

ГИБРИДНАЯ ЭКСПЕРТНАЯ СИСТЕМА ДЛЯ ПРОФОРИЕНТАЦИИ

А.В. ГАВРИЛОВ*

В статье рассматривается архитектура гибридной экспертной системы с двумя видами интерпретации правил-продукций и с возможностью запуска внешних программ, взаимодействующих с базой знаний, на примере экспертной системы для профориентации.

1. ВВЕДЕНИЕ

Как известно, разработка экспертных систем носит итерационный характер и обычно рассматривается как процесс, состоящий из следующих этапов [1]: идентификация, концептуализация, формализация, реализация, тестирование, перестроение прототипной версии.

Итерационный характер создания экспертных систем обусловлен отсутствием устоявшихся технологий и методик их разработки, хотя попытки их формулирования известны [1, 2]. Предложенные методики носят весьма поверхностный характер и оказывают помощь только при создании сравнительно простых экспертных систем и инструментария для их построения (к ним относятся, как правило, продукционные экспертные системы с использованием поверхностных знаний). Особенности методологические и технологические сложности возникают в процессе разработки гибридных экспертных систем, т.е. таких, в которых используются разные методы представления и обработки знаний, а также средства обработки данных, рассматриваемые обычно вне рамок инженерии знаний.

В данной статье описана архитектура такого рода экспертной системы на примере системы профориентации для применения в центре занятости населения.

Система предназначена для решения двух основных задач:

- 1) выбора подходящей профессии;
- 2) диагностики пригодности к выбранной профессии.

Эти две задачи решаются в диалоге с клиентом-посетителем центра занятости населения. Но в дальнейшем клиент имеет возможность обра-

* Доцент кафедры вычислительной техники, канд. техн. наук

таться к психологу для разъяснений, если диалог с системой по каким-либо причинам его не удовлетворил.

Решение первой задачи предполагается для клиента, не имеющего определенного представления о том, какой профессией он может или хочет заниматься. Таким клиентом может быть, например, выпускник школы или человек, потерявший работу, и который готов переквалифицироваться на другой вид деятельности.

Вторая задача решается в случае, если клиент выбрал для себя одну или несколько профессий и хочет получить оценку своих возможностей при попытке заняться ими.

2. СТРУКТУРА ЭКСПЕРТНОЙ СИСТЕМЫ

Экспертная система для профориентации представляет собой комплекс из четырех основных программ, программ для тестирования психофизиологических характеристик, набора модулей базы знаний, а также набора сервисных программ-утилит.

В состав основных программ входят следующие:

- программа PROF, ориентированная на диалог с клиентом и предназначенная для экспресс-консультирования его по вопросам профориентации и профдиагностики;
- программа PROFIN отличается от PROF отсутствием пользовательского интерфейса и может запускаться с передачей параметров из какой-либо другой программы, в частности, из банка данных о профессиях;
- программа PROFEXOR, ориентированная на специалиста по профориентации, позволяет просматривать и анализировать базы знаний, используемые в PROF и PROFIN, и пополнять их;
- программа EXTRACT, ориентированная на инженера по знаниям (достаточно квалифицированного разработчика баз знаний), более универсальна чем PROFEXOR, позволяет более глубоко анализировать базы знаний, содержащиеся в произвольных файлах, запускать их для тестирования и имеет более мощные средства редактирования баз знаний чем PROFEXOR.

Программа EXTRACT является, по существу, универсальной оболочкой экспертных систем, а PROFEXOR - специализированной оболочкой для повышения удобства использования в рамках системы PROFEX.

В состав программ-утилит входят:

- программа сравнения содержимого базы знаний и базы данных о профессиях COMPKD;

- программа манипулирования модулями, из которых состоит база знаний EDMODKB.

Программы тестирования психо-физиологических характеристик запускаются из базы знаний и включают в себя следующие тесты:

- оперативной памяти;
- концентрации и устойчивости внимания, переключения внимания;
- объема и распределения внимания;
- зрительно-моторной координации движений;
- умения читать чертежи, схемы.

База знаний содержит два главных модуля, соответственно для решения задач профдиагностики и выбора профессии. Первый из них содержит правила:

- для тестирования типов профессий, подходящих для клиента, в основу которого положен дифференциально-диагностический опросник (ДДО);
- для тестирования некоторых психологических характеристик, в основу которого положен тест Кэттела.

Опросник ДДО в процессе диалога определяет следующие пять типов профессий: “человек - человек” (например, преподаватель, тренер), “человек - знаковая система” (например, программист, бухгалтер), “человек - художественный образ” (например, музыкант, художник, маляр), “человек - техника” (например, токарь, шофер, диспетчер) и “человек-природа” (например, егерь, ветеринар).

Тест Кэттела в процессе диалога определяет следующие психологические характеристики личности: переоценивание себя, адекватность самооценки, сообразительность, конкретно-логический тип мышления, общительность, эмоциональная уравновешенность, самоуверенность, решительность, положительная установка к деятельности, импульсивность, подвижность, ответственность, принципиальность, аккуратность, социальная смелость, социальная активность, мягкость, готовность помочь, артистичность, женственность, эгоцентризм, мнительность, ревнивость, самостоятельное творческое решение, развитое воображение, проницательность, дипломатичность, депрессивность, впечатлительность, ранимость, тревожность, гибкость, адаптивность поведения, склонность к критике, независимость в суждениях, самоконтроль, целенаправленность, исполнительность, повышенная мотивация, напряженность, беспокойство, тактичность, безропотность, зависимость, застенчивость, абстрактно-образный тип мышления, утомляемость, раздражительность, подверженность чувств-

вам, ипохондрия, замкнутость, холодность, осторожность, пессимизм, беспринципность, непостоянство, неорганизованность, суровость, жесткость, мужественность, практичность, сдержанность, робость, откровенность, доверчивость, благожелательность, исполнительность, наивность, прямолинейность, хладнокровие, безмятежность, спокойствие, консервативность, склонность к морализации, конформизм, недисциплинированность, вялость, низкая мотивация, невозмутимость.

Второй главный модуль содержит правила:

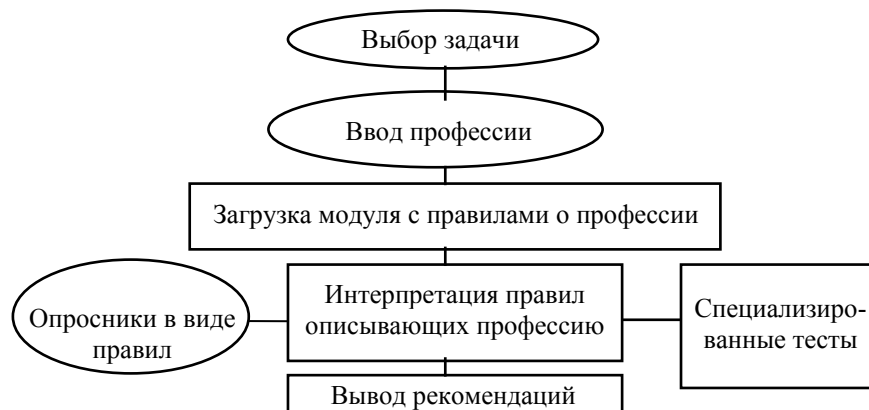
- реализующие тест ДДО, как и в первом модуле;
- реализующие тест Кэттела, как и в первом модуле;
- реализующие тест "Карта интересов".

В каждой из этих главных модулей есть ссылки на другие модули, содержащие знания о профессиях в виде правил-продукций.

Правила для профдиагностики по назначению делятся на три вида - для определения:

- пригодности к данной профессии;
- причины непригодности;
- родственников, рекомендуемых для дальнейшего тестирования профессий.

Структурная схема решения задачи при профдиагностике показана на рисунке.



Работа ЭС при решении задачи профдиагностики

Правила для выбора профессии используют рекомендуемые виды деятельности, а также классификацию профессий не только по типам, но и по классам, отделам.

3. ПРЕДСТАВЛЕНИ ЗНАНИЙ

В системе PROFEX используется представление знаний в виде правил-продукций с коэффициентами достоверности в процентах для представления нечеткости. Элементарные условия так же как, и заключения, сопровождаются коэффициентом достоверности. Правило имеет вид:

ПРАВИЛО номер правила :

ЕСЛИ

элементарное условие 1,

.....

элементарное условие n ,

ТО

заключение 1 КД = число

.....

заключение n КД = число.

Причем правила делятся на два вида в зависимости от формы интерпретации левой части:

- как конъюнкции элементарных условий с возможностью "отсеивания правила по порогу достоверности";
- как набора признаков с весовыми коэффициентами для вычисления взвешенной суммы их достоверностей для присвоения ее значения коэффициенту достоверности результата.

"Отсеивание правила по порогу достоверности" означает, что коэффициент при элементарном условии задает нижнюю границу достоверности факта, при которой еще имеет смысл далее интерпретировать данное правило. Если факт, соответствующий данному элементарному условию, имеет достоверность ниже пороговой, интерпретация правила прекращается (правило отсеивается).

Примеры правил такого вида для оценки пригодности к профессии "брокер" приведены ниже:

ПРАВИЛО 1: ЕСЛИ

Профессия=брокер КД=100% (Какую профессию Вы выбираете?) И

Тип=Человек-человек КД=60% (Какой тип профессии подходит для Вас?) И

Тип = Человек - знаковая система КД=50% (Какой тип профессии подходит

для Вас?) И

Концентрация и устойчивость внимания КД=90% (Можете ли Вы сконцентрировать и длительно удерживать свое внимание на чем-либо?) И

Объем и распределение внимания КД=90% (Вы способны легко следить сразу за несколькими объектами?) И

Гибкость, адаптивность поведения КД=90% (Вы хорошо приспосабливаетесь к изменяющимся условиям и поведению окружающих ?) И

Переключение внимания КД=90% (Можете ли Вы быстро переключать свое внимание?) И

Долговременная память КД=95% (У Вас хорошая долговременная память?) И

Ответственность КД=90% (Вы ответственный человек?) И

Общительность КД=90% (Вы общительный человек?) И

Хорошая дикция КД=90% (Вы обладаете хорошей дикцией?) И

Дизъюнкт=/*мед. противопоказания для брокера КД=100%:

органов дыхания или

сердечно-сосудистой системы. И

Оперативная память КД=90% (У Вас хорошая оперативная память?) И

Эмоциональная уравновешенность КД=80% (Вы эмоционально устойчивы?)

ТО

Можно=да КД=100 И

Отказ=нет КД=100

Пример правила для определения причины непригодности к профессии "брокер":

ПРАВИЛО 2: ЕСЛИ

Профессия=брокер КД=100% (Какую профессию Вы выбираете?)

И

Тип=/=Человек-человек КД=40% (Какой тип профессии подходит для Вас?)

ТО

Можно=нет КД=100. И

Отказ=Вам не подходит тип профессии "человек-человек" КД=100.

Использование второго вида интерпретации правил позволяет естественным образом вставлять в базу знаний обычно применяемые психологами опросники такие, как ДДО (дифференциально - диагностический опросник), тест Кэттела для оценки психических характеристик и т.п., основанных на подсчете баллов при ответе на вопросы.

Пример правила второго вида интерпретации из опросника ДДО :

ПРАВИЛО 41: ЕСЛИ

2d=Составлять таблицы, схемы, программы для вычислительных машин КД=100% (Из двух возможностей что бы Вы предпочли?) И

5d=Обсуждать научно-популярные книги, статьи КД=100% (Из двух возможностей что бы Вы предпочли?) И

9d=Искать и исправлять ошибки в текстах, таблицах, рисунках КД=100% (Из двух возможностей что бы Вы предпочли?) И

10d=Выполнять вычисления, расчеты КД=100% (Из двух возможностей что бы Вы предпочли?) И

12d=Разбираться в чертежах, схемах, таблицах (проверять, уточнять) КД=100% (Из двух возможностей что бы Вы предпочли?) И

15d=Составлять точные описания-отчёты о наблюдаемых явлениях КД=100% (Из двух возможностей что бы Вы предпочли?) И

19d=Заниматься черчением, копировать чертежи, карты КД=100% (Из двух возможностей что бы Вы предпочли?) И

20d=Работать на клавишных машинах (пишущей машинке, теле-тайпе) КД=100% (Из двух возможностей что бы Вы предпочли?)

ТО

Тип=Человек - знаковая система КД=100.

Здесь идентификаторы типа “20d” — условные обозначения объектов, представляющих собой ответы на соответствующий вопрос. Так, например, объект “20d” имеет кроме указанного в правиле второе возможное значение “вести борьбу с болезнями растений, вредителями леса, сада”.

Другой пример правила из теста Кэттела:

ПРАВИЛО 35: ЕСЛИ

15=нет КД=200% (Когда я планирую что-нибудь, я предпочитаю делать это

самостоятельно, без чьей-либо помощи?) И

15=иногда КД=100% (Когда я планирую что-нибудь, я предпочитаю делать это самостоятельно, без чьей-либо помощи?) И
 32=да КД=200% (Мне нравятся общительные, компанейские люди?) И
 32=не знаю КД=100% (Мне нравятся общительные, компанейские люди?) И
 49=возглавляя группу в туристском лагере КД=200% (Я бы скорее предпочел провести два летних месяца?) И
 49=затрудняюсь сказать КД=100% (Я бы скорее предпочел провести два летних месяца?) И
 66=неверно КД=200% (Я избегаю общественной работы и связанной с этим ответственностью?) И
 66=иногда КД=100% (Я избегаю общественной работы и связанной с этим ответственностью?) И
 83=да КД=200% (Я получаю большое удовольствие, рассказывая местные новости?) И
 83=иногда КД=100% (Я получаю большое удовольствие, рассказывая местные новости?) И
 100=где надо играть в команде или иметь партнера КД=200% (Я предпочитаю игры?) И
 100=не знаю КД=100% (Я предпочитаю игры?)
 ТО
 Конформизм КД=100.

Здесь в отличие от предыдущего примера результатом работы правила является не значение объекта, а атомарное утверждение “Конформизм”, имеющее смысл признака или психофизиологической характеристики личности.

В условиях правила числами обозначены также объекты-ответы на вопросы, так, например, объект ”100” может иметь одно из трех значений :

Я предпочитаю игры (это — вопрос)

- 1) где каждый играет за себя;
- 2) не знаю;
- 3) где надо играть в команде или иметь партнера.

Повторение двух разных значений в правиле означает, что каждое из них прибавляет веса результирующему утверждению.

Комбинация второго вида интерпретации правил и отсеивания правил по порогу достоверности может рассматриваться как моделирование в данной архитектуре синтаксическими средствами правил-продукций искусственной нейронной сети [3].

Факты могут быть двух видов:

- объект, значение, коэффициент достоверности;
- утверждение, коэффициент достоверности.

Экспертная система может работать со строковыми и численными значениями объектов. В качестве элементарных условий в левой части правила могут использоваться следующие отношения:

>, <, = для численных значений объектов;

=, != для текстовых значений объектов;

“истинно”, “ложно” для утверждений, “истинно” кодируется самим утверждением, а “ложно” его отрицанием с частицей “не” перед ним.

Для интерпретации правил используется обратный логический вывод.

Во внутреннем представлении правила и факты выглядят как предикаты пролога. Правило кодируется предикатом

$rule(N, List1, List2)$,

где N — номер (идентификатор) правила;

List1 — список элементарных условий;

List2 — список заключений.

Элементарное условие и заключение кодируется структурой $X(Object, Value, KD)$ или $Y(Sentence, KD)$,

где X, Y — двух- или трехбуквенное обозначение типа условия или заключения;

Object — название объекта;

Value — значение объекта;

Sentence — атомарное утверждение;

KD — коэффициент достоверности.

Факт кодируется предикатом вида $fact(X(Object, Value, KD))$ (факт-объект) или $fact(Y(Sentence, KD))$ (факт-утверждение).

Кроме этих основных предикатов база знаний во внутреннем представлении содержит предикаты для связи утверждений с внешними программами-тестами, запускаемыми, если факта-утверждения или его отрицания нет еще в базе данных $test(Sentence, NameProg)$, предикаты для связи главного модуля с модулями базы знаний, описывающими конкрет-

ные профессии $\text{mod}(\text{NameProf}, \text{NameMod})$, предикаты для описания возможных значений объектов $\text{obj}(\text{NameObj}, \text{ListValue})$ и вопросов, которые система задает пользователю при определении значения объекта $\text{quest}(\text{NameObj}, \text{Text})$ или при проверке утверждения, $\text{quest}(\text{Sentence}, \text{Text})$ и некоторые другие вспомогательные предикаты.

ЗАКЛЮЧЕНИЕ

Разработанные программы были опробованы в Городском центре занятости г. Новосибирска. Была создана экспериментальная база знаний примерно о двадцати профессиях. Работа с системой PROFEX показала ее работоспособность и эффективность принципов представления и обработки знаний, положенных в ее основу, а также и необходимость перевода системы на новую более современную платформу. Но, к сожалению, финансирование этих работ было приостановлено из-за отсутствия средств.

Программы системы PROFEX реализованы на языках PDC-Prolog и Turbo-Pascal 7.0 в среде MS DOS[♦].

Автор выражает благодарность Зайцеву М.Г., доценту кафедры ПСиБД НГТУ, который написал для системы программы-тесты психофизиологических характеристик, Петренко А.Ю., написавшему утилиту СОМПКД, руководству Новосибирского Центра занятости и фирмы “Сибирские компьютерные системы” за организацию финансирования этой разработки, психологам и профконсультантам Новосибирского центра занятости и фирмы “ГРАНТ”, в беседах с которыми рождались требования к архитектуре описанной системы, профессору Хабарову В.И., зав. кафедрой ПСиБД НГТУ, автору экспертной оболочки “Оракул”, совместимость с которой по данным мы пытались сначала соблюсти при проектировании системы PROFEX, а также студентам, которые принимали участие в разработке и обкатке системы, в частности, программы EXTRACT.

[1] *Построение экспертных систем* / Под ред. Ф.Хейес-Рота, Д.Уотермана, Д.Лената. — М.: Мир, 1987.

♦ По вопросам сотрудничества обращаться по адресу:
630111 Новосибирск 111, а/я 325, Гаврилову А.В.
e-mail: avg@osilab.cs.nstu.ru
тел. (3832)-46-02-19 Гаврилов А.В.

[2] УОТЕРМЕН Д. *Руководство по экспертным системам.* — М.: Мир, 1989.

[3] GAVRILOV A.V. *The methods of the combination of Logic and Associative processes in the Expert Systems* // Int. Workshop on Soft Computing SC'96. Proceedings. — Kazan, 1996. — P. 84-86.