

ОБОБЩЕНИЕ ЗНАНИЙ В ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЕ С СЕМИОТИЧЕСКОЙ МОДЕЛЬЮ ПРЕДМЕТНОЙ ОБЛАСТИ

Григорьев А.В.

Кафедра ПМИИ, ДонГТУ
grigorie@r5.dgtu.donetsk.ua

Каспаров А.А.

ВЦ, ДонГТУ
artemka08@ukr.net

Abstract

Grigoriev A. V., Kasparov A.A. Generalization of knowledge in intelligence systems. Deciding any problem in intelligence systems precedes a stage of educating a system, stage of making a knowledgebase.

Введение

Обучение на основе накопленных знаний о предметной области является одним из важнейших этапов работы интеллектуальной системы. Формы и методы представления знаний, применяемые на этапе обучения, определяют эффективность работы остальных механизмов интеллектуальной системы [1]. Процесс обучения предполагает выполнение внутрисистемной идентификации объектов, обобщение объектов и определение способа представления результатов. На этапе обучения необходимо учитывать: специфику задач решаемых интеллектуальной системой; концепцию представления предметных областей [2]; модель представления знаний в интеллектуальной системе; наличие эксперта и уровень его квалификации и т.д.

Цель данной работы - сравнительный анализ концепции обобщения знаний, принятой в среде мета-эвристической оболочки (МЭО), с прочими существующими концепциями построения механизмов обобщения.

МЭО предназначена для создания интеллектуальных систем проектирования [3]. Сфера применения МЭО - построение интеллектуальных САПР - требует учета основных положений САПР и системного анализа (СА) при построении модели предметной области. Спецификой представления предметной области в МЭО является применение семиотической модели (СМ), т.е. открытой формальной системы представления знаний [4]. К важнейшей особенности СМ можно отнести текстовый или грамматический подход к представлению моделей. Обучение в интеллектуальной системе с СМ предметной области может трактоваться как построение конструктивных процедур π и ψ (см. определение этих процедур в [4]): ψ - определяет принадлежность каждого объекта предметной области некоторому подмножеству объектов этой предметной области; π - идентифицирует различия объектов внутри каждого подмножества объектов предметной области.

Результат обобщения (обучения) трактуется как тип в СА, т.е. совокупность описаний структурной и функциональной компонент модели, рассмотренных во времени и пространстве (жизненный цикл), при наличии обязательных (системообразующих) и дополнительных (факультативных) признаков типа [5].

Грамматическая форма представления моделей в СМ позволяет трактовать тип как порождающую грамматику, а процедуры π и ψ строить на базе теоретико-множественных операций над порождающими грамматиками. Ориентация МЭО на модель САПР, предложенную в теории сложности САУ [6], требующую счетного ограни-

ченного числа решений, имеющихся в базе знаний САПР, позволяет трактовать тип, заданный в виде порождающей грамматики, как форму представления И-ИЛИ-дерева. Наличие И/ИЛИ-дерева решений обеспечивает последующее решения таких задач как проектирование, моделирование и изобретение.

Ориентация МЭО на САПР и представление модели предметной области в форме СМ приводит к следующим этапам процесса обобщения знаний:

- 1) типизация объектов предметной области, что предусматривает начальное распознавание объектов;
- 2) выделение общих и различных фрагментов в описании объектов, имеющих один тип;
- 3) построение ИЛШИИ-дерева решений.

1. Системный анализ в существующих концепциях представления предметных областей интеллектуальных систем

В настоящее время в большинстве разработок предпринимаются попытки рассмотрения модели предметной области с позиции СА. Достоинства и недостатки любой концепции могут быть определены исходя из полноты реализации положений СА.

Результаты анализа некоторых типичных концепций в сравнении с предлагаемой авторской концепцией представлены в таблице 1.

Таблица 1 - Некоторые современные концепции представления предметной области

Интеллектуальная система	Общие сведения о представлении предметной области	Достоинства	Недостатки
1. Экспертная система «Аргонавт» <i>UI</i>	Представлена в виде пространства координат D^n . Каждая координата X_i , где $i=1, n$ является характерным признаком предметной области.	Легко задавать обучающие выборки.	1. Не просматривается внутренняя структура объектов, их взаимосвязь. 2. Не учтено существование недоопределенных значений признаков.
2. Адаптивная система подготовки и аттестации специалистов [8].	Представлена множеством объектов $O = \{O_j, i \in I_0\}$, обладающих свойствами $F = \{f_{ij}, i \in I_0, j \in J, f_{ij}\}$ и множества отношений $Z = \{Z_k, k \in I_k\}$. Для динамики объекта вводится категория время.	Есть связи между объектами.	1. Явно выделено время, НО не учтено пространство. 2. Нет недоопределенностей.

Продолжение таблицы 1

Интеллектуальная система	Общие сведения о представлении предметной области	Достоинства	Недостатки
3. МЭО для построения интеллектуальных САПР [2,3]	Предметная область многоуровневая. Каждый уровень включает отношения: тип блока, блок, среда блока, свойства и значения границ блока, связь, массив, шкала, агрегация, обобщение, функция, неопределенный элемент Nil. Модель САПР рассматривается в рамках теории сложности систем автоматизации управления (ТС САУ) [6].	1. Наличие механизма поддержки представления о пространственно-временных связях. 2. Работа с недоопределенными моделями сложных объектов. 3. Обеспечение различных уровней представления, принятых в моделях САПР.	Сложность работы с данной концепцией как эксперта, так и механизмов МЭО.

Предлагаемая концепция представления предметной области для МЭО позволяет объединять возможности двух приведенных типичных подхода и, вместе с тем, максимально приблизиться к моделям СА.

2. Особенности подхода к типизации объектов

В предлагаемой концепции процесс типизации соответствует обобщению объектов по свойствам в других интеллектуальных системах. Специфика задачи обобщения заключается в системном подходе, т.е. в представлении объектов как совокупностей свойств, возникающих в точках взаимодействия этих объектов [2]. В пределах отдельного описания структуры объекта распознаваемые подобъекты рассматриваются с позиции «черного ящика», т.е. видимой является лишь их внешняя граница.

Организация процесса типизации в МЭО заключается в построении решающей функции f , которая определяет принадлежность любого объекта к интересующему нас подмножеству множества объектов данной предметной области. Построение данной решающей функции в интеллектуальной системе с семиотической моделью представления знаний и является целью работы процедуры ψ [4].

Существуют следующие варианты проведения этапа типизация объектов предметной области в интеллектуальной системе:

- задание типов непосредственно экспертом [3];
- задание экспертом обучающих, отрицательной и положительной, выборки для каждого типа объектов с последующим автоматическим проведением типизации новых объектов, информация о которых подается в систему (обучение «с учителем») [7, 9];
- ® полностью автоматизированная типизация объектов - фактически обучение «без учителя».

Иллюстрация и анализ некоторых подходов к обучению по свойствам приведены в таблице 2.

Можно сделать заключение, что для обеспечения высокой эффективности МЭО необходимо использовать несколько различных вышеуказанных вариантов типизации. Такой подход к построению интеллектуальных систем значительно расширяет возможности их промышленного применения, так как позволяет выбирать механизм приобретения знаний в зависимости от ситуации.

3. Специфика предлагаемого метода обобщения

В предлагаемой оболочке применяется метод обобщения, основанный на выделении общих и различных фрагментов описаний объектов, имеющих один тип. Это позволяет отнести предлагаемый метод к классу методов обобщения по структурам [4].

Под фрагментом понимается взаимосвязанная часть внутренней среды объекта, т.е. совокупность подобъектов и связей между ними.

Построение решающей функции ϕ , осуществляющей определение принадлежности фрагментов внутренних сред объектов на некотором их множестве, объединенным одним типом, есть задача выделения общей и альтернативных частей в сравниваемых объектах.

Таблица 2 - Некоторые подходы к обобщению по свойствам (признакам)

Интеллектуальная система	Общие сведения об обобщении по свойствам (признакам)	Наличие эксперта	Достоинства	Недостатки
Экспертная система «Аргонавт» [7].	В пространстве D^n определяют два набора векторов: T - положительные и F - отрицательные. По выборке примеров T , F строят функцию выбора. Выдвигаются гипотезы, а эксперт оценивает их качественно как положительные и отрицательные.	<i>РА</i>	1. Возможность построения многоуровневой функции выбора. 2. Возможность выделять наиболее существенные признаки.	1. Требуется опытный эксперт.
МЭО для построения интеллектуальных САПР [3]	Задание типов производится непосредственно экспертом.	№	1. Упрощается процесс обучения. 2. Уменьшается вероятность появления неточностей.	1. Требуется опытный эксперт.

Предлагаемое решение предусматривает, что в системе с СМ представления знаний за построение данной решающей функции возлагается на процедуру ψ

Задача обобщения по структурам тесно связана с активно развивающимися в последнее время методами поиска закономерностей, опирающихся на аппарат логики, а не на традиционный для этого класса задач аппарат математической статистики и теории вероятностей [4]. По своей постановке данная задача похожа на задачи, возникаю-

шие в индуктивной логике. Ее решение заключается в построении аксиом для последующего проведения, с их использованием, логического вывода.

Формирование аксиом проводится в соответствии с двумя схемами, являющимися модификацией схем индуктивного вывода С. Милля [4]:

а) Схема выделения общих фрагментов:

$$\begin{array}{l}
 a \in p_1, \\
 a \in p_2, \\
 a \in p_3, \\
 \dots\dots\dots \\
 a \in p_J, \\
 \hline
 \forall_{p_j \in P} : a \in p_j, j = \overline{1, J}.
 \end{array} \tag{1}$$

где

a - некоторый фрагмент внутренней среды объекта;

p_j - объект с порядковым номером j в пределах множества объектов данного типа;

P - множество объектов данного типа.

По схеме (1) находятся фрагменты внутренней среды, общие для всех объектов предметной области.

б) Схема выделения альтернативных фрагментов:

$$\begin{array}{l}
 b \in p_1, \\
 b \in p_2, \\
 \dots\dots\dots \\
 b \in p_m, \\
 b \notin p_{m+1}, \\
 b \notin p_{m+2}, \\
 \dots\dots\dots \\
 b \notin p_J, \\
 \hline
 (\forall_{p_i \in P' = \{p_1, p_2, \dots, p_m\}} : b \in p_i) \wedge (\forall_{p_j \in P \setminus P'} : b \notin p_j), \\
 i = \overline{1, m}; j = \overline{m+1, J}.
 \end{array} \tag{2}$$

где

b - некоторый фрагмент внутренней среды объекта;

p_i, p_j - объект с порядковым номером \mathcal{J} в пределах множества объектов данного типа;

P - множество объектов данного типа;

$P \setminus P'$ - подмножество множества объектов данного типа, для которых фрагмент b является общим.

С помощью схемы (2) определяются альтернативные фрагменты множества объектов одинакового типа в пределах данной предметной области.

Процедура П2 конструктивна, так как конечно множество объектов предметной области и, следовательно, конечно множество их фрагментов.

4. Порядок построения И/ИЛИ-дерева вывода на основе фрагментов внутренних сред объектов

Пусть $P = \{p_i\}$ - множество объектов некоторого типа, где $i = 1, K$ - номер объекта в базе знаний. Символом «@» будем обозначать фрагменты внутренней среды объектов. Тогда:

$$\forall p_i \in P : \&(@j), \quad (3)$$

где

@j - некоторый фрагмент среды (как совокупность связей).

Таким образом, можно построить P как И/ИЛИ-дерево, вершины которого представляют собой синтермы, объединяющие различные альтернативы возможных фрагментов внутренней среды объекта.

Алгоритм построения ШИЛИ-дерева - это алгоритм выполнения теоретико-множественных операций над порождающими грамматиками, представленными в форме БНФ (формы Бэкуса-Наура). С его помощью может быть получено множество различных комбинаций фрагментов описаний объектов. Среди данных комбинаций имеются в том числе и синтаксически и семантически правильные комбинации, т.е. составляющие описания исходных объектов. Количество различных комбинаций, включая пустое множество \emptyset , можно определить по формуле:

$$N_p = 2^n, \quad (4)$$

где

N_p - число комбинаций исходных фрагментов сред;

n - количество исходных фрагментов.

Приведем пример построения части И/ИЛИ-дерева на основе склеивания множества, состоящего из трех объектов одинакового типа: $P = \{p_1, p_2, p_3\}$, $|P| = 3$. На рисунке 1 схематично показаны три исходных объекта и общие и альтернативные фрагменты их структуры, выделенные после отработки процедуры обобщения по структурам.

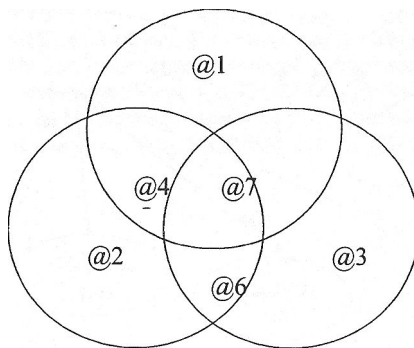


Рисунок 1 - Выделение общих и различных фрагментов описаний при обучении по структурам

В общем случае количество фрагментов составляет $n=7$. По формуле 1.4 можно вычислить число получаемых комбинаций после применения теоретико-множественных операций над множеством фрагментов: $N_p = 2^7 = 128$. Среди них находятся и комбинации-аксиомы, т.е. описания исходных объектов-прототипов:

$$\begin{aligned} p_1 &= @1 \& @4 \& @5 \& @7, \\ p_2 &= @2 \& @4 \& @6 \& @7, \\ p_3 &= @3 \& @5 \& @6 \& @7. \end{aligned}$$

Определим некоторое подмножество комбинаций, ПО которому и построим ИЛТЛИ-дерево: $P = \{p_1, p_4, p_5\}$. Здесь:

$$\begin{aligned} p_4 &= @1 \& @4 \& @5 \& @7, \\ p_5 &= @4 \& @5 \& @3. \end{aligned}$$

$$P_5 = @1.$$

Тогда в результате операции склеивания получим:

$$p' = @4 \& @5 \& (@1 \& @7 \vee @3) \vee @1, \tag{5}$$

где

p - идентификатор модуля знаний о структуре внутренней среды объектов одного типа;

$@j$ - некоторый фрагмент внутренней среды объекта.

На рисунке 2 изображена часть полученного ШИЛИ дерева.

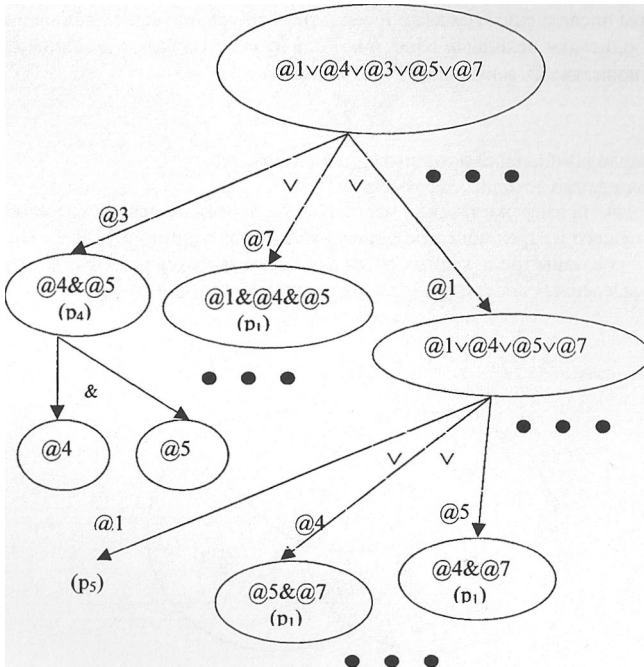


Рисунок 2 - Фрагмент И/ИЛИ-дерева

В вершинах дерева раскрыты значения синтермов, рядом со стрелками показан идентификатор фрагмента, при выборе которого мы двигаемся к следующей вершине по ветви вывода. Описанный алгоритм построения И/ИЛИ-дерева вывода был использован при реализации МЭО для создания интеллектуальных САПР [3].

Вывод

Сравнительный анализ специфики, достоинств и недостатков концепции обобщения знаний, принятой в МЭО, с прочими существующими концепциями построения механизмов обобщения показал наиболее полное соответствия решений, принятых в МЭО, области ее применения и особенностям СМ как формы представления предметной области. Предложенные методы решения задачи обобщения в среде МЭО обеспечили:

- возможность практической реализации МЭО;
- эффективное решение поставленных перед МЭО задач.

Литература

1. А.Ф. Колчин, О.В. Красников, С.Е. Чекменев. Об одном подходе к выбору комплекса программно-аппаратных средств для создания интеллектуальных систем.
2. А.В. Григорьев. Система уровней модели предметной области для интеллектуальных САПР, Москва: Физико-математическая литература, КИИ'2000, 2000, С. 633-642.
3. А.В. Григорьев, А.А. Каспаров. Особенности реализации мета-эвристической оболочки для построения САПР. В книге: Наукові праці Донецького державного технічного університету; серія: Проблеми моделювання та автоматизації проектування динамічних систем, Випуск 10, Донецьк: ДонГТУ,1999, С. 217-222.
4. Поспелов Д.А. Ситуационное управление: теория и практика. М.: Наука, 1986. - 288 с.
5. Зубенко Ю.Д. Системный анализ. Донецк: ДонГТУ, 1995. — с. 166.
6. Солодовников В.В., Тумаркин В.И. Теория сложности и проектирование систем управления. - М.: Наука. 1990. - 186 с.
7. М.О. Корлякова. Многоуровневая экспертная система на основе принципа обобщения примеров по признакам, Москва: Физико-математическая литература, КИИ'2000, 2000, С. 103-111.
8. А.П. Бень, В.И. Литвиненко, В.Е. Ходаков. Использование эпистемических логических исчислений для представления знаний в интеллектуальных интерфейсах, Москва:Физико-математическаялитература,КИИ2000, 2000,С. 158-164.
9. А.В. Куликов, М.В. Фомина. Разработка алгоритма обобщения знаний, Москва: Физико-математическая литература, КИИ'2000,2000,С. 135-141.