

УДК 004.9

А. О. Копырина, А. С. Копырин

Сочинский государственный университет, г. Сочи, email: lina1094@yandex.ru

КОНЦЕПЦИЯ ХРАНЕНИЯ ДАННЫХ В ЭКОНОМИЧЕСКОЙ ЭКСПЕРТНОЙ СИСТЕМЕ С ПРИМЕНЕНИЕМ ГИБРИДНЫХ ТЕХНОЛОГИЙ

Ключевые слова: экспертные системы, представление знаний, мивар

Задача анализа большого массива плохо структурированной экономической информации – выявление шаблонов и аномалий и их предметной интерпретации чрезвычайно важна во многих областях современной экономики. Для решения этого класса задач целесообразно использовать машинные знания и методы «мягких вычислений», которые помогут автоматизировать процесс выявления и интерпретации знания в большом диапазоне областей. Объектом исследования является процесс хранения данных и правил в экспертных системах. Предметом исследования является понятие и структура хранения правил вывода и обработки динамической информации. В работе представлен краткий обзор применяющихся в настоящее время моделей представления знаний в экспертных системах. Предложено использование гибридной модели с использованием технологии мивар для функционирования нечеткой системы вывода.

А. О. Копырина, А. С. Копырин

Sochi State University, Sochi, email: lina1094@yandex.ru

THE CONCEPT OF DATA STORAGE IN AN ECONOMIC EXPERT SYSTEM USING HYBRID TECHNOLOGIES

Keywords: expert systems, knowledge representation, mivar.

The task of analyzing a large array of poorly structured economic information-identifying patterns and anomalies and their subject interpretation is extremely important in many areas of modern economics. To solve this class of problems, it is advisable to use machine knowledge and “soft computing” methods that will help automate the process of identifying and interpreting knowledge in a large range of areas. The object of the study is the process of storing data and rules in expert systems. The subject of the study is the concept and structure of storing rules for the output and processing of dynamic information. The paper presents a brief overview of the currently used models of knowledge representation in expert systems. The use of a hybrid model using mivar technology for the operation of a fuzzy output system is proposed.

В настоящий момент существует потребность в применении современных IT-технологий в экономике в целом, так и в технологиях машинного обучения вместе с искусственным интеллектом в частности, особенно, когда это затрагивает создание систем поддержки принятия решений в широком спектре областей.

В начале 80-х годов в рамках искусственного интеллекта окончательно сформировалось новое направление – «инженерия знаний», цель которого – разработка и исследование экспертных систем. Отличительной особенностью экспертных систем, по сравнению с ранее существовавшими системами искусственного интеллекта, является то, что они способны решать не модельные, а вполне реальные проблемы, получая

при этом результаты, сопоставимые с теми, которые может получить эксперт.

Кроме того, экспертные системы используются в неформализованных или слабо формализованных областях знаний, которые ранее были мало доступны для возможностей вычислительных технологий. И, наконец, экспертные системы нацелены на общение с пользователями, уже не знакомыми с программированием, что позволяет резко расширить круг людей, способных решать свои прикладные задачи за счет использования компьютера на достаточно высоком профессиональном уровне. Все это вызвало многократный интерес к проблематике за последние десятилетия, как со стороны специалистов в области компьютерных технологий и ис-

кусственного интеллекта, так и со стороны пользователей.

Цель исследования

В современном информационном мире давно настал момент, когда стоимость информации заметно превысила стоимость структур для ее обработки и хранения. Таким образом, задача анализа большого массива плохо структурированной экономической информации – выявление шаблонов и аномалий и их предметной интерпретации чрезвычайно важна во многих областях современной экономики, например, при изучении логистики, а также при изучении показателей финансового рынка, управлении и оптимизации объемов производства с учетом энергосбережения и ресурсоемкости в жилищно-коммунальном хозяйстве. Для решения этого класса задач целесообразно использовать машинные знания и методы «мягких вычислений», которые вместе с улучшением специализированных экспертных систем и использованием эвристических алгоритмов помогут автоматизировать процесс выявления и интерпретации знания в большом диапазоне областей.

При решении этой задачи во многих случаях возникают трудности, связанные со сложностью применяемого математического аппарата и значительной сложностью последующей программной реализации разработанного алгоритмического ПО. Следовательно, совершенствование автоматизированных систем управления знаниями требует значительного программного обеспечения методов искусственного интеллекта. При этом в состав таких систем, как правило, входит база знаний, обеспечивающая реализацию разработанных алгоритмов анализа и обобщения регламентов и знаний.

Ранее авторами рассматривались проблемы разработки универсальной структуры базы знаний [1] и построения системы вывода, прежде всего, на основе базы знаний [2]. Данная статья посвящена описанию основных моделей представления данных в экспертных системах и их модификации для использования в экспертных системах с нечетким выводом.

Объектом исследования является процесс хранения данных и правил в экспертных системах. Предметом исследования является понятие и структура хранения правил вывода и обработки динамической информации.

Для достижения цели исследования необходимо решить ряд задач:

- изучить различные подходы к описанию баз знаний;
- предложить концептуальную модель хранения правил в нечеткой экспертной экономической системе.

Результаты исследования и их обсуждение

Общая структура экспертной системы содержит следующие три компонента: базу знаний, систему управления базой знаний, которая выполняет ввод, хранение, редактирование, добавление и преобразование информации во внутреннее представление, а также решатель (машины вывода) для обработки знания при принятии решений [3, 4].

Успешное цифровая трансформация бизнеса предприятий основана, в том числе, на внедрении интегрированных экспертных систем (ИЭС) с использованием современных интеллектуальных технологий: искусственных нейронных сетей (ИНС) и технологий нечеткой математики [5]. Существуют различные методы создания интеллектуальных аналитических приложений, которые дополнительно включают экспертные системы. В системах с базами знаний, в том числе и в экспертных системах, представление знаний является жизненно важным понятием, и решение о выборе метода представления знаний оказывает огромное влияние на любую из их составных частей. Обычно выделяют четыре группы методов представления знаний: логическое представление; сетевое представление; иерархическое представление; процедурное представление. Исходя из этого, модели знаний классифицируются аналогичным образом [6]:

1. Продукционные модели.

Продукционная модель наиболее соответствует процедурному характеру знаний. Эта модель описывает знания в структуре предложений типа «Если (левая часть правила), то (правая часть правила)» [4]. Характерными призна-

ками продукционной модели являются: наглядность, высокая модульность, простота внесения дополнений и изменений в базу знаний.

Системы с базами знаний, в основном основанными на этой модели, называются продукционными системами. Эти системы бывают двух диаметрально противоположных типов – с прямыми и обратными выводами.

В системе с обратными выводами дерево И / ИЛИ строится с использованием правил, связывающих факты и выводы в единое целое; оценка этого дерева на основе фактов, доступных в базе данных, является логическим завершением. Логические выводы могут быть прямыми, обратными и двунаправленными. При прямом выводе предоставленные данные служат начальной точкой, процесс оценки приостанавливается в узлах с отрицанием, а предположение, соответствующее самому верхнему уровню дерева (корню), используется в качестве заключения (если дерево не обработано полностью). Однако такой вывод характеризуется гигантским объемом данных, а также древовидными оценками, которые уже не имеют прямого отношения к заключению, а в этом нет необходимости. Преимущество обратных выводов заключается в том, что оцениваются только те компоненты дерева, которые имеют отношение к заключению, однако, если отрицание или утверждение невозможно, создание дерева бессмысленно. В двунаправленных выводах сначала оценивается небольшой объем полученной статистики и выбирается гипотеза (после случая прямых выводов), а затем запрашиваются данные, необходимые для выбора приемлемости этой гипотезы. На основе этих выводов может быть реализована более эффективная и гибкая система.

Системы с прямыми выводами среди систем, основанных на использовании знаний, имеют самую долгую историю, поэтому они в некотором смысле являются фундаментальными. Эти системы состоят из трех компонентов: базы правил, состоящей из набора продуктов (правил вывода), базы данных, содержащей набор фактов, и интерпретатора для получения логического заключения на основе этих знаний. База пра-

вил и база данных структурируют базу знаний, а интерпретатор соответствует механизму логического вывода. Вывод выполняется в форме цикла понимания-выполнения, и в каждом цикле выполняемая часть выбранного правила обновляет базу данных. В результате содержимое базы данных трансформируется из уникального в целевое, то есть целевая система синтезируется в базе данных. Другими словами, производственная система характеризуется простым циклом выбора и выполнения (или оценки) правил, но из-за необходимости периодического сопоставления с образцом в базе правил (идентификация) скорость вывода заметно снижается с увеличением количества последних (правил).

Достоинства модели знаний:

- простота создания и понимания, отдельных правил;
- простота модификаций и расширения базы знаний;
- простота механизма логического вывода;
- естественная модульность организации баз знаний.

Недостатки:

- неясность взаимных отношений правил;
- сложность оценки целостного образа знаний;
- крайне низкая эффективность обработки;
- отличие от человеческой структуры знаний;
- отсутствие гибкости в логическом выводе.

2. Семантические сети

Основная идея данной стратегии представления знаний состоит в том, что проблемная область рассматривается как совокупность объектов и связей между ними [7]. Семантическая сеть – это ориентированный граф, вершинами которого являются понятия, а дуги обозначают отношения между ними.

Основным практическим элементом семантической сети является фигура, состоящая из двух компонентов: дуги и полярные пары вершин. Можно предположить, что каждый такой фактор представляет собой простой факт или декларацию предметной области. Очевидно, что каждая вершина графа может быть связана с любым количе-

ством различных вершин. Таким образом осуществляется формирование сети фактов.

Базовую структуру семантической сети с точки зрения логики может рассматриваться как эквивалент бинарного предиката. Тогда предикат соответствует направленной дуге, а два его аргумента соответствуют двум вершинам, инцидентным этой дуге.

Семантические представления используют три основных типа объектов: концепции, события и свойства. Понятия обычно соответствуют объектам или абстракциям. Под событием понимаются действия, которые приводят к изменению ситуации. Свойства используются для уточнения или изменения концепций, событий или различных свойств.

При представлении знаний семантическими сетями особый практический интерес представляют следующие виды связей: «есть», «имеет», «есть». Отношение структуры «есть» отражает принадлежность к определенному типу объектов. Связь формы «имеет» и «есть» указывает, соответственно, на то, что одно понятие служит атрибутом другого. Используя эти отношения, довольно просто представить сложные наборы фактов.

Достоинствами моделей представления знаний с использованием семантических сетей являются большие выразительные возможности, наглядность и близость формы сети к семантической структуре предметной области.

Наиболее важным недостатком таких моделей является уязвимое развитие теоретических основ логического вывода, что увеличивает аспект произвольности, введенный разработчиком системы точной обработки знаний, в первую очередь основанной на семантической сети. Следовательно, процедуры вывода в таких системах регулярно создают угрозу противоречий, и, следовательно, необходимо уделять больше внимания процедурам проверки несогласованности в семантических сетях, чем, например, в производственных системах. Очень часто в семантических моделях эта функция возлагается на человека. Таким образом, при достаточно большом объеме знаний ответ на эту проблему резко усложняется, что ограничивает круг предметных областей, описываемых с помо-

щью семантических сетей, относительно небольшими задачами.

3. Фреймы.

Фреймовый подход к представлению знаний является частным случаем моделей представления знаний с помощью семантических сетей [8]. Здесь единицей представления является объект, называемый фреймом. Под фреймом понимается абстрактный образ для представления стереотипа восприятия, который служит для обозначения структуры знаний для восприятия пространственных структур. Фрейм представляет собой некоторую законченную структуру, которая при заполнении слотов значениями описывает вполне конкретный факт, событие, явление или процесс.

Фреймовые модели обеспечивают требования к структурированию и связности. Это выполняется из-за свойств наследования и вложенности, которыми обладают фреймы, то есть система именованных слотов нижнего уровня может действовать как слоты, а слоты могут дополнительно использоваться как вызовы любых процедур для выполнения.

Для многих предметных областей фреймовые модели являются основным способом формализации знаний.

4. Формально-логические модели.

Формально-логические модели представления знаний полностью основаны на описании предметной области в структуре набора аксиом. Логическая модель используется для представления ноу-хау в системе логики предикатов первого порядка и логики высказываний.

Основные преимущества использования логики предикатов для представления знаний заключаются в том, что можно напрямую запрограммировать мощный механизм вывода с хорошо понятными математическими резиденциями; есть возможность управлять логической целостностью базы знаний, то есть ее согласованностью и полнотой; документировать факты очень просто, что обеспечивает простоту и компактность базы знаний. Главный недостаток логической модели – отсутствие четких стандартов организации фактов в базе знаний. В связи с этим довольно большие базы данных сложно анализировать и обрабатывать.

Таким образом, полезно использовать логические модели представления знаний в тех предметных областях, где база знаний небольшая по объему и довольно простая по структуре.

Подведем итог полученному описанию и выберем оптимальную модель знаний для экспертной системы с применением нечеткой математики.

В практике создания интеллектуальных систем существует тенденция к использованию гибридных моделей для прояснения различных практических проблем. Экспериментальные исследования показывают, что использование этого подхода позволяет нам добиваться лучших результатов в увеличении скорости и точности решаемых задач.

Для решения широкого круга экономических задач, содержащих использование полуструктурированной информации, целесообразно использовать методы обучения вычислительных устройств и «мягких вычислений». При этом облик таких систем, как правило, включает в себя базу знаний, которая представляет собой реализацию разработанных алгоритмов анализа и синтеза правил и знаний. Авторы дополнительно опирались на гибридную модель базы знаний по использованию миварных технологий [1, 9].

Таким образом, применительно к проблеме формирования баз знаний экспертных диагностических систем актуальной задачей является разработка концепции эффективного хранения статистики для такой базы знаний. Предполагается, что устройство будет искать определенное состояние в обширном пространстве, которое будет удовлетворять всем внешним факторам принятия решения. Обратимся к опыту построения подобных систем.

Первым планировщиком, выполняющим планирование в области состояния, был STRIPS (Stanford Research Institute Problem Solver) [10], который должен был использоваться для решения проблемы формирования дизайна поведения робота, перемещающего объекты через множество комнат. Идея алгоритма STRIPS заимствована из системы General Problem Solver (GPS). Метод, используемый в GPS, известен как «анализ средств и целей» (англ.

means-ends analysis). Он подразумевает рассмотрение в текущем состоянии только тех ходов, которые имеют отношение к цели. STRIPS применяет действия, кроме задержки, достигая каждой цели индивидуально.

В 2018 году было предложено использовать миварные экспертные структуры для решения задач интеллектуального планирования. Из [11] следует, что миварные технологии также реализуют мышление метода GPS, но переводят его в формализм продукционного подхода и сетей Петри, что позволило избавиться от полного перебора и перейти к линейной сложности логического вывода.

Миварные технологии используются для решения различных задач в области логического искусственного интеллекта, в том числе задач распознавания образов, планирования поведения роботов и соблюдения правил дорожного движения, растущих виртуальных текстовых консультантов.

Следует отметить, что выбранная продукционно-фреймовая модель представления знаний с применением технологии мивар универсальна, так как она обладает: простотой механизма управления выводом; способностью отображения концептуальной основы организации памяти человека; легкостью и модульностью, поскольку значения слотов фреймов могут быть добавлены, удалены или изменены независимо от других; эффективностью при объединении декларативных и процедурных знаний в одной базе знаний; эффективностью при обработке семантической составляющей знаний – фреймы математически представляются как ориентированные графы с помеченными вершинами и дугами; универсальностью при отображении всего многообразия знаний.

Следует отметить, что современные понятия класса и объекта, используемые в языках программирования высокого уровня, которые поддерживают парадигму объектно-ориентированного программирования, довольно точно соответствуют таким классическим понятиям, как фрейм-образец и фрейм-экземпляр. В связи с этим структура разработанной базы знаний представлена на рис. 1.

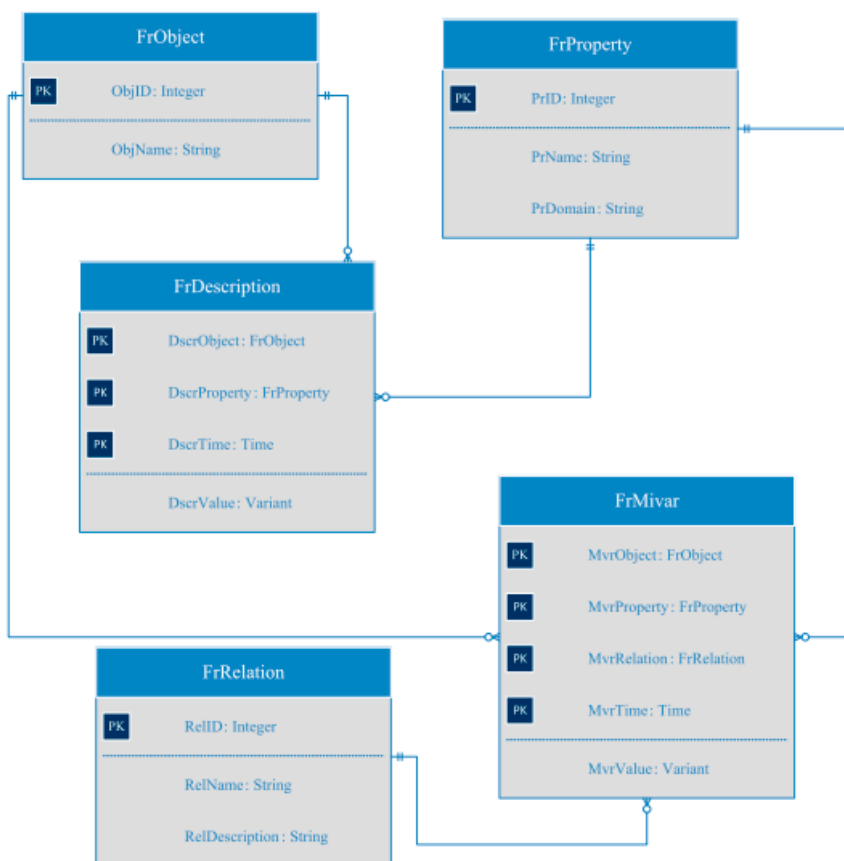


Рис. 1. Структура хранения экспертной системы [1]

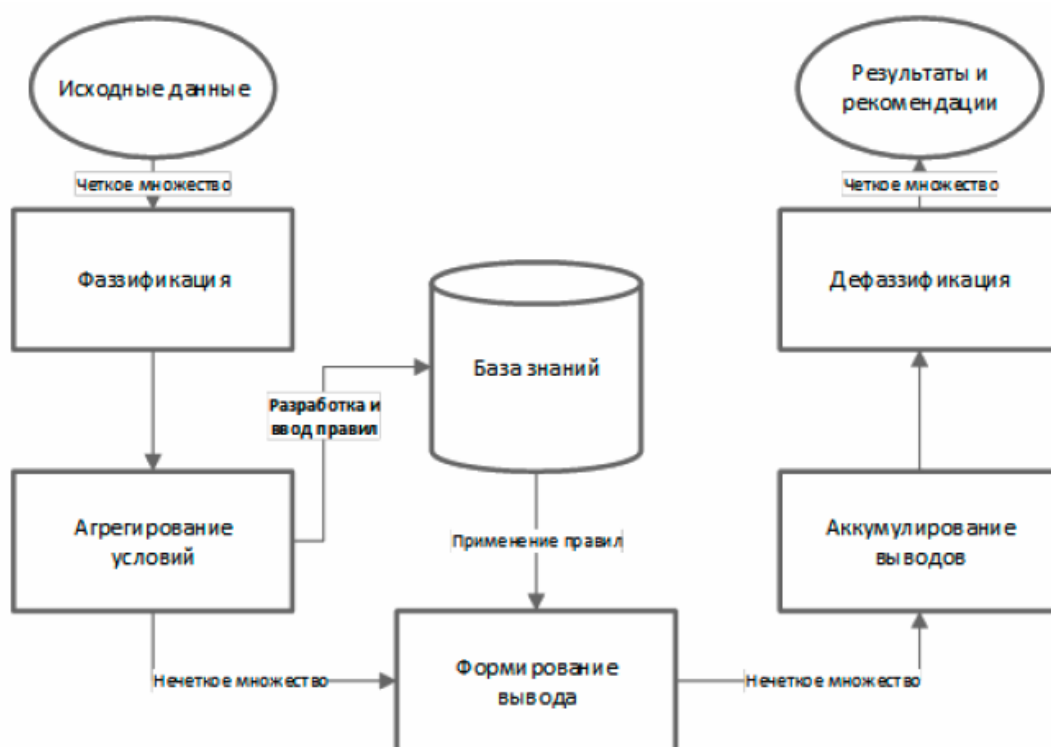


Рис. 2. Структурная схема формирования результатов с использованием нечеткой системы вывода [2]

База знаний включает в себя пять классов (примеры фреймов):

1. FrObject – образец фрейма для хранения массива объектов [фрейм имеет идентификатор и имя];

2. Свойство FrProperty – образец фрейма для хранения общего описания свойств (с идентификатором и именем, которое он имеет, поле, содержащее диапазон допустимых значений – домен PRD);

3. FrRelation – образец фрейма для хранения описаний отношений (атрибуты содержат имя, идентификатор и общее описание отношений);

4. FrDescription – образец фрейма, реализующий присвоение динамических свойств объектам (DscrObject – массив объектов, DscrProperty – массив свойств, DscrTime – период фиксации свойств, DscrValue – массив

значения, могут быть любого типа в зависимости от типа свойства);

5. FrMIVAR – образец фрейма, в котором хранятся МИВАРЫ.

Экспертная система подобного типа может функционировать даже при недостатке информации (т. е. неопределенности), так как согласно [2] предполагается использовать процедуру нечеткого вывода, которая представлена на рисунке 2.

Выводы

В работе представлен краткий обзор применяющихся в настоящее время моделей представления знаний в экспертных системах. Предложено использование гибридной модели с использованием технологии мивар для функционирования нечеткой системы вывода. В целом, авторы планируют в дальнейшем интегрировать данную концепцию в единую систему распознавания и интерпретации выбросов на экономических временных рядах

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта No19-01-00370.

Библиографический список

1. Копырин А.С. Development of the General Structure of the Knowledge Base for Neuro-Fuzzy Models, 2020.
2. Копырин А.С., Копырина А.О. Development of the generic system of inference rules by knowledgebase // Программные системы и вычислительные методы. 2021. № 1. С. 1–9.
3. Китанина К.Ю., Хромушин В.А., Дзасохов А.С., Хромушин О.В. Особенности построения экспертной системы на основе алгебраической модели конструктивной логики // Вестник новых медицинских технологий. 2015. № 3.
4. Аджит А., Семченко П.Н. Экспертные системы на основе правил // Вестник тихоокеанского государственного университета. 2013. № 3 (30). С.29-40.
5. Полковникова Н., Курейчик В. Разработка модели экспертной системы на основе нечёткой логики // Известия Южного федерального университета. 2014. № 1 (150). С. 83-92.
6. Горбатенко С.А., Даценко Н.В. Представление знаний в проблемно-зависимой информационной базе гуманитарной экспертной системы // Вестник Воронежского института МВД России. 2007. № 1. С. 137-140.
7. Юсупова Н.И., Гаянова М.М. Семантические сети и продукционные модели для анализа университетских образовательных программ в информационной системе // Вестник Уфимского государственного авиационного технического университета. 2006. Т. 7. № 2. С. 123-126.
8. Барышев М. В., Гагчин И. Ю., Гагчин Ю. А. Модели представления знаний экспертных систем // Научно-технический вестник информационных технологий, механики и оптики. 2006. № 29. С. 14-18.
9. Максимова А.Ю., Варламов О.О. Миварная экспертная система для распознавания образов на основе нечеткой классификации и моделирования различных предметных областей с автоматизированным расширением контекста // Известия Южного федерального университета. Технические науки. 2011. № 12 (125). С. 77-87.
10. Fikes R.E., Nilsson N.J. Strips: A new approach to the application of theorem proving to problem solving. Artificial Intelligence. 1971.
11. Эволюционные базы данных и знаний для адаптивного синтеза интеллектуальных систем. Миварное информационное пространство. [Электронный ресурс]. URL: <https://www.elibrary.ru/item.asp?id=21237254> (дата обращения: 23.09.2021).