

процедура анализа графической интерпретации совокупности схем БД, которая формирует словарь терминов, фиксирует связи между ними и формирует семантическую сеть.

Чтобы семантическая сеть была описанием деятельности некоторой большой системы, необходимо ее структурировать: разбить на фрагменты, упорядочить их иерархически так, чтобы следующий уровень описания детализировал предыдущий. Для структурирования семантической сети совокупность терминов разбивается на кластеры. К анализу ограниченного множества терминов кластера вторично привлекается эксперт. Машина предлагает ему информацию для формирования фрагментов модели ПРО: статического (элемента тезауруса) или динамического (ситуационного фрейма), их иерархического упорядочения.

Таким образом, интегрируя представления отдельных пользователей о данных и используя методы передачи содержания гипертекста с помощью семантической сети, мы формируем модель ПРО. Эксперт привлекается дважды: для семантического насыщения структур описания БД и для выяснения внутренней организации семантической сети. В последнем случае сформированные фрагменты сети описывают основные аспекты деятельности рассматриваемой ПРО, а их упорядоченная последовательность - модель ПРО.

Формирование декларативных знаний (модели ПРО) в составе действующих информационных систем позволит включить в их состав экспертные системы, что обеспечит современный научно-технический уровень их эксплуатации.

В целом разработанные методы извлечения и представления знаний могут быть применены для создания систем искусственного интеллекта.

Абламейко С.В., Волочкович С.А.

ИСПОЛЬЗОВАНИЕ ЗНАНИЙ ПРИ РАСПОЗНАВАНИИ ГРАФИЧЕСКИХ ИЗОБРАЖЕНИЙ

Целью данной работы является разработка методики распознавания графических изображений на основе использования априорных знаний об объектах изображения и их отношениях, организованных в виде базы знаний. Под объектом в дальнейшем будем понимать не только подлежащие распознаванию объекты графического документа, но и производные элементы, их описывающие. В качестве отношений между объектами будем рассматривать прежде всего топологические отношения: 1) "пересечение" объектов, т.е. ситуация, когда изображения объектов имеют хотя бы одну общую точку; 2) "соседство" объектов, т.е. попадание по крайней мере одной точки одного объекта в некоторую ϵ -окрестность другого объекта.

Предполагается организация базы знаний в виде иерархической семантической сети, вершинами которой соответствуют объекты графического изображения, а ребрам - отношения между ними. База знаний включает три уровня представления знаний:

1) поверхностный детальный уровень, содержащий описание параметров производных элементов, описывающих объекты графического документа; правила выделения производных

элементов из растрового представления исходного изображения;

2) синтаксический уровень, включающий описания простых объектов изображения (отдельных знаков) в терминах производных элементов и их отношений;

3) семантический уровень, содержащий описания сложных объектов изображения (например, надпись, состоящая из букв) в терминах простых объектов изображения и их связей.

Система распознавания графического изображения содержит следующие элементы:

1) система управления базой знаний, которая осуществляет выдвижение и проверку гипотез и на основе результатов проведенного анализа корректирует базу знаний;

2) база знаний (интенциональная семантическая сеть), содержит знания о составе графического документа, т.е. об объектах графического документа, их взаимном расположении и правилах выделения их из растрового представления графического изображения;

3) база данных (экстенциональная семантическая сеть) содержит описание уже распознанных на данный момент объектов графического документа, а также информацию об их взаимном расположении.

Сегментация исходного растрового представления с использованием базы знаний ведется целенаправленно путем выдвижения и проверки гипотез. Выдвижение гипотез осуществляется системой управления базой знаний с учетом 1) априорных знаний об общих закономерностях построения графических изображений заданного класса, 2) знаний об уже распознанных на данный момент объектах изображения.

Проверка гипотез проводится как целенаправленный поиск производных элементов изображения, подтверждающих данную гипотезу. Если гипотеза не подтверждается, все результаты, связанные с проверкой ее, аннулируются, что позволяет избежать потерь информации.

Область приложения данного подхода - обработка и распознавание изображений графических документов (чертежей, карт).

Основными результатами работы являются:

- организация базы знаний;
- алгоритм использования знаний для распознавания графических изображений.

Дробышев Ю.П., Скороходько Э.Ф.,
Темперанский В.А.

АВТОМАТИЗИРОВАННЫЙ СЛОВАРЬ ТЕРМИНОВ: КОНЦЕПЦИЯ, ПРОЕКТИРОВАНИЕ, РЕАЛИЗАЦИЯ

Терминологический словарь представляет собой языковое отображение системы понятий той или иной отрасли знания. От степени адекватности описания этой отрасли знания в словаре немало зависит эффективность использования последнего в качестве инструмента познания и научного исследования. Тем не менее традиционные методы лексикографической работы принципиально не позволяют полностью избавиться от многих