МАТЕМАТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ АВТОМАТИЧЕСКИХ СИСТЕМ УПРАВЛЕНИЯ ПРОФЕССИОНАЛЬНОЙ ОРИЕНТАЦИИ НА ОСНОВЕ МОДЕЛЕЙ НЕЧЕТКОЙ ЛОГИКИ

И. В. Самойло, канд. техн. наук
Первый Московский государственный медицинский университет им. И. М. Сеченова,
Москва, Россия

Предложена математическая модель профориентации обучаемых, в которой описаны все этапы профориентационного выбора — от ввода требований к студенту до поступления студента на кафедру. Для решения задачи разработки эффективных моделей профессиональной ориентации построена функция профессиональной ориентации студента, определены переменные этой функции. На основании представленной модели разработан алгоритм и создана автоматизированная система поддержки принятия решений для профессиональной ориентации студентов (абитуриентов) при поступлении на кафедру в вузы.

Ключевые слова: математическая модель, функция профессиональной ориентации, методы логического вывода, нечеткие методы, алгоритм, автоматизированная система управления.

Необходимость профессиональной ориентации личности требует проведения комплекса мер по содействию самоопределения личности, результативность которого отвечает как социально-экономическим целям личности, так и всего общества в целом. Применение математических методов в области профессионального выбора открывает возможности для более глубокого проникновения в суть и закономерности исследуемого процесса, более точного профориентационного прогнозирования, а значит, открывает возможности для эффективного управления процессом профессионального выбора.

Одними из методов построения моделей управления профессиональной ориентацией могут служить методы логического вывода, принятые в дедуктивных рассуждениях, с оценкой полученных результатов в рамках нечетких методов. В частности, возможно применение принципа резолюции (гиперрезолюции) в условиях неопределенности с оценкой степени истинности или ложности полученного результата. В данной модели любой элемент, входящий в состав формального описания конкретной задачи, должен иметь экспертную оценку истинности (достоверности) или степени принадлежности некоторому значению лингвистической переменной, характеризующей этот элемент. Область практического применения данной модели — создание автоматизированных экспертных систем профессионального выбора, которые должны обеспечить оптимальную профессиональную ориентацию при переходе между уровнями образования.

При построении и выборе математических моделей профессиональной ориентации можно выделить следующие основные этапы:

- Выбор и смысловое определение переменных и групп переменных, входящих в модели.
- Шкалирование или задание возможных значений этих переменных на некотором числовом множестве.
- Непосредственное построение функций профессионального выбора с помощью определения логических операций над переменными и группами переменных.

Полученные результаты могут быть использованы для создания экспертных систем, позволяющих лицу, принимающему решение (ЛПР), давать рекомендации по профессиональной ориентации и выбору студенту или абитуриенту.

Определим компоненты, из которых складывается профессиональная подготовка к тому или иному виду деятельности. Для этого сформулируем следующую задачу:

Пусть имеется учащийся, окончивший общеобразовательную среднюю школу, перед которым стоит задача выбора будущей профессии (факультета, кафедры и т. п.). В данном случае автор не рассматривает ни материальное положение учащегося, ни близость места его жительства к возможному месту учебы.

Аналогичную задачу можно сформулировать и для выбора профиля обучения при получении среднего образования (пойти учиться в физико-математический класс или в гуманитарный и т. д.).

Определим переменные и группы переменных, которые нужно использовать для построения функции, описывающей данное состояние обучаемого.

Группа переменных (O) — оценки. В общем случае для группы переменных O можно записать $O = \{O_1, O_2, O_3, ..., O_n\}$. Одними из переменных, которые более или менее адекватно могут описывать состояние обучаемого на данный момент времени, являются его учебные достижения, выраженные в оценках по отдельным предметам. Сюда также необходимо отнести и победы на предметных олимпиадах.

$$O_1 = \{x_1, x_2, x_3, ..., x_i\}$$
$$O_2 = \{x_{i+1}, x_{i+2}, x_{i+3}, ..., x_j\}$$
$$O_3 = \{x_{j+1}, x_{j+2}, x_{j+3}, ..., x_k\}$$
$$...$$
$$O_n = \{x_{l+1}, x_{l+2}, x_{l+3}, ..., x_m\},$$

где каждая из x-переменных принимает на заданном множестве одно из возможных целочисленных значений.

Группа переменных (C) — психологические тесты, направленные на выявление способностей, связанных с обучением и интеллектом.

$$C = \{y_1, y_2, ..., y_k\},$$

где y_1, y_2, ..., y_k — такие же переменные функции в моделях профессиональной ориентации, как и переменные x, и могут принимать на заданном множестве точно такие же числовые значения, и для их обозначения нужно было бы использовать символы x_{m+1}, x_{m+2}, x_{m+3}, ..., x_p.

Однако для того чтобы подчеркнуть, что они относятся к другой группе переменных, будем использовать для их обозначения символы y_1, y_2, ...,y_k.

Группа переменных (L) — характеристики личности учащегося.

Для определения личностных характеристик учащихся можно использовать различные психологические тесты. Например, можно выбрать оксфордский тест способностей личности.

$$L = \{z_1, z_2, z_3, z_4, z_5, z_6\},$$

где z_1 — самообладание;
z_2 — уверенность;
z_3 — активность;
z_4 — энергичность;
z_5 — чувство ответственности;
z_6 — правильность оценки реальности.

Или использовать многофакторный тест Р. Кеттела, тогда:

$$L = \{z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8, z_9, z_{10}\},$$

где z_1 — интеллект;
z_2 — эмоциональная устойчивость;
z_3 — доминантность;
z_4 — совестливость;
z_5 — смелость;
z_6 — мягкосердечность;
z_7 — мечтательность;
z_8 — радикализм;
z_9 — самодостаточность;
z_{10} — самоконтроль.

После выбора и смыслового определения переменных и групп переменных, входящих в функции профессионального отбора, и перед вводом логических операций над этими переменными необходимо осуществить их шкалирование и задать множество числовых значений, которые они могут принимать.

Решение подразумевает различные трактовки. Можно использовать линейные шкалы различной величины (5-балльную, 20-балльную, 100-балльную и т. д.) или нелинейные (например логарифмические). Следует отметить, что в любом случае значения, которые будут принимать переменные, должны округляться до целого.

Возникает вопрос, насколько правомерным является сравнение и логические операции над такими разнородными переменными, как, например, учебные оценки и результаты психологических тестов, и могут ли они при определенных условиях иметь равные права при осуществлении выбора.

Не вызывает сомнений, что любой человек является целой совокупностью различных неделимых качеств. Между качествами личности нет четкой детерминированности, и любое свойство человека проявляется через множество различных взаимоинтегрированных качеств. Поэтому одной из задач данной работы является построение модели профессиональной ориентации и отбора не по суммарному набору компетенций в отдельных областях (или по сумме показателей компетенций), а на основе логических взаимосвязей и сравнения различных показателей не в линейной или аддитивной, а в сенергетической модели. При этом качества объекта могут не находиться в отношении доминирования или даже иметь трудно сравнимый характер (т. е. быть принадлежащими множеству Эджворта–Парето).

Значения функций, которые необходимо построить в данной модели, будут давать не дифференциальную, а интегральную оценку компетенции и таким образом учитывать внутреннее взаимное влияние различных показателей, что невозможно осуществить при использовании линейных моделей.

Шкалирование переменных. В российской системе образования общепринятой является пятибалльная система оценок (на самом деле она является 4-балльной: 2 — неудовлетворительно, 3 — удовлетворительно, 4 — хорошо, 5 — отлично). Некоторые преподаватели для увеличения уровней градации вводят дополнительные оценки: три с плюсом, четыре с минусом и т. д. Причем оценка "три с плюсом" является более низкой, чем "четыре с минусом".

Надо отметить, что увеличение уровней градации оценок с каждым годом находит все большее и большее применение, примером чего является, например, введение в школах ЕГЭ (единый государственный экзамен) со 100-балльной шкалой.

Примерное соотношение между 100-балльной шкалой и традиционной системой школьных оценок представлено на рис. 1.

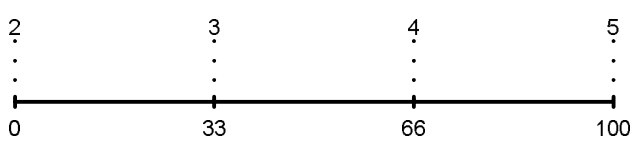

Рис. 1. Одно из возможных соотношений между традиционной системой школьных оценок и 100-балльной шкалой

В 100-балльной шкале результаты победителей предметных олимпиад всероссийского уровня могут иметь следующие значения: первое место — 100 баллов, второе место — 97 баллов, третье место — 94 балла. Для городского уровня: первое место — 90 баллов, второе место — 88 баллов, третье место — 85 баллов. Школьная оценка "отлично" будет соответствовать также 85 баллам по верхней границе и 80 баллам по нижней границе. Оценка "хорошо" — от 51 балла до 79 баллов, оценка "удовлетворительно" — от 34 до 50 баллов и ниже 33 баллов — оценка "неудовлетворительно".

При описании шкалирования величин переменных, относящихся к группам C и L, начнем с переменной, показывающей результаты IQ-тестов.

Как правило, результаты IQ-тестов представляют в числовом диапазоне от 80 до 150 баллов и для их сопоставления и логических операций с другими показателями необходимо все переменные привести к единой шкале. На рис. 2 представлен один из возможных путей такого перевода.

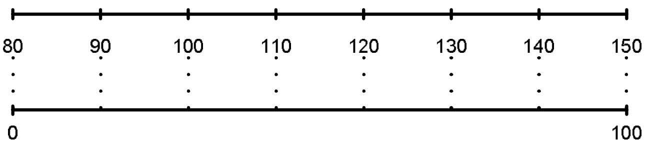

Рис. 2. Один из возможных способов перевода традиционных результатов IQ-тестов в 100-балльную шкалу

Общая формула для расчета результатов IQ-тестов y_5^*, приведенных к 100-балльной шкале, будет иметь вид:

$$y_5^* = \frac{y_5 - y_{5,\min}}{y_{5,\max} - y_{5,\min}} \cdot 100,$$

где $y_{5,\min}$ — нижняя граница баллов IQ-теста;
$y_{5,\max}$ — верхняя граница баллов IQ-теста;
y_5 — балл испытуемого, набранный в традиционной для IQ-теста шкале.

Отметим, что после расчета y_5^* округляется до целого значения.

Аналогичным образом можно поступить и с результатами остальных психологических тестов, если они не представлены в 100-балльной шкале.

В некоторых случаях, например, при использовании для моделирования профессиональной ориентации методов логического вывода, принятых в дедуктивных рассуждениях, с оценкой полученных результатов в рамках нечетких методов, где достоверность вывода определяется значениями истинности, необходимо использовать замкнутый интервал [0, 1]. Соответственно, необходимо проводить и шкалирование переменных в пределах данного интервала.

Сложность применения хорошо разработанного в аксиоматических системах методов вывода в правдоподобных рассуждениях, в том числе и в нечетких системах, определяется природой нечетких систем:

- в нечетких системах степень истинности может принимать бесконечное число значений из интервала [0, 1];
- в нечетких системах мы имеем дело со спектром лингвистических значений высказываний, которые принадлежат нечеткому множеству.

Значения лингвистических переменных могут быть сформированы с помощью модификаторов, которые придают переменной различный смысл. Между тем, каждое из значений лингвистической переменной образует нечеткое множество степеней принадлежности на ограниченном множестве значений независимой переменной, принадлежащей универсуму. Например, высказывания о результатах сдачи экзамена на "отлично", "хорошо", "удовлетворительно" могут иметь степени истинности (степени принадлежности лингвистическому значению "отличная оценка") 0,95; 0,73; 0,45; 0,3; 0,1. С другой стороны, каждое из этих значений лингвистической переменной характеризуется функцией принадлежности.

Пусть высказывание P = "успешно сданный экзамен" имеет степень истинности 0,95. Мы можем представить это высказывание как предикат с конкретизированным значением оценки, соответствующим понятию "отличная оценка". Отрицание высказывания $P(x)$ есть?$P(x)$ представляет подмножество значений лингвистической переменной "оценка" за исключением значения "отличная оценка".

Высказывание $P(x)|_{x=?i}$ можно интерпретировать как предикатную форму, определенную на множестве X, $X = [?_{\min}, ?_{\max}]$ — интервале действительных чисел. Если принять значение степени истинности высказывания $?_{P(x)|x=95} = 0{,}95$, где P = "отличная оценка", x — численное значение оценки, соответствующее экспертному представлению об отличной оценке, то степень истинности отрицания высказывания $P(x)$ есть:

$$?_{?P(x)} = 1 - ?_{P(x)} = 0{,}05.$$

Для успешного резолютивного вывода на множестве дизъюнктов необходимо наличие в двух дизъюнктах контрарных литералов P и $?P$. Тогда степень истинности противоречивого высказывания

$$??P??P? = \min(?_P, ?_{?P}) = \min(0{,}95 - 0{,}05) = 0{,}05.$$

Примем в качестве нечетких логических операций И и ИЛИ для высказываний P и Q:

$$?(P?Q) = \min(?_P, ?_Q),\ ?(P?Q) = \max(?_P, ?_Q).$$

Если степень истинности высказывания P обозначить $??P?$, а степень достоверности этого высказывания $cd(P)$ − (certainty degree), зависимость $cd(P)$ от $??P?$ может иметь вид, представленный на рис. 3.

Значение неопределенности 0,5 является границей между истинной (true) и ложной (false) областями.

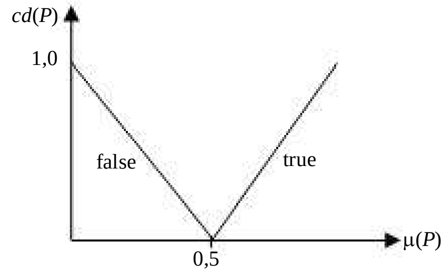

Рис. 3. Зависимость степени истинности высказывания P (??P?) от достоверности этого высказывания cd(P)

Отношение степени доверия $cd(P)$ некоторого высказывания на множестве степеней истинности $?(P(x = ?_i))$ в простейшем случае имеет линейную форму. Иные формы могут соответствовать более тонким требованиям, определенным условиями решаемой задачи.

Из приведенного на рис. 3 отношения следует, что все высказывания, степень истинности которых превышает 0,5, следует считать истинными. Это означает, что при нечетком резолютивном выводе в число истинных могут попасть и другие конкретизации.

В зависимости от требований конкретной задачи точка неопределенности может трансформироваться в область неопределенности. Например, если степень истинности не может оказаться ниже некоторого определенного экспертом значения $?_{inf}?P? > 0,5$, то область неопределенности соответствует интервалу $[1-?_{inf}, ?_{inf}]$ (рис. 4).

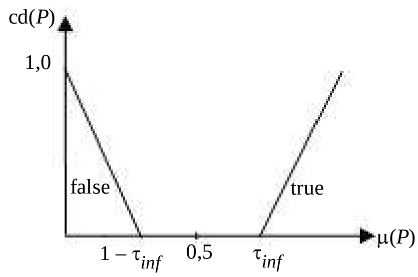

Рис. 4. Область неопределенности для функции степени доверия cd(P)

Можно сузить область решения, ограничив нечеткое множество значений лингвистической переменной только антонимными.

Заметим, что значение достоверности резольвент зависит от исходных значений степеней истинности высказываний. Так, при степени истинности высказывания $P_{?i}$, равной $\mu(P_{?i}) = 0,75$, степень достоверности резольвенты принимает значение 0,5, т. е. достигает степени индифферентности. Поэтому при решении задач, связанных с нечетким выводом, по крайней мере, следует задать некоторую нижнюю границу $\tau_{inf}(P)$ степени истинности. Можно сузить область решения, ограничив нечеткое множество значений лингвистической переменной только антонимными.

Пусть в данной интерпретации I степень противоречивости $P_{?i}$ $??P_{?i}$ есть и $?(P_{?i})$? [0,1]:

$$?(P_{?i}) = |?(P_{?i}) - ?(?P_{?i})|.$$

Определение. Степенью достоверности нечеткой резольвенты $R(C_1, C_2)$, где $C_1 = P_{?i}$? W, $C_2 = ?P_{?i}?U$ — дизъюнкты из S, есть степень противоречивости $?(P_{?i})$ высказывания $P_{?i}$.

Степень соответствия функции профессиональной ориентации данного студента требованиям кафедры оценивается при определении степени достоверности резольвент на множестве дизъюнктов.

Пусть имеется три кандидата: K_1, K_2, K_3 на поступление на кафедру "АБВ" (или в некоторое учебное заведение, например, университет), переменные функции профессионального выбора для каждого из них известны. Преподавателями кафедры (сотрудниками университета) Х было решено, что для поступления студенту (абитуриенту) необходимо иметь "хорошие" оценки по физике, математике, информатике, а также "хорошую" оценку по тестированию способностей к обучению.

Оптимальным методом вычисления степени истинности является экспертный опрос. Допустим, что хорошие оценки, это значение больше или равно 4,1 (для разных вузов или кафедр может варьироваться). Тогда для нахождения степени истинности можно взять статистику за последние два года — 10 и 11 классы. Причем, статистику можно брать и за более длительный срок. Для примера подойдут оценки четырех четвертей в 11 классе — 4, 4, 5, 5. Тогда степень истинности будет равна

$$\mu = (4 + 4 + 5 + 5)/\ (4{,}1\times 4).$$

Если степень истинности больше единицы, то принимаем ее равной 1. Для того чтобы найти степень истинности результатов тестирования, направленного на выявление способностей, связанных с обучением и интеллектом, или тестирования психологических особенностей личности, необходимо задать некоторую границу, при которой будем считать в качестве делителя — Граница, тогда μ = Значение/Граница.

Составим нечеткую гиперрезолюцию:

1. ¬хорошие_оценки_математика(x) ∨¬ хорошие_оценки_физика (x) ∨¬ хорошие_оценки_информатика(x) ∨¬ хорошие_оценки_тестирование_умственных_способностей (x) ∨¬ поступление_на_кафедру_АБВ(x).

2. хорошие_оценки_математика (K_1) $\mu_2 = 0,75$ степень истинности факта 2.

3. хорошие_оценки_математика (K_2) μ_3=0,9

4. хорошие_оценки_математика (K_3) μ_4=0,95

5. хорошие_оценки_физика (K_1) μ_5=0,85

6. хорошие_оценки_физика (K_2) μ_6=0,8

7. хорошие_оценки_физика (K_3) μ_7=0,75
8. хорошие_оценки_информатика (K_1) μ_8=0,9
9. хорошие_оценки_информатика (K_2) μ_9=0,65
10. хорошие_оценки_информатика (K_3) μ_{10}=0,75
11. хорошие_оценки_тестирования_способностей к обучению (K_1) μ_{11}=0,95
12. хорошие_оценки_тестирования_способностей к обучению (K_2) μ_{12}=1,0
13. хорошие_оценки_тестирования_способностей к обучению (K_3) μ_{13}=0,6

Решение.

14. поступление_на_кафедру_АБВ (выбранную студентом или абитуриентом K_1 специальность): из дизъюнктов 1, 2, 5, 8, 11 степень достоверности резольвенты 0,5 получена следующим образом:

из 1 и 2 $\zeta_1 = |0,75-0,25| = 0,5$

из 1 и 5 $\zeta_2 = |0,85-0,15| = 0,7$

из 1 и 8 $\zeta_3 = |0,9-0,1| = 0,8$

из 1 и 11 $\zeta_3 = |0,95-0,05| = 0,9$

$cd = \mu_{14} = \min(0,5;\ 0,7;\ 0,8;\ 0,9) = 0,5$

15. поступление_на_кафедру_АБВ (выбранную студентом или абитуриентом K_2 специальность):

$cd = \mu_{15} = 0,8$ (взято примерно, не вычислялось как в пункте 11)

16. поступление_на_кафедру_АБВ для K_3:

$cd = \mu_{16} = 0,9$ (взято примерно, не вычислялось как в пункте 11).

Таким образом, в соответствии с принятой экспертной оценкой качества студентов или абитуриентов (оценка была не менее 0,7) по двум признакам выдержавшими конкурс признаются два кандидата: K_2 (0,8), K_3 (0,9).

Если принять граничное значение степени истинности $\tau_{inf} = 0,85$, то конкурс выдерживает лишь один кандидат K_3, для которого степень достоверности $\mu_{16} = 0,9 > \tau_{inf} = 0,8$. Введение нижней границы степени достоверности сужает область решений, повышает требования к выполнению принимаемого решения.

На основании вышепредставленной модели в рамках нечетких методов разработан алгоритм работы автоматизированной экспертной системы поддержки решений при профессиональной ориентации. Реализован работающий прототип такой системы, сформирован четкий механизм управления профвыбором (кафедральным выбором) студентов:

1. Студент-первокурсник (абитуриент) заходит на стартовую страницу системы, заносит школьные оценки и (или) вносит результаты единого государственного экзамена, система проводит оценку достоверности результата с помощью нечеткой логики.

2. Студент-первокурсник (абитуриент) проходит тестирование психологических особенностей личности и способности к обучению с оценкой достоверности результата с помощью нечеткой логики.

3. Автоматизированная экспертная система (АЭС) проверяет соответствие данного студента требованиям кафедры (учебного заведения). Если "да", то с помощью управляющей образовательной среды (УОС) направляется дальнейшее развитие студента. Если "нет", то с помощью УОС корректируются знания студента, создаются оптимальные условия преодоления кафедрального "барьера", кроме того, у студента есть возможность отказаться от борьбы за интересующую его кафедру и продолжить образование на той кафедре, на которой позволяют его достижения.

4. Последующие тестирования проходят для второкурсников в конце учебного года в момент выбора кафедры и для третьекурсников — по результатам осенней сессии с целью подтверждения выбора либо отказа от него.

Выводы

1. Для решения задачи разработки эффективных моделей профессиональной ориентации построена функция профессиональной ориентации студента (абитуриента), определены переменные этой функции.

2. Для определения степени соответствия функции профессиональной ориентации требованиям кафедры (образовательного учреждения) применены методы логического вывода, принятые в дедуктивных рассуждениях, с оценкой полученных результатов в рамках нечетких методов, т. е. вне двоичной логики. В рамках данной модели любая переменная функции профессиональной ориентации должна иметь экспертную оценку истинности (достоверности) или степени принадлежности некоторому значению лингвистической переменной, характеризующей эту переменную.

3. На основании математического аппарата методов логического вывода с оценкой полученных результатов в рамках нечетких методов разработан алгоритм работы автоматизированной экспертной системы поддержки решений при профессиональной ориентации.

Литература

1. Борисов В. В., Круглов В. В., Федулов А. С. Нечеткие модели и сети. — М: Горячая линия — Телеком, 2007. — 284 с.

2. *Ашинянц Р. А.* Логические методы в искусственном интеллекте. — М.: МГАПИ, 2001. — 224 с.

3. *Сумкин К. С., Морозова Т. Ю.* Об использовании нечетких множеств для разграничения прав доступа информационной сети// Наукоемкие технологии, 2008. Т. 9. № 7. С. 12—14.

4. Искусственный интеллект: Справочник. В 3-х кн. Кн. 2. Модели и методы//Под ред. Д. А. Поспелова. — М.: Радио и связь, 1990. — 304 с.

5. *Аверкин А. Н.* Нечеткие множества в моделях управления и искусственного интеллекта // Под ред. Д. А. Поспелова. — М.: Наука. 1986. — 312 с.

SOFTWARE OF AUTOMATED SYSTEMS OF MANAGEMENT OF VOCATIONAL GUIDANCE BASED ON FUZZY LOGIC MODEL

I. V. Samoylo
The First Moscow State Medical University named after I. M. Sechenov, Moscow, Russia

***It is suggested a mathematical model of career counseling of students, which describes all stages of career-oriented choices — from entry requirements to the student till the student's joining the department. For the solution of the problem of the developing effective models of professional orientation, variables of vocational guidance of the learner are defined. There are used methods of inference, adopted in deductive reasoning with an estimate of the extent to which the functions of vocational guidance are correspond to the requirements of the department within the framework of fuzzy methods. Based on the presented model the algorithm was worked out and automated decision support system for professional orientation of students (applicants) when joining the department
(the higher education institutions) was created.***

Keywords: mathematical model, the function of career guidance, methods of inference, fuzzy methods, the algorithm for automated learning management systems.

Самойло Ирина Владимировна, доцент.
E-mail: samoilo@yandex.ru

♦