

П.М. Зацепин, В.Б. Авцинов, Д.Н. Трошкин

Реализация экспертных оболочек на основе продукционной системы представления знаний и семантической сети

В последние десятилетия человечество все больше интересуется искусственным интеллектом (ИИ). Одним из наиболее интересных направлений ИИ являются экспертные системы (ЭС). ЭС – программы для компьютера, которые могут воспроизводить процесс решения проблемы человеком-экспертом. Применение экспертных систем целесообразно в тех областях человеческой деятельности, где невозможно привести точные алгоритмы решений, но можно построить систему эвристических правил (правил, основанных на практическом опыте человека-эксперта). К этим областям относятся: проектирование, планирование, постановка диагноза, перевод, реферирование, ревизия, выдача рекомендаций и т.п. Наиболее универсальными по своей структуре являются экспертные оболочки (ЭО) – это ЭС, из которых удалены базы знаний, но реализован механизм их заполнения. ЭО позволяют экспертам сосредоточиться собственно на заполнении баз знаний, что существенно сокращает время, необходимое на построение ЭС.

Существует несколько способов построения баз знаний для ЭС: системы логической продукции, семантические сети (СС), фреймы [1;2]. В данный момент в регионе является актуальной задача реализации экспертных оболочек, пригодных для постановки диагнозов. Подходящими способами представления знаний для этой цели являются система логической продукции и семантические сети.

Семантическая сеть – это информационная модель предметной области, представляющая собой направленный граф, у которого вершинами являются объекты, а дугами – отношения между ними. Она состоит из структур трех типов: элементов, свойств и указателей. Элементом может являться любой факт, вопрос, заключение, слово, фрейм и т.д. Причем у каждого из них есть атрибут, с помощью которого система определяет, к какому классу принадлежит элемент. Свойством является структура, описывающая элемент. Свойство определяет семантику объекта с помощью указателей на другие объекты в соответствии с базой правил системы.

В системе логической продукции знания представляются в виде правил вида: если (условие), то (вывод). В них существенно упрощается реализация нечетких рассуждений, которые позволяют более точно отразить ситуацию, делают процесс мышления более похожим на человеческий. Для реализации нечетких рассуждений используется теория нечетных множеств – достаточно молодая область математики, появившаяся в 60-е годы нашего столетия. Определение нечетких подмножеств, операции с ними используются в работе согласно [3].

Экспертная оболочка с базой знаний в виде логической продукции реализована по следующему алгоритму: в правилах каждой пары атрибут-значение ставится в соответствие коэффициент принадлежности к выводу из этого правила. Пары атрибут-значение являются нечеткими подмножествами вывода из правила. Если в правило входят несколько атрибутов, связанных цепочкой "и", то необходимо вычислять алгебраическое произведение нечетких подмножеств, если атрибуты связаны цепочкой типа "или", то сумму. База знаний (правил) в системе содержится в цифровой форме. Это позволяет существенно экономить память, повысить быстродействие системы, уменьшить количество ошибок, связанных с символьным вводом (ошибки правописания, опечатки при вводе и т.п.). Под атрибуты отведен одномерный массив строк, под значения атрибутов – двумерный. В текстовом виде также хранятся диагнозы. Для экономии оперативной памяти целесообразна их динамическая подкачка. Принцип построения правил следующий: правила с цепочками типа "или" разбиваются на отдельные правила с одинаковым выводом, таким образом все правила имеют одинаковую структуру: ЕСЛИ ... И ... И ... , ТО Правила разбиваются на локальные и глобальные. Локальные служат для переопределения достоверности значений одних атрибутов через другие, глобальные – для связи атрибутов и диагнозов. Условные части правил (посылки) реализуются в виде массивов, сопоставляющих номера атрибутов,

входящих в правило, и номера их значений с коэффициентом достоверности. Заключение локальных правил представляются в виде массива, ставящего в соответствие номеру правила номера атрибута и его значения, которые являются выводом из правила. Заключение глобальных правил предоставляются в виде массива, ставящего в соответствие номеру правила номер диагноза, который следует из этого правила. Реализация прямой цепочки рассуждений в системе проведена следующим образом: пользователь вводит начальную ситуацию (номера атрибутов, их значений и коэффициенты достоверности, с которыми эти значения определены). Система в соответствии с введенной ситуацией находит подходящие правила и, согласно правилам алгебраических операций на нечетких подмножествах, переопределяет текущие коэффициенты достоверности для значений атрибутов или диагнозов, входящих в консеквент. Далее запоминается номер отработавшего правила, номер атрибута, послужившего отправной точкой; система новой отправной точкой считает атрибут, определенный из вывода отработавшего правила. Так система проходит по цепочке: ЕСЛИ А, то В; ЕСЛИ В, то С, ЕСЛИ ..., до тех пор, пока не перестанет находить правила, в условную часть которых входит вывод из предыдущего правила. Тогда система "вспоминает" предпоследний атрибут, номер отработавшего с ним правила и продолжает поиск подходящего правила, начиная со следующего и т.д. Вся процедура проделывается для каждого номера атрибута в начальной ситуации. Реализация обратной цепочки рассуждений в экспертной системе имеет 2 стратегии: поиск в ширину и поиск в глубину. Поиск в ширину предусматривает нахождение всех возможных начальных ситуаций, приведших к имеющемуся результату. Но (в системе с элементами нечеткой логики это особенно заметно) количество ситуаций возрастает катастрофически быстро, если только к диагнозу приводит более 1 правила или правила содержат более одной пары атрибут-значение. Более удобной является стратегия поиска в глубину, когда с каждым шагом назад система запрашивает, в каком направлении ей продолжать поиск. Эту стратегию легко реализовать следующим образом: пользователь вводит номер имеющегося диагноза, система находит, из каких правил он следует и выдает на экран эти правила. Определив

атрибут, который его интересует, пользователь вводит номер этого атрибута, система находит правила, выводом из которых является определение значения этого атрибута и т.д.

Экспертную оболочку с базой знаний в виде семантической сети можно реализовать по следующей схеме: семантическая сеть реализуется в стандартной форме представления графов — это матрица, у которой элементом является тип связи между двумя объектами. Для удобства назовем ее матрицей связей. Объекты представляют собой текстовое описание и атрибут, которые хранятся в базе данных. Правила состоят из ссылок на объекты и тоже являются равноправными объектами, только они содержат ссылки на объекты в базе правил. В ней каждая ссылка связывается следующими видами отношений:

1. Класс — элемент класса (IS-A).
2. Свойство — значение ("цвет" — "красный").
3. Связи типа "часть-целое" (PART-OF).
4. Логические отношения (и, или, не).
5. Пример элемента класса ("например").
6. Функциональные связи ("объект-свойство", "свойство-значение").
7. Атрибутивные связи (количественные, временные, пространственные).

Проблема поиска решения в базе знаний типа СС сводится к задаче поиска фрагмента сети, соответствующего некоторой подсети, представляющей вопрос. Для этих целей используется механизм дедукции. Пользователю в начале работы предоставляется возможность самому выбрать необходимый объект, с которого начинается процесс вывода. Далее, путем просмотра матрицы связей, система выявляет все объекты, у которых имеется интересующая нас связь. После пересмотра этих объектов "подходящие" записываются и с ними проделывается то же самое, что и с первым. Это продолжается до тех пор, пока не будет найден интересующий нас элемент или правило. Если процесс вывода пошел по ложному пути, то необходимо его прервать и выбрать еще несколько объектов для рассмотрения, что существенно сужает область поиска.

Среди отношений большое значение имеют иерархические структуры. Для примера рассмотрим две самые распространенные связи: IS-A и PART-OF. С помощью связи IS-A осуществляется наследование свойств.

Свойства высших типов автоматически переносятся на низшие. Например, "объект" IS-A "класс-объекта", при этом все признаки "класса" переносятся на "объект". Так можно избежать дублирования информации, представив "объект" как "класс объекта" + некоторые индивидуальные признаки "объекта". Для разбиения информации по уровням детализации служит связь PART-OF. Она позволяет представить так информацию в СС, что появляется возможность отвечать на вопросы типа "из чего состоит объект?". Логические отношения строятся при помощи операций булевой алгебры. Создание нового отношения сводится к внесению в базу правил способов его интерпретации. Естественно, определение правил должно происходить через уже имеющиеся в базе правил отношения с использованием процедур для работы с графом. Здесь перечислены самые важные процедуры, которые присущи подавляющему большинству экспертных систем с использованием СС:

1. Установление связи. (Служит для добавления новых элементов в процессе вывода решения).

2. Аннулирование связи. (Удаление, если обнаружена ошибка).

3. Подсчет или выявление вершин, соединенных дугой. (Этой процедурой поль-

зуется интерпретирующий механизм при ответе на вопросы типа "сколько?" или "из чего?". Например, при ответе на вопрос "из чего состоит А?", система ищет в СС объект А и в качестве ответа выводит на экран список вершин, соединенных дугой PART-OF).

4. Проверка наличия-отсутствия связи между заданными вершинами. (В данном случае это служит для проверки гипотезы или правила).

5. Определение типа вершины. (Для систем, которые ведут диалог с пользователем на естественном языке, существуют такие типы объектов, как "существительное", "глагол", "время").

6. Подсчет вершин, принадлежащих одному типу.

В экспертной оболочке для эффективной работы с семантической сетью необходимо создать язык представления знаний. В качестве средства для его реализации был выбран язык программирования Форт. Это позволило создать рабочий словарь, который будет пополняться в дальнейших разработках. Оболочка содержит редактор связей, интерпретирующий механизм (вывод решений) и блок для работы с файлами данных.

В настоящий момент ЭО реализованы и проходят апробирование.

Литература

1. Уотермен Д. Руководство по экспертным системам. М., 1989.
2. Таунсенд К., Фохт Д. Проектирование и программная реализация экспертных систем на персональных ЭВМ. М., 1990.
3. Кофман А. Введение в теорию нечетких множеств. М., 1982.