МЕТОДОЛОГИЯ ИСПОЛЬЗОВАНИЯ ЭКСПЕРТНЫХ СИСТЕМ ДЛЯ ОЦЕНКИ УРОВНЯ ЗНАНИЙ

Г.Д. Когай, Э.Г. Кесарева, Г.М. Яворская

Нынешнее состояние контроля знаний представляет собой элективное соединение старого с новым, субъективного с объективным, ручного труда с компьютером… Возросшие затраты непроизводительного педагогического труда на проверку знаний студентов стали тормозом в развитии высшего образования. В частности, к недостаткам современных форм Контроля хода обучения следует отнести: использование грубой четырехбальной шкалы оценок; большие затраты труда и времени на устные опросы (на коллоквиумах, зачетах, экзаменах), проверку письменных контрольных и расчетно-графических работ; низкую оперативность в использовании результатов контроля для управления ходом учебного процесса; абсолютно неудовлетворительную объективность оценки знаний студентов, невозможность сопоставления оценок, полученных ими у разных преподавателей или тем более в разных ВУЗах.

Экспертная система оценки уровня знаний является одной из важнейших компонент автоматизированного обучения. Естественно, система оценки уровня знаний «судит» о степени усвояемости материала путем задания вопросов и сопоставления ответов со значениями, заданными в ее базе знаний. Рассмотрим, какие основные виды вопросов существуют.

Вопросы как обучающее воздействие можно различать как показано на рисунке 1.

Каждая из представленных на рисунке 1 линий различения, естественно, может быть продолжена. Зарубежные специалисты проблемами классификации вопросов занимались достаточно много. Среди предложенных ими классификаций следует назвать под-ход, при котором вопросы разделяются по двум уровням, в зависимости от задействованных в ответе когнитивных операций. К низкому уровню относят вопросы на репродукцию, к высокому – вопросы на выявление процесса, отношений, связей, на применение, догадку, требующие осуществить синтез, высказать мнение.

К. Холл [1] применил указанную спецификацию вопросов к компьютерному обучению. Вопросы он распределил в зависимости от психологического содержания обучения, выделив обучение на основе ассоциации, а также обучение по изучению понятий и решению задач, выдвигающее определенные требования к ответам учащихся. Интересную классификацию вопросов по типу задаваемой в них мыслительной активности предложили Б. Смит и М. Мюкс [1]. Они выделяют вопросы на: определение, описание, обозначение, обоснование, сообщение, подстановку, оценку, высказывание мнения, классификацию, сравнение и противопоставление, осуществление вывода, объяснение, руководство и управление классом. Проанализировав какие именно вопросы обычно ставят учителя, авторы пришли к выводу, что большинство предпочитают вопросы на описание, общую характеристику и объяснение. Схема, предложенная Смитом и Мюксом, будучи в общем полезной, требует уточнений. Например, представляется целесообразным включить вопросы на определение функций объекта, на отнесения понятия к роду, на определение системы существенных признаков понятия, на определение структуры объекта, на принятие решения, на осуществление контроля за правильностью решения, на обоснование своего рассуждения, на определение типа задачи, на отличие задачи от решавшейся ранее и т. д. [1]

При проверке знаний студента с помощью линейных тестов, оценку ответа студента на вопрос теста можно рассматривать как результат эксперимента, который ставит педагог [12]. Результатом такого эксперимента является некоторая величина (или совокупность величин) Z, зависящая от состояния знаний студента β. Если бы зависимость $Z = f(\beta)$ была детерминированной (то есть функциональной), то вынесение решения об оценке знаний не вызывало бы затруднений, определив по педагогическому опыту вид этой зависимости f и измерив Z, можно было бы точно определить неизвестное состояние знаний $\beta = f^{-1}(Z)$. Затем осталось бы выбрать решение, касающееся дальнейшего обучения для известного теперь уровня знаний β, с тем чтобы с наименьшими затратами достичь цели обучения.

На практике все сложнее. Как правило, зависимость результатов тестирования от состояния знаний является *статистической*, то есть в принципе при любом состоянии знаний могут наблюдаться любые значения Z (неверный ответ может случайно дать студент отлично знающий материал, и наоборот, полностью правильный ответ может случайно выбрать студент абсолютно не знающий материала) и только вероятности этих значений p(Z/ β_j) меняются в зависимости от β_j.

Набор условных распределений

$$p(Z / \beta_j),\ j = \overline{1,n}. \qquad (1)$$

может быть оценен исходя из предыдущего опыта и может уточняться по результатам использования системы. Задача состоит в том, чтобы на основании наблюдаемого в эксперименте значения Z определить оценку α. Решение о той или иной оценке называется *статистическим*, так как оно основано на наблюдении случайной величины Z.

Пусть, студент может находится в двух состояниях: обладать знаниями по исследуемому предмету и не обладать таковыми знаниями. Обозначим эти состояния, соответственно β_1, β_2. Педагог принимает решение по оценке знаний студента. Например, пусть оценка производится по четырехбальной шкале: "отлично", "хорошо", "удовлетворительно", "неудовлетворительно", обозначим эти оценки соответственно: $\alpha_1, \alpha_2, \alpha_3, \alpha_4$.

В таблице 1 приведена матрица потерь принятия того или иного решения в зависимости от возможных состояний студента, которая может быть построена с использованием понятия субъективной вероятности из теории полезности [2].

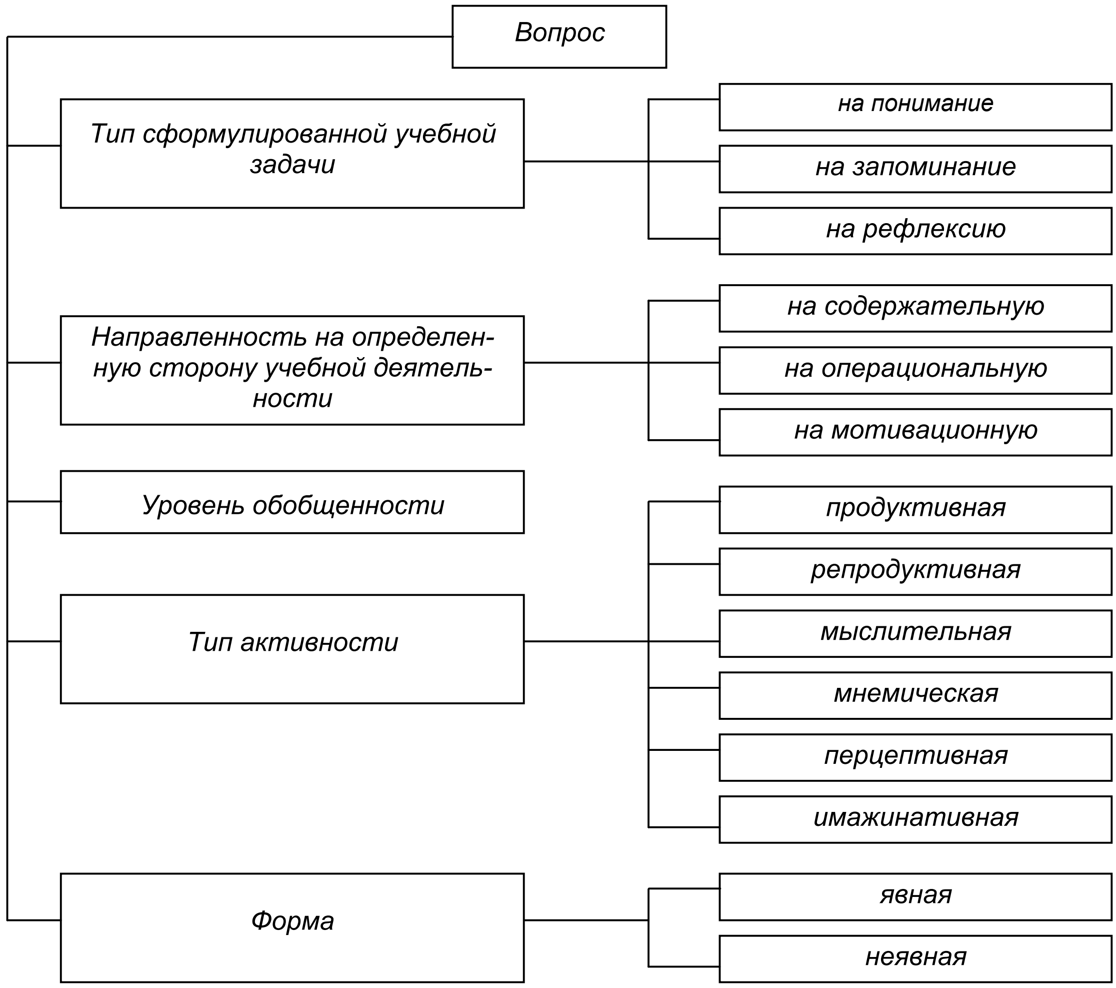

Рис.1. Виды вопросов как обучающих воздействий

Таблица 1

Матрица потерь принятия того или иного решения в зависимости от возможных состояний студента

	β_1	β_2
α_1	0	1
α_2	0,1	0,6
α_3	0,8	0,4
α_4	1	0

Допустим педагог, прежде чем поставить оценку, обращается к своему педагогическому опыту и статистическим данным (статистическому прогнозу) относительно ответов на вопросы теста и качества знаний. Для простоты будем считать, что число возможных ответов на вопросы теста (множество ситуаций в статистическом прогнозе) равно трем:

Z1 – “абсолютно правильный ответ”;

Z2 – “есть понимание, но ответ не верен”;

Z3 – “абсолютно неправильный ответ”.

Из предыдущего опыта субъекту известно, что статистическому прогнозу соответствуют следующие значения условных вероятностей (таблица 2).

Таблица 2

Значения условных вероятностей

$p(Z/\beta j)$	β_1	β_2
Z_1 (правильно)	0,7	0,1
Z_2 (неверно)	0,2	0,5
Z_3 (неправильно)	0,1	0,4

Данные таблицы 2 означают, что если студент знает материал, то с вероятностью 0,7 выберет правильный ответ теста, с вероятностью 0,2 – неверный ответ и с вероятностью 0,1 – неправильный; в случае отсутствия знаний варианты выбора ответов соответственно – 0,1; 0,5; 0,4. Конечно, результаты испытаний дают не идеальные результаты, но бывают и худшие варианты предсказаний. Во всяком случае, можно считать, что данные испытаний дают разумные измерения, и несут информацию о состоянии студента.

Задача педагога состоит в том, чтобы выбрать конкретное решение в зависимости от наблюдаемого исхода тестирования. Для этого он должен иметь какой-то алгоритм, правило, которое каждому наблюдаемому значению Z ставит в соответствие некоторое решение. Этот алгоритм в случае очного обучения и взаимодействия студента и педагога формируется в процессе обучения, практических занятий и характеризует профессиональные качества педагога. В случае дистанционного обучения, когда результаты тестирования являются единственным способом получения информации о знаниях студента, использования экспертных систем оценки знаний, необходима автоматизация алгоритмов принятия решений об оценке:

$$\alpha = \delta(Z). \qquad (2)$$

С математической точки зрения алгоритм принятия решения равносилен заданию функции которая отображает множество возможных ответов на множество решений по оценке знаний. Эта функция называется *статистической решающей функцией* [3,4]. Число различных решающих функций M в случае конечных множеств Z поддается подсчету. Если число значений величины Z равно k, а число оценок равно m, то число различных решающих функций $M = m^k$.

В рассматриваемом примере m – число возможных оценок равно четырем, а число ответов на вопрос теста - k предполагалось равным 3, тогда число возможных статистических решающих функций - M равно 4^3=64. Восемь первых и восемь последних функций приведены в таблице 3 в лексикографическом порядке.

Таблица 3

Статистические решающие функции

	δ_1	δ_2	δ_3	δ_4	δ_5	…	δ_{58}	δ_{59}	δ_{60}	δ_{61}	δ_{62}	δ_{63}	δ_{64}
1	α_1	α_1	α_1	α_1	α_1	…	α_4	α_4	α_4	α_4	α_4	α_4	α_4
2	α_1	α_1	α_1	α_1	α_2	…	α_3	α_3	α_3	α_4	α_4	α_4	α_4
3	α_1	α_2	α_3	α_4	α_1	…	α_2	α_3	α_4	α_1	α_2	α_3	α_4

В таблице приведены вообще все возможные решающие функции, а не только разумные. Например, δ_{57} отражает образ действий субъекта, который решил поступать наперекор данным тестовых испытаний и если ответ абсолютно верный ставит оценку "неудовлетворительно", а если ответ был абсолютно неверным ставит оценку "отлично".

Очевидно, что если из всего множества решающих функций выбрать какую-нибудь конкретную и остановиться на ней, то участие педагога в выборе решения уже не требуется, нужно механически подставить полученный результат эксперимента в функцию и получить готовое решение. При этом важно, чтобы одна решающая функция годилась на все случаи так как она определена на всех возможных значениях наблюдаемой величины. В этом и состоит основная задача теории статистических решений: эта теория помогает субъекту из всего множества возможных решающих функций выбрать одну, наилучшую в некотором смысле. Для этого необходимо ввести количественную меру качества решающих функций: определить критерий оптимальности, позволяющий сравнивать функции между собой. Пусть задана некоторая решающая функция $\delta i(Z)$. Если наблюдается значение величины Z_k, то функция вынесет решение $\alpha = \delta i(Z_k)$. В случае если студент находиться в состоянии βj, то данное решение приведет к потерям

$$l(\alpha, \beta_j) = l(\delta i(Z_k), \beta_j). \qquad (3)$$

В данном состоянии β_j студент может показать различные ответы Z_k с вероятностями $p(Z_k/\ \beta_j)$. Средние потери данной i-той решающей функции будут определяться по формуле

$$Lij=L(\delta i,\beta_j)=\sum_k p(Z_k/\beta_j)l(\delta i(Z_k),\beta_j), \qquad (4)$$

где суммирование ведется по всему множеству значений случайной величины Z.Число Lij показывает, насколько плохо (в среднем) "работает" решающая функция δi, если студент находится в состоянии βj. Таким образом, качество решающей функции не абсолютно, оно зависит еще и от состояния студента.

Таблица 4

Обобщенная матрица потерь

	β_1	β_2	…	β_n
δ_1	L_{11}	L_{12}		L_{1n}
δ_2	L L_{21}	L L_{22}		L_{2n}
…				
δ_m	L L_{m1}	L L_{m2}		L_{mn}

Можно построить таблицу, строки которой соответствуют решающим функциям, а столбцы – состояниям студента. Эта таблица называется обобщенной матрицей потерь (термин "обобщенная" указывает на то, что здесь мы имеем дело не с решениями, а с решающими функциями).

Имея обобщенную матрицу потерь, можно поставить задачу о выборе оптимальной решающей функции. Действительно, теперь мы имеем задачу о принятии решения, только под решением мы понимаем выбор конкретной решающей функции. Для выбора решений можно воспользоваться известными критериями, используемыми в теории полезности: минимаксным, Гурвица, Сэвиджа, Байеса и др. Можно, ввести понятие рандомизированной решающей функции как распределения вероятностей на множестве нерандомизированных решающих функций. Это позволяет эффективно решить проблему накопления базы знаний экспертной системы оценки обучения.

Например, для уже введенных условий нашего примера поставим задачу отыскания статистических решающих функций, оптимальных по минимаксному критерию. Прежде всего, необходимо построить обобщенную матрицу потерь. Это таблица с 64-мя строками - по числу решающих функций, и двумя столбцами - по числу состояний студента; ее элементы вычисляются по формуле (4). Затем необходимо пост-роить платежное множество на плоскости потерь, возникающих для двух состояний студента $\beta_1 и \beta_2$. Так как знания студента могут быть оценены как отличные, хорошие, удовлетворительные и неудовлетворительные, то есть количество состояний равно 4 (β_1, β_2, β_3 и β_4), а количество ответов на вопросы теста может быть сколь угодно большим, то вычисление числа возможных статистических решающих функций, а тем более оптимальной решающей функции становится практически невозможным.

Современные средства создания экспертных систем, а в их числе и редактор EXSYS, позволяют упростить процесс принятия решения о выставлении оценки. После задания каждого вопроса экспертная система в зависимости от ответа обучаемого присваивает каждой оценке заданный в ее базе знаний коэффициент вероятности и в соответствии с выбранным ответом задает следующий (более легкий или более сложный) вопрос. Таким образом достигается нелинейность теста, сокращение затрат времени и автоматизация принятия решения о выставлении той или иной оценки.

Экспертная система оценки уровня знаний должна моделировать умственную деятельность преподавателя-эксперта на экзамене, зачете, коллоквиуме или на любом другом виде контроля успеваемости. Опытный преподаватель не станет задавать каждому студенту одни и те же вопросы. Обычно преподаватель задает вопрос или задание средней сложности, так, чтобы можно было предварительно судить об уровне знаний студента. Если студент справился с заданием средней сложности, преподаватель задает вопрос более высокой степени сложности. Если студент справляется со сложными заданиями, преподаватель не станет задавать ему легкие задания, а после нескольких вопросов выставляет оценку. Если же студент не справился с заданием средней сложности, преподаватель задает ему более легкое задание в зависимости от того, в чем была ошибка студента. Если и с более легким заданием студент не справляется, после задания нескольких простейших вопросов, преподаватель с уверенностью может выставить неудовлетворительную оценку, не задавая слабому студенту сложных заданий. Проанализировав, таким образом, деятельность преподавателя-эксперта при оценке знаний студентов,, была разработана модель этой деятельности для использования ее в экспертной системе оценки уровня знаний.

Согласно схеме, экспертная система оценки уровня знаний, моделируя деятельность преподавателя, задает вопрос достаточно высокой сложности, чтобы определить первоначальные степени вероятностей выставления оценок.

Затем, в соответствии с выбранным ответом, система задает следующий вопрос:

- если студент справился с заданием, то система присваивает высокие значения коэффициентам уверенности отличной и хорошей оценок, низкие коэффициенты уверенности – удовлетворительной и неудовлетворительной оценкам и задает более сложный вопрос;

- если студент не справился с заданием, то система присваивает низкие коэффициенты уверенности высоким оценкам и высокие коэффициенты уверенности – низким оценкам и задает более легкий вопрос в соответствии со связями, указанными в ее базе знаний.

Далее система обрабатывает ответы на последующие вопросы, присваивает соответствующие коэффициенты уверенности оценкам.

Сложность представляет задать коэффициенты уверенности различным оценкам в соответствии с ответом студента. Преподаватель-эксперт выставляет эти коэффициенты, полагаясь на свой опыт. То есть, если студент справился с определенным заданием, преподаватель может сказать, что с уверенностью, например, 0,8 студент отлично знает материал, с уверенностью 0,7- студент хорошо знает материал, с уверенностью 0,4 – студент удовлетворительно знает материал, и с вероятностью 0,2 – студент не знает материал, а просто случайно или по подсказке выбрал правильный ответ.

СПИСОК ЛИТЕРАТУРЫ

1. Машбиц Е.И. Психолого-педагогические проблемы компьютеризации обучения. - М.: «Педагогика», 1988. – 192 С.

2. Гладких Б.А. Лекции по исследованию операций. – Томск: Издательство Томского университета, 1979.-120 С.

3. Сибаров Ю. Г. и др. Охрана труда в вычислительных центрах. - М.: «Машино-строе-ние», 1990.

4. Боброва - Голикова Л.П., Мальцева О.М., Коханова Н.А., Строкина А.Н. Эргономика и безопасность труда. - М.: «Машиностроение», 1985. – 250 С.