

УДК 001.2

С.И. Карась

МОДЕЛИ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ И КОГНИТИВНАЯ ПСИХОЛОГИЯ: ДВЕ СТОРОНЫ ОДНОЙ МЕДАЛИ

Моделирование знаний и когнитивную психологию объединяет прикладная предметная область – инженерия знаний. Для извлечения знаний используются индивидуальные и групповые коммуникативные методики, а также не прямые когнитивные методы обнаружения знаний. Полученные знания представляются с помощью ряда моделей: фреймовой, продукционной, сетевой. Важное место в моделировании знаний занимает онтологический подход, который реализуется в двухуровневых информационных системах.

Ключевые слова: извлечение знаний, модели знаний.

Что объединяет моделирование знаний и психологические методы, или что есть медаль? На эту роль претендует прикладная область науки – инженерия знаний (knowledge engineering). Этот термин введен Е. Шортлиффом и Е. Фейгенбаумом как привнесение принципов и инструментария исследований из области искусственного интеллекта в решение трудных прикладных проблем, требующих знаний экспертов [1]. Центральным моментом инженерии знаний является создание базы знаний, т.е. извлечение экспертных знаний и их представление в оптимальной модели. Для извлечения знаний используются коммуникативные и когнитивные методы, с помощью которых системный аналитик (инженер по знаниям) помогает эксперту формализовать необходимые сведения.

Традиционно из широкого ассортимента психологических методов в инженерии знаний применяются различные виды диалогов (от ознакомительного опроса в начале разработки до критического обзора ближе к ее завершению) и групповые методы (от мозгового штурма до заочной Дельфийской методики) [2, 3]. Когда эти прямые методы не дают желаемого результата, используются не прямые методы извлечения знаний.

Не прямые методы близки к одному из направлений психологии – психо-семантике, которая занимается реконструкцией индивидуальных семантических пространств и основана на разных формах субъективного шкалирования [3]. Экспертом в определенной шкале оцениваются смысловые расстояния между элементами знаний. Эксперт отмечает различия и степень сходства между концептами, атрибутами, между элементами знаний разного типа и некоторым заданным элементом. Построение семантического пространства эксперта – это переход на язык более высокого уровня абстракции, хотя применяемые для этого методы инженерии знаний достаточно просты.

К непрямым методам инженерии знаний относятся различные виды сортировок – по отношению эксперта к концептам в рамках решения определенного круга задач или по взаимоотношениям концептов предметной области между собой с построением иерархической диаграммы кластеров концептов. В результате применения этих методов можно построить концептную

карту предметной области, смоделировать потоки данных, жизненный цикл объектов или последовательность действий (построить сети Петри).

Как не прямой способ исследования знаний широко используется матричный подход. Метод матрицы атрибутов заключается в количественной оценке экспертом связи между каждым из набора атрибутов и каждым объектом предметной области. Метод репертуарных матриц (repertory grids) предложен автором теории личностных конструктов Дж. Келли. Целью его применения является выявление персональных конструктов эксперта, т.е. тех признаков или свойств, по которым объекты сходны между собой и, следовательно, отличны от других объектов. Конструкт – некоторая ось, относительно которой располагаются подмножества элементов, количественно связанных с конструктом. При этом объекты связаны определенным контекстом — репертуаром предметной области, а эксперт оценивает каждый объект предметной области по каждому конструкту. Если аналитик изменяет репертуар (набор элементов знаний), то выявляются конструкты разного уровня общности [4].

Перечисленные методы направлены на извлечение экспертных знаний, но не менее важным аспектом инженерии знаний является их моделирование. Среди традиционных моделей представления знаний, имеющих психологические корни, можно выделить следующие:

1. Модель фреймов. Марвин Минский называл фреймом иерархически организованную структуру памяти человека, являющуюся основной для хранения информации [5]. В инженерии знаний фреймом называется структура для описания стереотипной ситуации, состоящая из характеристик этой ситуации, называемых слотами, и их значений. В этой модели знания человека представлены относительно большими единицами со сложной внутренней структурой. Каждый фрейм может быть дополнен информацией о способах его применения, последствиях этого применения, действиях, которые надо выполнить в определенных ситуациях. Каждый слот в качестве значения может содержать не только числовую или символьную константу, но и функцию, правило, с помощью которого можно найти значение слота, приказ на выполнение программы, ссылку на другие фреймы и слоты того же фрейма и т.п. Благодаря этому организуется сеть фреймов, в явном виде вводится иерархическая структура, основанная на различной степени абстракции характеристик объектов и ситуаций.

2. Сетевая модель имеет аналогом структуру долговременной памяти человека и пригодна для представления знаний любых типов. Она начала развиваться в виде семантических сетей с целью моделирования естественного языка. В первой модели такого рода (модели Куиллиана) сетевая структура использовалась для представления семантических отношений между концептами (словами, обозначающими некоторые понятия предметной области) [6]. Обычно используемая для моделирования знаний, семантическая сеть состоит из двух множеств: узлов и соединяющих узлы ребер, которые отражают бинарные отношения между ними [7]. В качестве узлов в модели используются объекты предметной области, их свойства и значения, а также события, процессы, явления. Ребра могут иметь или не иметь тип, направление и количественную оценку отношений между узлами. Главное преимуще-

щество сетевой модели — способность выражения достаточно тонких смысловых оттенков знаний.

3. Продукционная модель как способ представления знаний предложена Э. Постом в 1943 г. Основу этой модели составляет множество так называемых продукционных правил, которые применяются в тех областях, где значительная часть знаний основана на опыте специалистов, эмпирических закономерностях и ассоциациях [8, 9]. Модель базируется на логике «из ситуации следует действие», представляет систему правил вида «Если..., то...» и обеспечивает способ представления выводов, рекомендаций, указаний, стратегий. Эта модель широко используется в слабоформализованных областях.

С конца прошлого века активно разрабатывается онтологический подход к моделированию знаний. С одной стороны, онтология — раздел метафизической философии, в котором рассматриваются всеобщие основы, принципы бытия, его структура и закономерности. С другой — онтологией называются представленные на некотором языке знания о предметной области [10]. Онтология — формальная система понятий предметной области, их свойств, отношений между ними и правил операций над ними. Онтология является моделью предметной области и может служить основой баз знаний и баз данных.

Для чего разрабатываются онтологии и что они дают? Главная цель создания онтологии — стандартизация и максимальная формализация моделей знаний. Онтология облегчает:

- совместное использование людьми или программными агентами структуры информации о предметной области;
- совмещение разных частей модели предметной области;
- изменения и доступность моделей для новых пользователей;
- отделение структуры знаний предметной области от конкретных данных.

Можно выделить несколько этапов разработки онтологии. После определения предметной области надо осуществить поиск существующих онтологий и рассмотреть возможность их повторного использования. Многие онтологии доступны в электронном виде и могут быть импортированы в среду проектирования новой онтологии.

Затем важно получить полный список терминов предметной области, разработать иерархию классов и описать их свойства. Следующим шагом является описание отношений между элементами онтологии. На этом этапе важно определиться, на каком уровне иерархии находится понятие, т.е. является ли оно классом объектов, конкретным объектом или свойством одного из объектов? Глубина детализации определяется приложением онтологии. Последним этапом является создание отдельных экземпляров классов путем ввода конкретных значений их свойств.

Онтологический подход к моделированию предметной области дает возможность разрабатывать надстроечные информационные системы и обеспечить интероперабельность первичных информационных систем. В частности, в настоящее время развивается парадигма двухуровневых информационных систем [11, 12].

В стандартных одноуровневых информационных системах конкретные данные согласованы с объектной моделью. Примером двухуровневого подхода является стандарт разработки информационных систем OpenEHR [13]. Одной из основных целей OpenEHR является объединение медицинских записей, созданных в информационных системах многих учреждений, в одной электронной карте пациента.

На первом уровне стандарт OpenEHR опирается на стабильную базовую информационную модель-онтологию, называемую референтной моделью. Эта модель согласована с моделью архетипов. Архетип является частной моделью отдельных элементов знаний (медицинских терминов, процессов, явлений), соответствует реальным потребностям медицинских работников и может изменяться медицинскими аналитиками. Конкретные медицинские данные представлены на втором уровне системы в соответствии с референтной моделью и архетипами первого уровня. Такая архитектура значительно уменьшает вариабельность используемых данных в разных информационных системах, позволяет передавать информацию между системами и совмещать электронные записи пациентов.

Но без объекта моделирования процесс моделирования невозможен. Методической основой извлечения экспертных знаний являются когнитивные и коммуникативные психологические методы. И в процедуре «добычи» знаний до сих пор не решены многие проблемы. Так, в настоящее время отсутствуют верифицированные методы извлечения и представления скрытых экспертных знаний, которые являются исключительно важными для разработки экспертных систем.

Представляется, что разработка современных и доступных для аналитиков психологических методов извлечения знаний ускорит и увеличит эффективность разработки баз знаний и интеллектуальных компьютерных программ.

Литература

1. Shortliffe E.H. Knowledge engineering for medical decision making / E.H. Shortliffe, D.G. Buchanan, E.A. Feigenbaum // Rev. Computer-based Clin. Decision Aids. 1979. Vol. 67, № 9. P. 1207-1223.
2. Приобретение знаний / С. Осуга, Ю. Сазки (ред.). М.: Мир, 1990.
3. Червинская К.Р. Медицинская психодиагностика и инженерия знаний / К.Р. Червинская, О.Ю. Щелкова. СПб.: Ювента, 2002.
4. Франсella Ф. Новый метод исследования личности: руководство по репертуарным личностным методикам / Ф. Франсella, Д. Баннистер. М.: Прогресс, 1987.
5. Минский М.Л. Фреймы для представления знаний / М.Л. Минский. М.: Энергия, 1979.
6. Quillian M.R. Semantic memory / M.R. Quillian // Semantic Information Processing / Ed. by M. Minsky. Cambridge, Massachusetts: MIT Press, 1968. P. 216-270.
7. Кузнецов П.П. Семантические представления / И.П. Кузнецов. М.: Наука, 1986.
8. Уотермен Д. Руководство по экспертным системам / Д. Уотермен. М.: Мир, 1989.
9. Уэно Х. Представление и использование знаний / Х. Уэно, М. Исидзука (ред.). М.: Мир, 1989.
10. Кяецев А.С. Определение структурных свойств онтологий // Изв. РАН. Теория и системы управления. 2008. №2. С. 69-78
11. Алпатов А.П. Госпитальные информационные системы: архитектура, модели, решения / А.П. Алпатов, Ю.А. Прокопчук, В.В. Костра. Днепропетровск: УГХТУ, 2005.
12. Гусев А.В. Медицинские информационные системы / А.В. Гусев, Ф.А. Романов, И.П. Дуданов, А.В. Воронин. Петрозаводск: Изд-во ПетрГУ, 2005.
13. ISO 18308 Requirements for an Electronic Health Record Reference Architecture [Электронный ресурс]. Режим доступа: <http://www.openehr.org/standards/iso.html>