» ответов на них. Наиболее важными являются меры, основанные на теории подобия символьных последовательностей, позволяющие адекватно оценивать лексическое и структурное сходство текстов естественных языков. Система компьютерного тестирования знаний «Control» использует «библиотеку» мер аналогичности различных видов, предоставляющую преподавателю средства оценки ответов, предъявленных в разнообразных семантических формах (причем эта библиотека может пополняться новыми типами указанных форм). Каждой семантической форме предъявления ответа соответствует так называемое решающее правило системы «Control», которое оценивает ответы учащегося, предъявленные в указанной форме. В результате для компьютера оказывается возможным игнорировать несущественные ошибки в ответе, учитывать сокращения и синонимы, разрешить, запретить или ограничить перестановки слов, и т.п. В статье приводятся примеры «открытых» тестовых заданий как по естественнонаучным, так и по гуманитарным дисциплинам, варианты возможных правильных и неправильных ответов на них, и соответствующие оценки, выставленные системой «Control». Статистический эксперимент показал, что эти оценки совпадают с оценками, выставленными преподавателями в результате письменного опроса. Включение в сферу компьютерного тестирования «открытых» тестовых заданий требует введения ряда новых понятий для их классификации, которые также обсуждаются в статье.

Ключевые слова: аналогия, меры сходства, тестирование знаний, открытые ответы

ACM Classification Keywords: I.2.6 Artificial Intelligence – Learning – Analogies; K.3.1 Computers and Education - Computer Uses in Education

Введение

«Внешнее» тестирование знаний учащихся средней и высшей школы призвано решить, в частности, следующие задачи: 1) устранение субъективных факторов в оценке знаний (и, как следствие, коррупции в сфере образования), 2) возможности сравнения по единым критериям как учебных заведений, так и систем образования регионов одной страны и разных стран. Нет сомнения – эти задачи важны. Почему же многие известные ученые и педагоги, прекрасно осознающие эту важность, не покидают лагерь критиков тестирования?

На мой взгляд, если выразить суть предмета разногласий несколькими словами, то ими будут «проблема свободных (открытых) ответов». Какое место в тестировании могут (или должны) занимать вопросы, допускающие ответы в свободной форме? В «традиционных» методиках проверки знаний они преобладают. Но существующие технологии тестирования от них почти отказались: «Для задач, допускающих вариации формы ответа, *правильный ответ должен быть включен в формулировку теста*» — вот тезис, являющийся подлинным предметом дискуссий. Различия вариантов этого тезиса, вроде «Приводится список ответов, среди которых есть как верные, так и неверные» или «Ответ должен быть сконструирован из частей, указанных в тесте, но не объединенных в нем надлежащим образом», не носят принципиального характера.

Ответы в свободной форме «плохи» тем, что их проверка требует значительных затрат при письменном тестировании; и, как обычно считается, невозможна (недостойна) при компьютерном. Цель данной статьи – оспаривание этого «обычного» мнения о недостоверности.

Выводы по аналогии – базис алгоритмов оценки ответов «открытой» формы

В работах [Леоненко и Поддубный, 1996], [Леоненко, 2002], [Леоненко, 2007] разработана методология оценки ответов свободной формы и необходимый ей математический аппарат. Методы, предложенные в этих работах, суть различные типы *выводов по аналогии* [Уемов, 1971]. Объединяет эти типы то, что во всех случаях признаком, переносимым с модели («правильного ответа») на прототип (фактический ответ), является *качество* фактического ответа, которому соответствует *оценка*, выставленная за ответ.

Я охарактеризую здесь кратко те типы выводов по аналогии, которые используются в системе компьютерного тестирования CONTROL [Леоненко и Поддубный, 1996], [Леоненко, 2009]:

1. *Различные виды парадейгмы*. Парадейгма (см. [Уемов, 1971]) используется, в частности, при сравнении ответов, представляющих собой числовые множества и последовательности (процедуры сравнения здесь исходят из меры пересечения множества-модели и множества-прототипа).

2. *Аналогии через интерпретации формул*. Модель и прототип, представляющие собой математические выражения, сравниваются путем вычисления их значений на одних и тех же значениях аргументов. Разновидности аналогий этого типа зависят от вида сравниваемых выражений (функции, уравнения, тождества) и других факторов. При этом можно разрешить учащемуся «свободно» выбирать обозначения переменных формулы.

3. *Аналогии, базирующиеся на оценке подобия текстов*. Большинство типов таких аналогий, применяемых в системе CONTROL, используют теорию подобия символьных последовательностей, описанную в [Леоненко и Поддубный, 1996], [Леоненко, 2002], [Леоненко, 2007]. Разные виды аналогий этого класса связаны с учетом либо игнорированием весов элементов текста, структуры текста, а также с дополнительными ограничениями.

4. *«Комбинированные» аналогии, применяющие к модели и прототипу несколько разнотипных процедур сравнения*. В частности, такие аналогии используются, когда форма предъявления одинаковых по смыслу прототипов может быть разной (скажем, текст или число); или когда прототип содержит разнородные части (текст и формулу, либо семантически различные фрагменты текста).

Применение выводов по аналогии к оценкам подобия текстов различной природы позволяет системе компьютерного тестирования CONTROL предъявлять студенту вопросы, предполагающие открытую форму ответов, в том числе на естественном языке. CONTROL может *игнорировать несущественные ошибки в словах ответа, учитывать сокращения и синонимы, разрешить, запретить или ограничить перестановки слов*; и т.д. В следующем разделе приведены примеры, иллюстрирующие это.

Примеры оценок системой CONTROL тестовых заданий с «открытыми» ответами

Следующие далее примеры тестовых задач относятся к различным учебным дисциплинам средней школы и вуза. Они представлены на различных этнических языках, во-первых, с целью показать применимость упомянутых выше методов к этим языкам, и во-вторых, поскольку часть вопросов и ответов взята из учебников и сборников тестов, я счел неуместным переводить их.

Для каждого тестового вопроса я приведу несколько возможных правильных либо приемлемых ответов (буду их нумеровать G1, G2 и т.д. – от «Good»), и несколько неправильных ответов (нумерую их B1, B2 и т.д. – от «Bad»). Система CONTROL адекватно распознаёт эти ответы. В случаях, когда ответ является приемлемым, но не «отличным», ниже приводится оценка, выставленная за него.

Вопросы и возможные ответы на них набраны курсивом, мои комментарии – прямым шрифтом. Первые два вопроса взяты из школьного учебника по информатике [Жалдак і Морзе, 2000].

Вопрос1. *Від чого залежить набір послуг, які можна отримати, використовуючи комп'ютер?*

G1. *Від набору фізичних пристроїв, що приєднані до комп'ютера; а також від наявних на ньому програм та файлів даних (інформаційного забезпечення).* Это «эталонный» ответ из учебника [Жалдак і Морзе, 2000].

G2. Від програмного та апаратного забезпечення.

G3. Від техніки та тих програм, що є на цьому комп'ютері.

V1. Від програм, що є на цьому комп'ютері. (Предполагаем, как это имеет место в цитируемом учебнике, что в правильном ответе упоминание аппаратного обеспечения является обязательным).

V2. Від моделі комп'ютера.

Вопрос2. Що називають файлом?

G1. Запис (сукупність даних), розміщений на деякому носії, який може бути прочитаний чи переписаний на інший носій окремо (незалежно) від інших записів (ответ из учебника [Жалдак і Морзе, 2000]).

G2. Поіменовану купність даних. (Грамматически правильно было бы: «Поіменовану сукупність даних». В каждом слове ответа допущены ошибки. Но такой ответ на вопрос по информатике следует признать правильным).

V1. Дані на магнітному диску.

V2. Запис у каталозі.

Вопрос3. Чему равно $(a+b)^2$?

G1. $a^2+2ab+b^2$

G2. $b^2+aa+2ab$

V1. $(a+b)(a+b)$

V2. $(a+b)^2$.

Вопрос4. Чему равно отношение **длины** окружности к ее **диаметру**?

G1. Оно равно числу π

G2. Pi

G3. 3.14

V1. Система CONTROL сочтет неправильным любой ответ, не содержащий принятых обозначений числа π или его численного значения (с некоторой приемлемой точностью).

Вопрос5. Введите формулу, определяющую **объем** прямого кругового конуса с высотой **H** и радиусом основания **R**.

G1. CONTROL сочтет правильной любую формулу, эквивалентную $\pi R^2 H / 3$.

V1. Что угодно, не эквивалентное формуле $\pi R^2 H / 3$.

Вопрос6. What the RHOMBUS is? Please, give the verbal definition.

G1. Rhombus is a parallelogram with equal sides.

G2. It is an equilateral rectangle.

V1. A parallelogram with different sides.

Вопрос7. З якими історичними подіями періоду Київської Русі пов'язана дата: **988 рік**?

G1. Хрещення Русі.

G2. Крещение Руси. (Этот ответ дан на другом языке, чем вопрос. Разумеется, нельзя требовать от системы тестирования, чтобы она распознавала имеющие одинаковый смысл ответы, представленные на любых языках. Но в данном случае это возможно, благодаря близости русского и украинского).

G3. Принятие Киевской Русью христианской религии.

V1. Любой ответ, в котором не упомянуто христианство или Русь.

Вопрос8. Де і коли відбулась перша велика битва козаків під проводом Богдана Хмельницького з польським військом?

G1. Під Жовтими Водами 5-6 травня 1648 року. («Эталонный» ответ из сборника тестов по истории Украины; CONTROL оценивает его на «отлично»).

G2. Жовті Води, травень 1648. (CONTROL оценит этот ответ на «хорошо»).

G3. 1648, *Жовті Води*. («Удовлетворительно»).

B1. Під Корсунем у 1648 році.

Вопрос9. Введіть формулу для вичислення **производной** сложной функции $F(1/g(x))$ (используйте апостроф для обозначения производных: F' , g' и т.п.).

G1. Любая формула, эквивалентная формуле $F'(1/g(x)) g'(x)/g^2(x)$.

B1. Что угодно, не эквивалентное формуле $F'(1/g(x)) g'(x)/g^2(x)$.

Вопрос10. Введіть **общее уравнение прямой на плоскости** xOy .

G1. $Ax+By+C=0$

G2. $ux+vy+w=0$

G3. $C+yB= -Ax$

B1. $y=kx+b$

B2. $Ax+By=0$.

Вопрос11. Что такое указатель (**pointer**)? Дайте определение этого понятия.

G1. Переменная, значение которой — адрес области памяти.

G2. Это величина, равная адресу какой-то другой величины.

G3. Это тип данных, представляющих собой адреса.

B1. Это переменная, объявленная с помощью \wedge .

B2. Это адрес оперативной памяти.

Вопрос12. Какую задачу выполняет следующий фрагмент программы?

```

...
var S : string; k : byte;
begin
S:=Edit1.Text;
k:=Length(S);
while ( k>0 ) and ( S[1]=' ' ) do
begin DELETE( S, 1, 1); k:=k-1 end;
Edit1.Text:=S;
...

```

G1. Вычеркиваются все пробелы в начале строки.

G2. Удаление ведущих пробелов.

G3. Уничтожаем крайние левые пробелы в строке.

B1. Любой ответ, не содержащий упоминания об удалении пробелов, а также о расположении оных в строке.

Вопрос13. Запишіть формулу, що виражає закон де Моргана для кон'юнкції.

G1. $\neg(P\&Q) \leftrightarrow (\neg P \vee \neg Q)$

G2. $\neg(a\&b) \leftrightarrow \neg a \vee \neg b$

G3. $(\neg(x\&y) \rightarrow (\neg x \vee \neg y)) \& ((\neg x \vee \neg y) \rightarrow \neg(x\&y))$

Сравнивая эти правильные ответы, вы видите, что студент может произвольно выбирать переменные, входящие в искомую тавтологию. Ясно, что указание вроде «в ответе используйте переменные P и Q» снизило бы качество вопроса.

B1. $\neg(P\&Q) \leftrightarrow (\neg P \& \neg Q)$

B2. $\neg(P\&Q) \leftrightarrow \neg(P\&Q)$.

Вопрос14. Що називають тавтологією у логіці висловлювань? Дайте визначення і наведіть приклад тавтології.

G1. Формула, що має значення "істина" незалежно від значень істинності своїх підформул; наприклад $P \rightarrow P$. (Оценка CONTROL: «отлично»).

G2. Висловлювання, таблиця істинності якого містить лише одиниці; наприклад $P \rightarrow (P \vee Q)$. («Отлично»).

G3. Формула, таблиця істинності якої містить лише одиниці; наприклад $(P \vee Q) \rightarrow P$. («Удовлетворительно», формула не являється тавтологією).

G4. Формула, що завжди має значення "істина"; наприклад $P \rightarrow P$. («Удовлетворительно», определение неточное).

B1. Формула, що завжди має значення "істина"; наприклад $(P \vee Q) \rightarrow P$. («Неудовлетворительно»).

B2. Логічну помилку (повторення того самого), наприклад "Масло масляне". («Неудовлетворительно», ведь вопрос касался исчисления высказываний, где термин "тавтология" имеет иной смысл).

Вопрос15. Запишіть формули Виєта для уравнения:

$$ax^2+bx+c=0 \quad (a \neq 0)$$

и объясните **смысл** величин, входящих в эти формулы.

G1. Чтобы CONTROL оценила ответ на «отлично», он должен включать:

a) любые два равенства, эквивалентные равенствам

$$x_1+x_2 = -b/a \quad \text{и} \quad x_1x_2 = c/a;$$

b) указание, что x_1 и x_2 – корни исходного уравнения.

При этом вместо x_1 и x_2 разрешено использовать другие обозначения для корней: p и q , и.т.п.

G2. Ответ, в котором отсутствует пункт b, но верны оба равенства, CONTROL оценит как «хороший».

G3. Ответ, в котором присутствует пункт b, но неверно одно из равенств, будет оценен как удовлетворительный.

B1. Любые иные ответы будут сочтены неудовлетворительными.

О типологии «открытых» тестовых заданий и ответов на них

Приведенные выше примеры показывают, что для компьютерного тестирования оказываются допустимыми (в частности) вопросы следующих – общепринятых в педагогической практике по любой предметной дисциплине – типов:

1. «Свободно» сформулировать описание некоторого факта или ситуации (или же «свободно» дополнить, завершить предложенное описание).
2. Сформулировать определение некоторого понятия.
3. Дать объяснение некоторому факту или ситуации.
4. Привести примеры, иллюстрирующие некоторое понятие или ситуацию.
5. Привести контр-примеры к чему-либо.
6. Описать границы или условия, необходимые для чего-либо.
7. Исправить ошибки в предьявленном определении, описании или выражении.
8. Дать иную, равносильную формулировку предложенного в задании определения, утверждения, формулы.
9. Выразить символически некоторое утверждение, сформулированное на естественном языке (или, наоборот, дать словесное истолкование некоторой формулы).

Включение таких «открытых» тестовых заданий в число подлежащих автоматической проверке требует уточнения ряда понятий, связанных с компьютерным тестированием. Ниже обсуждаются некоторые из них.

Назовем темой вопроса множество тех понятий и связей между ними, которые должны быть упомянуты в правильном и полном ответе на данный вопрос.

Назовем тематической формой вопроса структуру, которая связывает понятия из темы вопроса в данной формулировке вопроса.

Используя эту терминологию, можно сказать, что приведенные выше типы 1 – 9 суть примеры некоторых видов *тематических форм* вопроса (поскольку ясно, что вопросы, различающиеся по виду 1 – 9, могут иметь одну и ту же тему вопроса).

Назовем формой изображения вопроса структуру, в которой видео- или аудио-изображения понятий и связей темы вопроса предъявляются учащемуся.

Традиционно применяемые «закрытые» формы тестовых заданий («выбрать верный ответ», «дополнить», «установить соответствие» и др.), как представляется, являются видами форм изображения вопроса.

Каждый опытный преподаватель, по-видимому, будет согласен с тем, что *выбор тематической формы вопроса ограничивает множество пригодных для неё форм изображения вопроса*. Иными словами, не любая форма изображения вопроса пригодна для заданной его тематической формы. Например, не просто совместить упомянутые выше тематические формы 1 – 9 с формой изображения «выбрать верный ответ» без снижения «качества» вопроса.

Назовем семантической формой предъявления ответа на вопрос структуру изображения понятий из темы вопроса в ответе, которую данный вопрос предписывает (в явном или неявном виде) учащемуся.

Имеется тесная связь между формой изображения вопроса и семантической формой предъявления ответа на него. В большинстве случаев учащийся узнает, *как именно ему формулировать ответ, именно из поставленного вопроса (или из контекста, его сопровождающего)*. Поэтому тематическая форма вопроса оказывает – через «навязываемое» ею ограниченное множество форм изображения вопроса – значительное, если не решающее, влияние на семантическую форму предъявления ответа на данный вопрос.

В частности, вопрос, поставленный в какой-нибудь из перечисленных выше тематических форм 1 – 9, предполагает соответствующую семантическую форму предъявления ответа. Так, ответом на вопрос формы 4 *не должно быть определение* понятия (если, конечно, оно не сопровождается *примером*); а для вопроса формы 2 все обстоит наоборот.

Придерживаясь традиционной терминологии, следовало бы сказать, что вопросы видов 1 – 9 предполагают *открытую* форму предъявления ответа, поскольку не предписывают учащемуся ограничений на его формулировку. Однако при этом, очевидно, *не предполагается*, что учащийся свободен в выборе *семантической* формы предъявления своего ответа. То, что имеется здесь в виду под «открытостью», означает свободу выбора учащимся средств аудио- или видео-изображений понятий, предполагаемых темой вопроса – но не выбора самих понятий и той их взаимосвязи, которая «навязана» тематической формой вопроса.

Назовем синтаксической формой предъявления ответа на вопрос избранные учащимся аудио- или видео-средства изображения, а также структуру взаимосвязи аудио- или видео-элементов в предъявленном учащимся ответе.

Таким образом, в данной терминологии следует сказать, что тематические формы вопроса 1 – 9 предполагают *свободную синтаксическую форму* ответа на них.

Систему компьютерного тестирования знаний «Control» можно рассматривать как «библиотеку», предоставляющую преподавателю *средства оценки ответов* предъявленных в разнообразных *семантических* формах (причем эта библиотека может пополняться новыми типами указанных форм). Каждой семантической форме предъявления ответа соответствует так называемое *решающее правило*

системы «Control», которое оценивает ответы учащегося, предъявленные в указанной форме [Баранов и Леоненко, 2004].

При этом большинство решающих правил *не накладывает никаких ограничений на синтаксические формы* предъявления соответствующих ответов (или стремится свести такие ограничения к минимуму).

Заключение

За несколько лет применения системы «Control» в Одесской национальной академии связи им. А.С.Попова разработаны банки задач по курсам "Информатика", "Аналитическая геометрия на плоскости", "Дифференциальное исчисление", "Численные методы решения инженерных задач", "Теория линейных электрических цепей", "Математическая логика".

Были выполнены статистические оценки эффективности системы CONTROL. В эксперименте участвовали все студенты первого курса Академии (более 500 человек). Половина сдавала регулярные «зачеты» системе CONTROL, а половина писала в то же время письменные контрольные работы. Задачи тестирования в обеих группах совпадали.

Эксперимент показал [Баранов, 2004], что итоговые уровни качества знаний студентов в обеих группах статистически совпали. Это позволяет утверждать, что компьютерное тестирование может использоваться в учебном процессе с не меньшей эффективностью, чем традиционные методы контроля знаний.

Библиография

- [Leonenko, 2002] L. Leonenko. Analogical inferences in computer assisted knowledge testing systems // 6th Multi-Conference on Systemics, Cybernetics and Informatics (SCI 2002), Orlando, Florida, USA, 2002. Proceedings, Vol. XVIII, p.371-376.
- [Баранов, 2004] В. Ю. Баранов. Комп'ютерне тестування з інформатики: підсумки педагогічного експерименту в Одеській національній академії зв'язку // Теорія та методика навчання математики, фізики, інформатики. – Вип. 4. – Т. 3 – Кр.Ріг.: НМетАУ, 2004. – с. 6 –12.
- [Баранов и Леоненко, 2004] В. Ю. Баранов, Л. Л. Леоненко. Компьютерное тестирование знаний: об алгоритмах оценки «свободных» ответов. // «ИНФОТЕХ-2004». Материалы междунаrod. науч.-практич. конф. Киев-Севастополь: НТО РЭС Украины, 2004, с. 163 – 171.
- [Жалдак і Морзе, 2000] М. І. Жалдак, . Н. В Морзе . Інформатика-7. Експериментальний навчальний посібник для учнів 7 класу загальноосвітньої школи. – К.: ДіаСофт, 2000. – 208 с.
- [Леоненко, 2007] Л. Л. Леоненко. Алгоритмы оценки аналогичности текстов и их применение в компьютерном тестировании // Сб. трудов VII международной конф. "Интеллектуальный анализ информации". – К.: Просвіта, 2007, с. 210–220.
- [Леоненко, 2009] Л. Л. Леоненко. Оценивание «открытых» ответов в системе компьютерного тестирования знаний CONTROL // Сб. трудов IX международной конф. "Интеллектуальный анализ информации". – К.: Просвіта, 2009, с. 240–246.
- [Леоненко и Поддубный, 1996] Л.Л. Леоненко, Г.В. Поддубный. Теория подобия конечных последовательностей и ее приложения к распознаванию образов // Автоматика и телемеханика, 1996, № 8, с.119-131.
- *[Уемов, 1971] А.И. Уемов. Логические основы метода моделирования. М.: Мысль, 1971. – 311 с.

Информация об авторе



Леонид Леоненко – Одесская национальная академия связи им. А. С. Попова, доцент, ул. Конная, 22, кв. 6, Одесса, 65029, Украина; e-mail: Leonid.Leonenko@gmail.com
Основные области научной деятельности: неклассическая логика, общая теория систем, компьютерное тестирование знаний