

УДК 681.3

ИНСТРУМЕНТАЛЬНЫЙ КОМПЛЕКС ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ СЕТЕВОЙ МОДЕЛИ ПРЕДМЕТНОЙ ОБЛАСТИ

В.П. Гладун, В.Ю. Величко

Институт кибернетики им. В.М. Глушкова НАН Украины

e-mail: aduis@rambler.ru

В основу функционирования разработанного программного инструментального комплекса положены методы решения задач выделения закономерностей на основе пирамидальных сетей, а также методы использования выделенных закономерностей для принятия решений. Теоретические основы алгоритма построения пирамидальной сети описаны во многих публикациях [1-4].

Рассмотрим кратко основные функции и состав программного комплекса. Основные функции состоят в:

- выделение закономерностей (знаний), которые характеризуют классы объектов (ситуаций);
- использование обнаруженных закономерностей с целью классификации объектов (ситуаций) и выбора управляющего решения соответственно классу.

Основные блоки комплекса и взаимосвязь между ними представлены на Рис. 1.

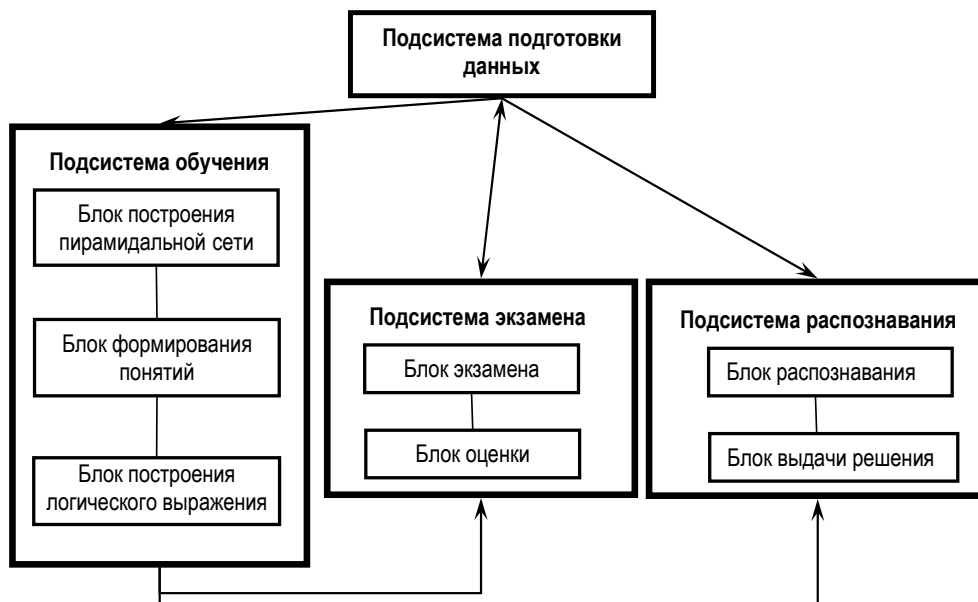


Рис. 1. Основные блоки комплекса и взаимосвязь между ними

Комплекс включает следующие четыре подсистемы.

1. **Подсистема подготовки данных** предназначена для формирования внутреннего представления признаков описаний ситуаций, используемых в качестве исходных данных в подсистемах обучения, экзамена и распознавания.

2. **Подсистема обучения** осуществляет выделение закономерностей (знаний), которые характеризуют классы ситуаций.

Входными данными для подсистемы обучения служит обучающая выборка, которая включает примеры ситуаций, характеризующихся различными типами (классами) ситуаций. Каждый класс ситуаций должен быть представлен некоторым количеством примеров, достаточным для того, чтобы на основании их анализа была выделена закономерность, характеризующая данный класс. Примеры ситуаций задаются признаковыми описаниями. Признаковое описание ситуации должно включать имя ситуации, класс, которому ситуация принадлежит, и набор значений признаков, которые

характеризуют ситуацию. Выходные данные для подсистемы обучения – закономерности, которые характеризуют классы ситуаций.

Процесс обучения состоит из следующих стадий:

- представление исходных данных в виде пирамидальной сети;
- формирование понятий на основе пирамидальной сети;
- построение логических выражений, соответствующих сформированным понятиям.

Блок построения пирамидальной сети реализует первую стадию процесса обучения, когда внутреннее представление объектов обучающей выборки преобразуется в пирамидальную сеть. Поскольку основным процессом при выделении закономерностей является поиск сочетаний значений признаков, характеризующих группы сходных объектов, как правило, приходится многократно просматривать объекты обучающей выборки, что с ростом числа объектов и числа значений признаков приводит к «информационному взрыву». Предоставление данных в виде пирамидальной сети позволяет избежать этой опасности, в первую очередь, за счет особенностей алгоритма построения сети, обеспечивающего выделение общих сочетаний значений признаков в процессе ввода объектов в сеть, а также за счет ассоциативности и иерархичности самой сети. При этом для полного построения сети необходимо только два просмотра обучающей выборки. Пирамидальная сеть является динамической структурой, которая перестраивается в зависимости от поступающей в нее информации.

Выделенные общие сочетания значений признаков отображаются в структуре сети. Сеть содержит узлы двух типов: рецепторы и концепторы. Рецепторы - это узлы, которые не имеют входных связей и соответствуют значениям признаков, характеризующих объекты обучающей выборки. Концепторы соответствуют сочетаниям значений признаков, которые являются общими для нескольких объектов, а также объектам в целом. Концепторы первого типа называются промежуточными узлами и имеют входные и выходные связи, концепторы второго типа - главными узлами, которые могут иметь выходные связи только в том случае, когда описание объекта, соответствующего данному главному узлу, является частью описания некоторого другого объекта.

Основной процесс построения пирамидальной сети, при котором в сети фиксируются общие сочетания значений признаков (строятся все промежуточные узлы), реализуется еще при первом просмотре описаний объектов обучающей выборки. При втором просмотре сеть только корректируется с тем, чтобы обеспечить однозначное представление каждого объекта в виде отдельной пирамиды (с одним главным узлом).

3. Подсистема экзамена предназначена для тестирования степени обученности комплекса.

Входными данными для подсистемы экзамена служит экзаменационная выборка, включающая примеры ситуаций, которые не вошли в обучающую выборку, но относительно которых известно, какому классу они принадлежат. Примеры ситуаций задаются признаковыми описаниями, включающими имя ситуации, класс, которому ситуация принадлежит, и набор значений признаков, которые характеризуют ситуацию.

Выходные данные подсистемы – ее ответ относительно класса, к которому отнесена каждая ситуация, и статистические данные о количестве правильных, неправильных и неопределенных ответов подсистемы. Полученные результаты сопоставляются с информацией о реальной принадлежности ситуации классу, на основании чего можно судить о том, насколько хорошо обучен комплекс.

4. Подсистема распознавания предназначена для идентификации класса, которому принадлежит новая ситуация и выдачи управляющего решения по нормализации ее.

Входными данными для подсистемы распознавания являются признаковые описания новых ситуаций, которые включают имя ситуации и набор значений признаков,

которые характеризуют ее. Распознавание выполняется на основании сопоставления описания новой ситуации и понятий, сформированных на этапе обучения. Выходные данные подсистемы – ее ответ относительно класса, к которому отнесена распознаваемая ситуация.

Блок выдачи решения, как и *блок оценки* результатов экзамена, может выдавать как точные, так и неопределенные ответы.

В случае неопределенного ответа подсистема выдает дополнительную информацию о том, насколько распознаваемая ситуация похожа на ситуации, соответствующие разным классам, или вообще не похожа на ситуации из обучающей выборки.

Для оценки степени схожести используется функция уверенности, которая рассчитывается на основе анализа конъюнкций, принимающих участие в распознавании данной ситуации. Функция уверенности отражает процентное соотношение информативности конъюнкций, описывающих закономерности, характеризующие разные классы ситуаций.

Выводы

Растущая пирамидальная сеть является сетевой памятью, самонастраивающейся на структуру входной информации. В самоструктурирующихся системах структура данных адаптируется к задаче (выделяются и определяются классы объектов), в результате чего достигается оптимизация решения. В отличие от сетей нейроразподобных элементов эффект адаптации достигается без введения априорной избыточности сети. В самоструктурирующихся системах удастся не только найти зависимости, обеспечивающие диагноз или прогноз, но и создать их логические описания.

Существенными особенностями разработанного инструментального программного комплекса поддержки принятия решений являются: независимость от проблемной области, что значительно упрощает процесс использования программного комплекса; обеспечение возможности эффективного использования комплекса непрофессиональными в области программирования пользователями.

Типичными прикладными задачами, для решения которых использовался разработанный инструментальный программный комплекс, являются: прогнозирование новых химических соединений и материалов с заданными свойствами, прогнозирование в генетике, геологии, прогнозирование солнечной активности, медицинская и техническая диагностика, планирование действий роботов, прогнозирование нарушений в работе сложных агрегатов и т.п. В качестве примера приведем данные о задачах прогнозирования существования неорганических соединений с заданными свойствами. В качестве обучающей выборки использовались таблицы, содержащие признаковые описания двойных, тройных и четверных систем химических элементов, образующих и не образующих химические соединения. Обучающие выборки для двойных, тройных и четверных систем включали соответственно 1333, 4278 и 4963 описания, а экзаменационные выборки – 692, 2156 и 2536 описаний. Каждый химический элемент описывался 87 признаками. Описания двойных, тройных и четверных систем состояли соответственно из 174, 261 и 348 признаков. Точность прогноза составила до 99%.

Литература

1. Гладун В.П. Планирование решений. - Киев: Наук. думка, 1987. – 168 с.
2. Гладун В.П. Процессы формирования новых знаний. – София: СД «Педагог 6».– 1994. – 189 с.
3. Гладун В.П. Партнерство с компьютером. Человеко-машинные целеустремленные системы. – Киев: Port-Royal. – 2000. – 118 с.
4. Victor Gladun, Vitaly Velichko, Yurii Ivaskiv. Selfstructurized Systems. International Journal "Information Theories & Applications". FOI ITHEA, Sofia. – Volume 15 – 2008, Number 1 – P. 5-13.