

УДК 681.3.16

Л.Д. ГРЕКОВ

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Украина

ПРИМЕНЕНИЕ ФРЕЙМОВЫХ МОДЕЛЕЙ ЗНАНИЙ НА ОСНОВЕ ОНТОЛОГИЧЕСКИХ СИСТЕМ В ЗАДАЧАХ ЛОГИСТИКИ

Предложено применение фреймовых моделей представления знаний для учета земельных ресурсов в информационно-поисковых системах. Построение модели базируется на использовании онтологического подхода, что позволяет построить сложные взаимосвязи между семантическими объектами предметной области, и реализовать запросы произвольной сложности более простыми методами по сравнению с реляционными базами данных. Приведены примеры построения системы онтологий для учета земельных ресурсов НАК «Нафтогаз України».

фреймовая модель, онтологическая система, реляционная база данных, база знаний, семантическое подобие

Введение

Современный подход в управлении сложными производственными системами, к которым, безусловно, относятся и системы учета земельных ресурсов, предполагает не только активное использование ЭВМ для хранения и обработки информации, но и включение в контур автоматизации интеллектуальных функций, что позволит аппарату управления сосредоточиться на принятии наиболее ответственных решений, оставив рутинные вопросы средствам автоматизации. Например, в задаче создания электронного реестра земельных ресурсов актуальной функцией является формирование оперативной информации о состоянии земельных участков и объектов недвижимого имущества с обобщением высокого уровня данных и переходом на требуемый уровень детализации, оценки достоверности получаемой информации, выполнения сложных запросов с поиском подобных по семантическому критерию объектов и т.д. Для этого каждый объект системы автоматизации должен быть представлен некоей структурой, обладающей не только набором необходимых атрибутов, но и поведением, описывающим взаимодействие с другими объектами в информационной среде.

Применение моделей искусственного интеллекта позволяет реализовать приведенные выше задачи.

1. Применение фреймовой модели представления знаний в задачах логистики

Среди известных и широко используемых моделей представления знаний [1] фреймовые модели приобрели популярность благодаря возможности описания сложных процессов и явлений достаточно унифицированными конструкциями, а также реализации таких структур с использованием современных информационных технологий.

Определение 1. Фрейм – это структура данных, обладающая внутренними характеристиками (элементами), а также связью с другими фреймами.

Элементы фрейма называют *слотами*. Различают слоты, описывающие статические характеристики объекта, динамические характеристики (поведение), а также слоты, связывающие данный объект с родительскими либо семантически близкими объектами. Основной логический вывод во фреймовых моделях – механизм наследования. Дополнительно управлять выводом можно с помощью слотов, отвечающих за поведение – демонов и присоединенных процедур.

Например, в задачах учета земельных ресурсов актуальным является назначение земельного участка для прокладки нефтепровода и определение технической возможности выполнения такой задачи. В этом случае операцию назначения удобно предста-

вить в виде взаимодействия двух фреймов – «Ресурс» и «Элемент нефтетранспортной системы», имеющих инверсионные слоты, ссылающиеся друг на друга (рис. 1).



Рис. 1. Организация межфреймового взаимодействия

В отличие от организации такой связи, например, в реляционных базах данных, с необходимостью организации связей между отношениями и поддержкой ограничений целостности, фреймовая модель организует связь *автоматически*, в какой бы фрейм не заносилась информация. Преимуществом фреймовой модели является также то, что структура фреймов и их взаимодействие расположены в оперативной памяти компьютера, а не в файловой системе, как в случае с базами данных. Это позволяет, во-первых, сократить время доступа к информации, а также «держать» в памяти компьютера довольно сложный механизм взаимодействия фреймов (логику приложения), не прибегая к услугам триггеров, хранимых процедур, запускаемых опять же из файловой системы.

2. Реализация фреймовой модели с помощью онтологической системы

В настоящее время известно несколько определений онтологии.

Онтология – это формально представленные на базе концептуализации знания. Концептуализация предполагает описание множества объектов и понятий, знаний о них и связей между ними. Таким образом, *онтологией называется эксплицитная спецификация концептуализации* [2]. Формально онтология состоит из терминов, организованных в таксономию, их определений и атрибутов, а также связанных с ними аксиом и правил вывода.

Часто набор предложений, составляющих онтологию, имеет форму логической теории первого порядка, где термины словаря являются именами унарных и бинарных предикатов, называемых соответственно *концептами* и *отношениями*. В простейшем случае онтология описывает только иерархию концептов, связанных отношениями категоризации. В более сложных случаях в нее добавляются подходящие аксиомы для выражения других отношений между концептами и для того, чтобы ограничить их предполагаемую интерпретацию. Таким образом, онтология представляет собой базу знаний, описывающую факты, которые предполагаются всегда истинными в рамках определенного сообщества на основе общепринятого смысла используемого словаря.

Приведем наиболее полное определение онтологии.

Определение 2. Онтология – это структурная спецификация некоторой предметной области, ее формализованное представление, которое включает словарь (или имена) указателей на термины предметной области и логические выражения, которые описывают, как они соотносятся друг с другом. Иными словами, онтология представляется тройкой

$$O = \langle X, R, F \rangle,$$

где X – конечное множество концептов (понятий, терминов) предметной области, которую представляет онтология O ;

R – конечное множество отношений между концептами (понятиями, терминами) заданной предметной области;

F – конечное множество функций интерпретации (аксиоматизация), заданных на концептах и/или отношениях онтологии O . Заметим, что единственным ограничением, накладываемым на множество X , является его конечность и непустота.

Таким образом, онтологии обеспечивают словарь для представления и обмена знаниями о некоторой предметной области и множество связей, установленных между терминами в этом словаре.

Построение онтологии (онтологический инжи-

ниринг) является мощным когнитивным инструментом, позволяющим определить наиболее значимые для решения задачи концепты и связи между ними.

Алгоритм онтологического инжиниринга можно представить в виде следующей последовательности операций: выделение концептов – базовых понятий данной предметной области; определение «высоты дерева онтологий» – числа уровней абстракции; распределение концептов по уровням; построение связей между концептами – определение отношений и взаимодействий базовых понятий.

В задачах учета земельных ресурсов множество X содержит такие концепты, как: земельный участок, объект недвижимости, нефтепровод, нефтеперекачивающие станции, узлы подключения станций, очистные сооружения, средства диагностики, и др.

Множество отношений между концептами R – это: назначение ресурса, освобождение ресурса, формирование связи для реализации технологической операции и другие операции, связывающие между собой два или более концептов. И, наконец, функции интерпретации F – это практическая организация межфреймового взаимодействия (например, приведенный на рис. 1 случай).

Следует также отметить, что онтологический подход позволяет также расширять описание предметной области, т.е. делает систему открытой к изменениям, что противоречит концепции разработки баз данных – структура базы должна быть неизменной. Это обстоятельство чрезвычайно важно в задачах автоматизации учета земельных ресурсов, поскольку возможно довольно динамичное расширение как спектра выполняемых задач, так и изменение структуры системы нефтеснабжения.

Онтология предметной области в общем случае представляет собой граф с размеченными вершинами и дугами, то есть пару $\langle T, L \rangle$, где T – множество вершин графа, $T = \{t_1, \dots, t_n\}$, представляющих понятия онтологии, L – множество дуг графа, представляющих связи между терминами онтологии, $L = \{l_1, \dots, l_n\}$. Каждая вершина $t_i \in T$ характеризуется своим именем и типом, $t_i = \{termname, type\}$. Каждая

дуга $l_j \in L$ характеризуется своим именем, множественностью начальной вершины и множественностью конечной вершины.

Иерархия классов в онтологической системе представляет отношение «is-a»: класс A – это подкласс B , если каждый экземпляр B также является экземпляром A . Отношение подкласса транзитивно: если B – это подкласс A , а C – подкласс B , то C – подкласс A .

Свойства включают две основные категории:

- свойства объекта (Object properties), которые связывают индивидуальные элементы между собой;
- свойства типов данных (Datatype properties), которые связывают индивидуальные элементы со значениями типов данных, такими как целые числа, числа с плавающей запятой и строки.

К классам и свойствам могут применяться различные ограничения. Например, ограничения мощности множества указывают на число связей, в которых может участвовать класс или индивидуальный элемент.

Значение слота может зависеть от значения другого слота. Например, если на земельном участке расположено определенное техническое сооружение нефтетранспортной системы, то данное сооружение должно содержать географические координаты, принадлежащие земельному участку. Эти два отношения, «принадлежит» и «размещается», называются обратными отношениями. Излишне хранить информацию и о том, и о другом. Тем не менее, с точки зрения приобретения знаний удобно иметь оба блока информации доступными в явном виде [7].

Многие фреймовые системы позволяют определить для слотов значения по умолчанию. Если значение определенного слота одинаково для большинства экземпляров класса, то можно определить это значение как значение слота по умолчанию. Затем, когда создается каждый экземпляр класса, имеющего этот слот, система автоматически заполняет значение по умолчанию. После этого можно изменить это значение на любое другое, которое позволяют фацеты. То есть, значения по умолчанию созданы

для удобства: в любом случае они не накладывают какие-либо ограничения на модель или никак ее не меняют [7].

3. Семантический анализ предметной области в онтологическом представлении

Одним из главных преимуществ онтологического представления в сравнении с базами данных является возможность построения иерархии классов. В связи с этим актуальным представляется вопрос о сравнении разнотипных объектов, т.е. определения их семантического подобия – функции, принципиально не существующей в базах данных.

Известно множество трактовок понятия подобия. Два объекта могут считаться *подобными*, в случае если они обладают общими характеристиками (или атрибутами). Формализуя понятие подобия, в настоящей работе предлагается ввести функцию семантического. Пусть задана мера подобия – функция с вещественным значением

$$sim(x, y) : S^2 \rightarrow [0, 1]$$

на множестве S семантических концептов онтологии, определяющая степень подобия между x и y .

Очевидно, что функция sim должна быть рефлексивной и симметричной, т.е.

$$\forall x, y \in S \text{ выполняется}$$

1. $sim(x, x) = 1$;
2. $sim(x, y) = sim(y, x)$.

Будем считать, что два объекта e_1 и f_1 являются *семантически подобными*, если существует два семантических пути

$$e_1, P_1, e_2, P_2, e_3, \dots, e_{n-1}, P_{n-1}, e_n \text{ и}$$

$$f_1, Q_1, f_2, Q_2, f_3, \dots, f_{n-1}, Q_{n-1}, f_n$$

семантически соединяющих $e_1 \rightarrow e_n$ и $f_1 \rightarrow f_n$ соответственно, и таких, что для каждой пары свойств P_i and Q_i , $1 \leq i < n$, выполняется одно из условий $P_i = Q_i$ или $P_i \subseteq Q_i$, или $Q_i \subseteq P_i$.

Таким образом, мера подобия ставит в соответствие паре концептов некоторое вещественное чис-

ло от 0 до 1, где 0 (1) присваивается полностью различным (одинаковым) понятиям.

В общем случае метод производит присвоение каждой из категорий понятий, например, класса объектов, специфической меры, которая определяется как функция переменных, значения которых вычисляются по связанным между собой категориям понятий, например, свойство, подкласс, с помощью соответствующего метода. Затем различные компоненты вычисления подобия объединяются с помощью взвешенной суммы. Могут использоваться и другие операторы объединения, но при этом могут возникнуть трудности нахождения решения.

Весы позволяют настраивать важность компонента при оценке подобия, нулевой вес означает полное игнорирование компонента.

Например, для двух классов c, c' в общем случае подобие вычисляется по формуле:

$$Sim_C(c, c') = \sum \omega_i Msim_i(\lambda(c), \lambda(c'))$$

где ω_i – вес соответствующего компонента сравнения, $\sum \omega_i = 1$,

λ – компонент сравнения классов,

$Msim_i$ – среднее значение подобия по одному из компонентов сравнения вида

$$MSim_C(\lambda, \lambda') = \frac{\sum_{(c, c') \in Pairing(\lambda, \lambda')} Sim_C(c, c')}{\max(|\lambda|, |\lambda'|)}$$

В качестве методологии нахождения семантического подобия между объектами в данной работе предложены онтологии.

4. Применение среды описания онтологий *Protégé* для представления логистической системы

Построение онтологий – сложный и занимающий много времени процесс. Чтобы облегчить его, со середины 90-х годов начали создаваться первые среды для процесса разработки онтологий. Одной из наиболее популярных в настоящее время является среда разработки Protege [3, 4].

Protege включает редактор онтологий, позво-

ляющий проектировать онтологии, разворачивая иерархическую структуру абстрактных или конкретных классов и слотов. Структура онтологии сделана аналогично иерархической структуре каталога. На основе сформированной онтологии Protege может генерировать формы получения знаний для введения экземпляров классов и подклассов.

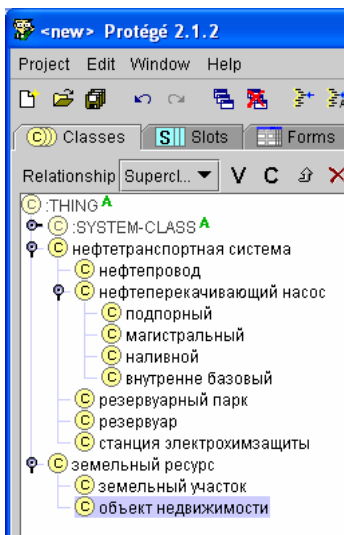


Рис. 2. Иерархия фреймов

На рис. 2 приведен фрагмент иерархии фреймов разрабатываемой электронной системы учета земельных ресурсов.

Модель содержит всю необходимую информацию для учета земельных ресурсов, объектов нефтетранспортной системы и получения оперативной информации о текущем состоянии всех элементов. Слоты фрейма «Нефтепровод» определяют год ввода в эксплуатацию, диаметр, длину на балансе, срок эксплуатации, а инверсионный слот «размещен на» связывает элемент нефтетранспортной системы с множеством земельных ресурсов, на которых он расположен в настоящий момент времени. Аналогичный инверсионный слот имеет и фрейм «Земельный ресурс».

Цепочки инверсионных слотов обеспечивают динамическую связь между фреймами, позволяя автоматически отслеживать связанные элементы автоматизированной системы.

Ввод первичных данных можно осуществлять в системе Protégé во вкладке Instances.

Заключение

Представленный подход позволяет с единых позиций описать структуру базы знаний, необходимой для хранения и извлечения информации о процес-

сах, протекающих в электронной системе учета земельных ресурсов.

Основным преимуществом использования фрейм-модели над традиционными базами данных является более сложные правила обеспечения целостности информации, возможность реализовывать запросы с учетом наследования (что принципиально невозможно для реляционных баз данных), возможность интеграции онтологического описания с другими программными средствами.

Предлагаемая онтологическая структура позволяет решить задачу автоматического формирования земельных ресурсов для нефтетранспортной системы, а также осуществлять запросы о текущем состоянии системы в любой момент времени.

Предложены метрики оценки семантической близости концептов, позволяющие количественно определить эффективность принимаемых решений. Получены практические результаты, позволяющие определить эффективность использования онтологической модели, как более приемлемую для задач поиска.

Литература

1. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем / Учебник для вузов. – С-Пб.: Питер, 2000. – 580 с.
2. Gruber T.R. A translation approach to portable ontologies. Knowledge Acquisition. – 1993. – No. 5 (2). – P. 199-220.
3. Musen M. Domain Ontologies in Software Engineering: Use of Protégé with the EON Architecture // Methods of Inform. in Medicine. – 1998. – P. 540-550.
4. Protege 2000 User Guide [Электронный ресурс]. – Режим доступа: <http://protege.stanford.edu>.

Поступила в редакцию 8.10.2007

Рецензент: д-р техн. наук, проф. А.Ю. Соколов, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.