

Репрезентация знаний

В.С. Выхованец

Старший научный сотрудник, к.т.н., доцент

Институт проблем управления РАН, valery@vykhovnets.ru, Москва

Под знанием будем понимать субъективно обоснованное личностное убеждение, выраженное в некоторой материальной форме и признанное объективно истинным. По своей сути знание – проверенный на практике результат субъективного отражения объективной действительности, представляемый в виде понятий и суждений, утвержденных некоторой последовательностью умозаключений.

Помимо задач извлечения, представления и актуализации знаний имеется другая актуальная задача – их репрезентация, заключающаяся в изменении формы представления знаний. Такая задача, в частности, возникает в CAD/CAM/PDM системах, когда знания, накопленные в процессе функционирования таких систем, преобразуются в другую форму, отличную от исходной.

При репрезентации знания преобразуются в новую форму с учетом некоторой активной проблематики, возможно отличающейся от проблематики, которая была использована при их извлечении и актуализации, а сам процесс преобразования описывается знанием о накопленных ранее знаниях, или метазнанием.

Задача репрезентации знания не может быть решена в полном объеме средствами, заранее встроенными в CAD/CAM/PDM системы. Для ее решения необходимо уметь извлекать, представлять и актуализировать метазнания, описывающие процесс репрезентации предметных знаний.

Репрезентацию знаний предлагается осуществлять на основе построения понятийной модели. Понятийная модель включает понятийную структуру предметной области репрезентации, правила извлечения фактов и правила выражения объема и содержания понятий.

1. Понятийная структура

Понятийную структуру представим множеством понятий, на котором заданы отображения одних понятий в другие, задающие способы их абстрагирования.

1.1. Понятия

Понятие будем рассматривать как сложный когнитивный феномен, не имеющий однозначной семиотической репрезентации. Последнее связано с такими свойствами мышления как предметность (ориентация на открытие еще неизвестных свойств предметов познания), рефлексивность (саморегуляция познавательной деятельности), субъектность (активная перестрой-

ка мыслительных механизмов в процессе познания) и продуктивность (постоянное производство новых содержательных обобщений) [1].

Однако семиотическая репрезентация понятий видится единственной и принципиально необходимой формой представления и обработки знаний. По этой причине описываемый далее проблемный подход и пополняемое множество форм многоаспектного выражения понятий является некоторым неизбежным компромиссом между понятием как когнитивным феноменом и понятием как сложно организованной совокупностью существенных признаков.

Представим понятие множеством сущностей, объединенных по общности своих признаков. Имя понятия определим как его знаковое выражение. Схему понятия зададим набором признаков, характерных для этого понятия. Интенционал будем рассматривать как наборы значений взаимосвязанных признаков, позволяющие выделять сущности, которые принадлежат понятию и составляют его экстенционал.

Следует заметить, что понятие обладает фрактальностью, заключающейся в том, что признаки можно рассматривать как простые понятия, а сущности – как единичные. Более того, одна и та же сущность, будучи объективной реальностью, может быть использована для решения различных прикладных задач, в рамках которых она характеризуется различными признаками и образует различные понятия.

1.2. Абстрагирование

Понятия образуются при абстрагировании. Абстракция определяется как один из основных процессов познавательной деятельности человека, позволяющий мысленно вычлнить и превратить в самостоятельный объект рассмотрения отдельные свойства, стороны, элементы или состояния сущностей предметной области.

Известны следующие способы абстрагирования [2]: обобщение, типизация, ассоциация и агрегация (рис. 1). Обобщение и типизация выражают общность понятий, проявляющуюся при дифференциации признаков. Агрегация и ассоциация раскрывают объединение понятий, заключающееся в интеграции их признаков.

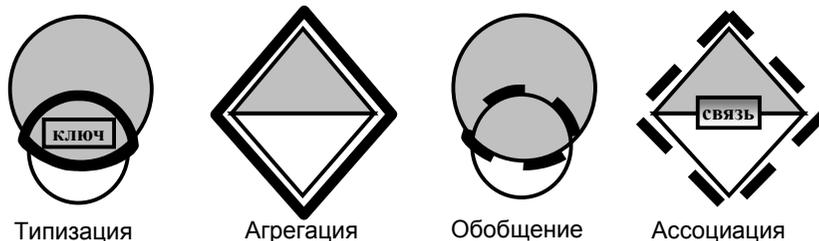


Рис. 1. Абстракции понятий

При обобщении происходит порождение нового понятия на основе одного или нескольких подобных понятий, когда порождаемое понятие сохраняет общие признаки исходных понятий. Типизация является частным случаем обобщения. В отличие от обобщения, при типизации имеется возмож-

ность для каждой сущности из экстенционала понятия-типа узнать ее исходное понятие. Для этого используется множество признаков, называемое ключом.

Ассоциация выражает специфическое соединение сущностей, которое позволяет от сущности одного понятия перейти к одной или нескольким сущностям других понятий, для чего используется набор признаков, называемый связью. Агрегация является предельным случаем ассоциации, когда между сущностями устанавливаются все возможные связи.

1.3. Понятийная структура

Под понятийной структурой будем понимать совокупность понятий, для которых заданы способы их образования (абстрагирования). Носителем понятийной структуры является множество понятий, а ее сигнатурой – множество отображений обобщения, типизации, агрегации и ассоциации.

На рис. 2 приведен пример понятийной структуры, где показаны понятия «Система», «Система1», «Система2», «ЭлементА», «ЭлементВ», «ЭлементС», «Управляет», «А1», «А2», «В1», «В2», «С1» и «С2».

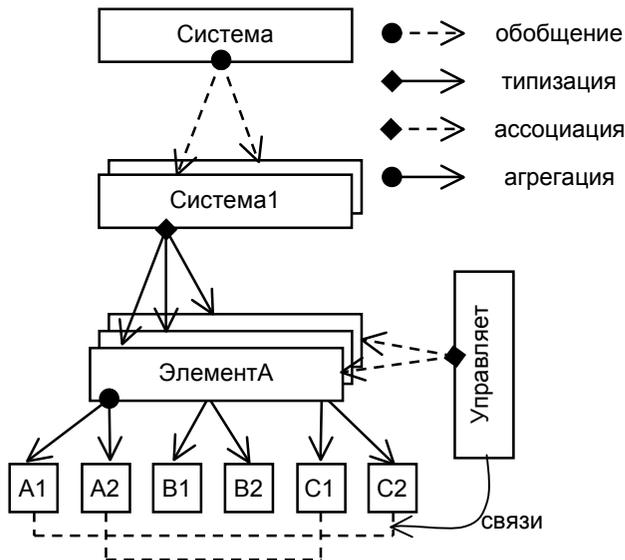


Рис. 2. Понятийная структура

Первичными понятиями предметной области являются понятия-сущности «А1» и «А2», «В1» и «В2», «С1» и «С2», причем такие, что результатом их типизации являются понятия «ЭлементА», «ЭлементВ», «ЭлементС» соответственно. Понятие «Управляет» описано как ассоциация понятий «ЭлементА» и «ЭлементС». Последнее означает, что между сущностями этих понятий устанавливаются ассоциативные связи такие, что появляется возможность от одной сущности понятия «ЭлементА» перейти к одной или нескольким сущностям понятия «ЭлементА», и наоборот. В свою очередь понятие «Система1» определено как агрегация понятий «Элемен-

тА», «ЭлементВ» и «ЭлементС», а понятие «Система» – как обобщение понятий «Система1» и «Система2».

Понятийная структура предназначена для выражения наиболее устойчивых знаний относительно предметной области и строится для каждой репрезентации. Для решения различных прикладных задач, например, для формирования различных документов, одна и та же предметная область, представленная текущим состоянием баз данных или знаний, будет описываться различными понятийными структурами. В этом случае эксперт для каждой репрезентации должен строить свою, присущую только этой репрезентации понятийную структуру.

Более того, для сложных предметных областей понятийная структура может зависеть от текущего состояния имеющихся данных и знаний. По этой причине предусмотрим гибкие средства для формирования понятийных структур, например, на интерпретируемом универсальном или специализированном языке программирования.

На рис. 3 показан пример фрагмента текста на языке программирования VBA, который был использован для адаптивного формирования понятийных структур при репрезентации знаний в виде текстовых документов в среде MS Word.

```
Set Module = TemplateOpen(SetupModel, SetupTemplates)
If Module Is Nothing Then
    Err.Raise ModelError, SetupCaption, ModelError
End If
Set ModelCollection = New Collection
Module.Common ModelErrors, ModelCollection, _
    SetupCaption, SetupConnect, _
    SetupRegion, SetupObject, SetupSystem, _
    SetupClass, SetupClone, SetupDocument
If ModelErrors <> "" Then
    Err.Raise ModelError, SetupCaption, ModelErrors
End If
Set col = ModelCollection.Add New Collection, "Volume"
Module.Volume ModelErrors, col, SetupClass
If ModelErrors <> "" Then
    Err.Raise ModelError, SetupCaption, ModelErrors
End If
```

Рис. 3. Формирование понятийной структуры

В рассматриваемом случае понятийная структура представляется коллекцией ModelCollection, состоящей из понятий, соединенных между собой связями обобщения (дерево) и ассоциации (перекрестные связи между узлами). Для формирования стандартных фрагментов понятийной структуры используется внешний подгружаемый модуль Module, содержащий повторно используемый код, например, для создания общей части всех понятийных структур (метод Module.Common), или стандартную понятийную структуру тома, в который входит создаваемый документ (метод Module.Volume).

Помимо стандартных фрагментов, при формировании понятийной структуры могут использоваться и специфические фрагменты кода.

2. Извлечение фактов

Правила извлечения фактов необходимы для извлечения фактов из имеющихся описаний предметной области и наполнения понятийной структуры конкретным предметным содержанием. Последнее означает, что понятийная структура выражает наиболее устойчивые представления эксперта относительно предметной области, в то время как факты являются переменной частью понятийной структуры и, по своей сути, составляют экстенционалы выделенных для репрезентации понятий.

Для извлечения фактов необходимо использовать выразительные средства, определяемые как форматами хранения данных и знаний относительно репрезентируемой предметной области, так и особенностями той CAD/CAM/PDM системы, в которых эти данные и знания хранятся. Например, для извлечения фактов из баз данных может быть использован язык структурированных запросов (SQL), а также те стандартные программные компоненты, которые необходимы для такого извлечения.

На рис. 4 приведен фрагмент функции MS SQL Server, предназначенной для извлечения значений признаков понятия Volume (том проектной документации).

```
CREATE FUNCTION [dbo].[fnModel_Volume] ( @intVolume smallint )
RETURNS @tabModel TABLE (
    fldIndex smallint identity primary key,
    fldName varchar(64) NULL,
    fldValue varchar(256) NULL)
AS BEGIN
    DECLARE @strTags varchar(1024), @strVolume varchar(16)
    SELECT @strVolume = CONVERT(varchar, fldVolume),
           @strTags = fldTags
    FROM dbo.tabType_Volumes
    WHERE fldVolume = @intVolume
    INSERT @tabModel (fldName, fldValue)
    SELECT ", @strVolume
    UNION ALL SELECT fldName, fldValue
    FROM dbo.fnTags_Split(@strTags, default, default)
    RETURN
END
```

Рис. 4. Извлечение фактов из предметной области

Вызов функции fnModel_Volume осуществляется из активного шаблона документа, а возвращаемые данные используются для наполнения понятийной структуры фактами. В частности, функция возвращает таблицу, описывающую некоторую сущность и имеющую три поля – номер признака fldIndex, имя признака fldName и его значение fldValue. Однако не видится

препятствий для задания правил извлечения фактов непосредственно в шаблоне документа, или в соответствующих полях базы данных.

3. Выражение понятий

Для репрезентации знаний, накопленных в понятийной модели предметной области, необходимо выразить эти знания в одной из возможных выходных форм – в виде рисунков, изображений, текста, звука, речи, и др. Понятно, что выразительные средства, используемые для такой репрезентации, будут зависеть от требуемой формы представления.

Для представления знаний в виде текстов на естественных и формальных языках правила выражения понятий могут быть оформлены в виде специальной разметки ранее упоминавшегося шаблона документа. В этом случае шаблон содержит фрагменты текста, предназначенные для выражения тех или иных закономерностей предметной области (умозаключений), появление которых в итоговом тексте связывается с построенной ранее понятийной моделью.

Для разметки шаблона используется специальный язык разметки, позволяющий задать действия (операции) над фрагментами текста при их переносе в выходной документ, которые выполняются в соответствии с имеющейся понятийной моделью.

В качестве примера на рис. 5 приведена формальная грамматика специализированного языка для текстового выражения понятий. При описании формальной грамматики использованы следующие соглашения:

- левая часть правила (нетерминальное понятие языка) отделяется от правой части (выражение нетерминального понятия в тексте) стрелкой;
- терминальные понятия языка задаются строками, заключенными в апострофы;
- строки, разделенные вертикальной чертой, выбираются альтернативно;
- строка в фигурных (квадратных) скобках может быть повторена один и более раз (ноль или один раз).

Первичными понятиями рассматриваемого языка являются Текст, Понятие, Предикат, Выражение, Императив и Формат. Текст представляет собой произвольный фрагмент текста, допускаемый используемым приложением, в том числе, содержащий специальные знаки, таблицы, рисунки, и т.п. Понятие – это составное имя понятия, позволяющее выполнить его идентификацию в понятийной модели. Например, для понятийной структуры на рис. 2 допустимо имя «Система.Система1.ЭлементВ». Если окажется, что идентифицируемое понятие не является одиночным понятием-сущностью, то значением идентификатора будет список сущностей, извлеченный из экстенционала этого понятия. Язык поддерживает также возможность определения предметных предикатов. Предикаты и значения признаков понятий, а также стандартные математические, логические и строковые операции используются в Выражениях. В свою очередь понятие Императив позволяет задать некоторую последовательность действий, которую необходимо выполнить в процессе репрезентации. Синтаксис и семантика этого понятия зависит от используемого приложения. И, наконец, первичное понятие

Формат необходимо для использования встроенной в язык развитой системы форматного преобразования.

Шаблон →
[Разметка] | [Разметка] Шаблон
Разметка →
Текст | Форматирование | Представление | Изменение |
Определение | Выполнение | Вычисление |
Выбор | Ветвление | Итерация
Форматирование →
'[LAYOUT] Текст '[FORMAT] Формат '['
Представление →
'[GET] Понятие '['[FORMAT] Формат] '['
Изменение →
'[SET] Понятие '['[VALUE] Выражение] '['
Определение →
'[DEFINE] Предикат ['(' {Понятие} ')'] '['[AS] Выражение] '['
Выполнение →
'[EXEC] Императив '['
Вычисление →
'[VALUE] Выражение '['[FORMAT] Формат] '['
Выбор →
'[IF] Выражение 'THEN' Текст '['ELSE' Текст] '['
Ветвление →
'[CASE] Выражение 'THEN' Текст
[{'[CASE] Выражение 'THEN' Текст}] '['ELSE' Текст] '['
Итерация →
'[FOR] {Понятие} '['[WHILE] Выражение '['[REPEAT] Текст
['[BETWEEN] Текст '['[END] Текст]]] '['

Рис. 5. Формальная грамматика языка репрезентации

4. Система репрезентации знаний

Для экспериментальной проверки полученных результатов реализована система текстовой репрезентации знаний (рис. 6), где используются созданные для каждого типа документа языки разметки. После грамматического разбора правил выражения понятий происходит генерация итогового документа по активному шаблону путем отображения накопленных в понятийной структуре фактов в виде фрагментов текста, построенных по правилам, задаваемым соответствующим языком разметки и выражающим предметное знание, подлежащее репрезентации.

Разработанная информационная система, являющаяся надстройкой над используемыми CAD/CAM/PDM продуктами, выполняет автоматическое формирование документов на основе обработки первичных данных и метазнаний. При этом реализованы принципы создания автоматизированных систем, включая принцип открытой архитектуры, позволяющий адаптиро-

вать и динамично развивать функции репрезентации знаний в соответствии с целевыми задачами пользователя на основе дополнения метазнаний, развития библиотек типовых фрагментов, формирования перекрестных связей между документами и их частями.

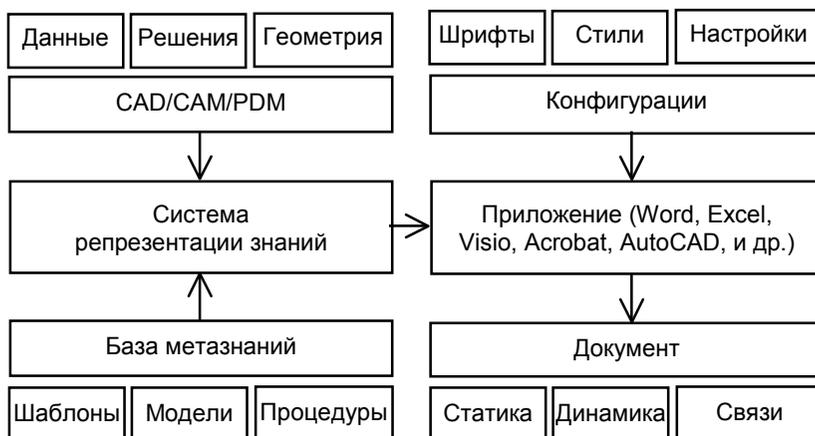


Рис. 6. Система репрезентации знаний

5. Заключение

Репрезентация знаний, являющихся результатами проектирования в CAD/CAM/PDM системах, осуществляется на основе построения понятийной модели предметной области, которая включает понятийную структуру, правила извлечения фактов и правила выражения понятий. Понятийная структура представляется множеством понятий, соединенных между собой связями обобщения (дерево) и ассоциации (перекрестные связи между узлами). К каждому узлу-понятию привязаны действия (правила извлечения фактов), которые необходимы для извлечения сущностей этого понятия из имеющихся данных и знаний с учетом проблематики репрезентации. Сама репрезентация фрагмента знания, представленного извлеченными фактами и структурированного понятийной структурой, осуществляется на основе правил выражения понятий в одной из возможных выходных форм – в виде изображений, текста, звука, речи, и др., для чего используются специальные языки репрезентации.

Литература

1. Кравцов Л.Г. Методологические проблемы психологического анализа мышления в понятиях // Материалы Первой российской конференции по когнитивной науке. Казань, Казанский гос. ун-т, 2004.
2. Brodie M.L., Mylopoulos J., Schmidt J.W. On Conceptual Modeling: Perspectives from Artificial Intelligence, Databases and Programming Languages. New York: Springer-Verlag, 1984.