

ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

Э. Н. МАНУКЯН, А. Р. МАНВЕЛЯН

АРХИТЕКТУРА И ВОЗМОЖНОСТИ НЕДЕТЕРМИНИРОВАННОЙ
БАЗЫ ДАННЫХ, ПОДДЕРЖИВАЕМОЙ В ДИАЛОГОВОЙ
ИНФОРМАЦИОННОЙ СИСТЕМЕ «АНИ»

Первейшей задачей при разработке баз данных (БД) является проблема генерации логической структуры данных (ЛСД) БД. При этом информация, описывающая конкретную предметную область (ПО), терпит следующие четыре этапа преобразований [1].

1. На основе тщательного анализа ПО формируется ее описание.

2. Формируется полученное описание и составляется информационно-логическая структура (ИЛС), описывающая систему реальных объектов ПО [2].

3. Составляется ЛСД конкретной БД.

4. Определяется структура размещения данных на физических носителях ЭВМ.

Не располагая специальными средствами для описания реальных объектов, разработчики БД вынуждены пользоваться языками описания конкретных систем. В результате чего разработанная ЛСД из-за ограничений, присутствующих в конкретных системах, и влияния субъективных факторов, отражает лишь некоторое искаженное подмножество данных и информационных связей ПО. Это приводит к трудностям при общении пользователей с системой, неэффективности использования памяти и увеличению отклика системы.

Решением данной проблемы, на наш взгляд, является разработка архитектуры БД и языка описания ЛСД, позволяющих максимально приблизить ЛСД конкретной БД к ИЛС ПО, тем самым упростить реализацию этапа 3.

Эти соображения и ряд других легли в основу конструирования системы управления базами данных (СУБД) диалоговой информационной системы «АНИ», которая разрабатывается в ЕрПИ. Ниже приводится описание организационных единиц и архитектура БД обслуживаемых СУБД «АНИ». В БД выделяются три функционально различных раздела: раздел данных (РД), знаний (РЗ) и словаря (РС).

Раздел словаря является хранилищем идентификаторов организационных единиц логической структуры (ЛС) РД, РЗ и их синонимов.

РС используется системой для обеспечения интерфейса между пользователями и БД, в которой идентификаторы представлены в закодированном виде.

Логическая структура раздела данных имеет вид ориентированного мультиграфа. Организационными единицами, с помощью которых создается граф, является «объект исследования» (ОИ), соответствующий вершинам графа, «отношение», соответствующее дугам графа и показывающее кем или чем является один «объект исследования» для другого, и «атрибут», являющийся описанием некоторого свойства ОИ и набором элементов.

Вершинами графа ЛСД могут быть представлены реальные объекты, факты, а дугами — реально существующие отношения между ними. Системой не ставятся ограничений на структуру мультиграфа, на количество и типы «атрибутов» и их элементов, а также обеспечивается возможность хранения описаний структур данных вместе с их значениями, что позволяет создавать и вести недетерминированные базы данных [3]. Система обладает средствами идентификации организационных единиц по датам их «рождения» и «смерти», что обеспечивает возможность имитировать ПО как «в пространстве», так и «во времени».

При ведении РД системой различаются отношения трех типов: простые, основные и определяющие. Отношение называется определяющим если им обуславливается существование ОИ в БД. Основные отношения обратно направлены определяющим. По определяющим (основным) отношениям ЛСД представляется набором несвязанных древовидных графов. Добавление ОИ в базу происходит по запросам пользователей путем создания определяющего отношения. При удалении ОИ удаляются все ОИ, связанные с ними по определяющим отношениям. Простым отношением могут быть связаны любые два ОИ. Таким образом, введение трех типов отношений позволяет контролировать создание и удаление ОИ в РД. Это становится особенно необходимым в БД со сложной логической структурой.

Раздел знаний служит для хранения описаний основных понятий ПО и отношений между ними. Логическая структура РЗ представляется в виде ориентированного псевдографа. Организационными единицами РЗ являются «понятия», «отношения» между «понятиями» и «атрибуты понятий».

Поскольку структура знаний отражает элементы объективной действительности и не зависит от конкретной ПО, необходимо поддерживать определенные закономерности в логической структуре РЗ, поэтому ведение РЗ происходит под жестким контролем системы.

В РЗ различаются две «категории понятий», описывающие:

а) свойства ОИ РЗ и идентифицирующиеся именами соответствующих атрибутов;

б) набор ОИ, обладающих свойствами, которые удовлетворяют определенным условиям. Члены такого набора называются экземплярами.

первой категории отношения «объекты с частным проявлением свойства» (тип G) и «возможное свойство» (тип F).

Указанные отношения не зависят от ПО и отражают объективно существующие отношения между понятиями реального мира. Для представления взаимоотношений между реальными объектами и системой разрешается соответствующие им «понятия» связывать отношениями, имена которых могут совпадать с любым существующим в РЗ «понятием» второй категории. Например, на рис. отношение 7. Основной объем знаний в предметной области хранится в виде атрибутов «понятий». При описании РЗ атрибуты и отношения между «понятиями» могут быть объявлены как «обязательные» и «возможные». Создание РЗ производится путем постепенного наращивания ЛС. При манипуляции «понятиями» второй категории различаются полуопределенные и конечно-определенные «понятия». В отличие от экземпляров конечно-определенных «понятий» экземплярам полуопределенных «понятий» разрешается добавлять атрибуты и связи, не противоречащие содержащимся в РЗ знаниям. При этом производится соответствующая коррекция РЗ.

Для организации связей между разделами БД важную роль играет раздел словаря. При ведении БД автоматически создаются и поддерживаются связи между основными словами и соответствующими им «понятиями» в ОП. Они широко используются системой для обеспечения доступа к требуемым данным. Системой автоматически может поддерживаться связь между некоторыми понятиями и их экземплярами (тип H). Например, на рис. связи 14 и 21.

Система с предлагаемой архитектурой не ставит ограничений на сложность ЛС БД, а предложенные организационные единицы позволяют адекватно описать информационную структуру системы реальных объектов и понятий ПО. Четко выделенный РЗ позволяет:

а) уменьшить объем РД за счет выделенного хранения у «понятий» общих для их экземпляров данных;

б) обеспечить контроль и допущение запросов пользователей.

Возможность автоматической организации системой отношений типа H позволяет создавать инверсные формы хранения данных в виде триад [4] и использовать некоторые элементы реляционного исчисления при выделении информации на РД. При наличии отношений типа H у всех «понятий» РЗ, ЛС БД можно рассматривать как инвертированную реляционную структуру [5]. При полном отсутствии таких отношений ЛС БД представляется сетевой. В общем случае системой обеспечивается несколько различных путей к конкретному ОП через РД и РЗ, что позволяет в зависимости от конкретной ситуации выбрать стратегию поиска, обеспечивающую максимальную эффективность ассоциативного доступа. Таким образом, путем автоматического выбора системой подмножества «понятий» и формирования у них связей типа H достигается и поддерживается оптимальность ЛСД БД. Такая архитектура БД позволяет не только решать вышеописанные проблемы, но и создает

Է. Ն. ՄԱՆՈՒՅԱՆ, Ա. Ռ. ՄԱՎԵՅԱՆ

**«ԱՆՆԻ» ԳԻՒՂՈՒՄՆԻՆ ԻՆՏԵՐՈՒԱՑԻՈՆ ՇԱՄԱԿԱՐԳՈՒՄ ԱՊԱՀՈՎՎՈՂ
ՉԳԵՏԵՐՄԻՆԱՑՎԱՆ ՏՎՅԱԸՆԵՐԻ ԲԱԶԱՅԻ ԿԱՌՈՒՑՎԱԾԲԸ
ԵՎ ՇՆԱՐԱՎՈՐՈՒԹՅՈՒՆՆԵՐԸ**

Ա մ փ ո ւ փ ո ս մ

Գիտարկվում են խնդիրներ, որոնք առաջանում են տվյալների բազաների տրամաբանական կառուցվածքների նախադժման էտապում: Առաջարկվում է տվյալների բազայի այնպիսի կառուցվածք, որի ապահովման դեպքում տվյալների բազայի դեկոմպարման համակարգի կողմից հնարավորություն է ընձեռնվում մոդելավորել արտաքին առարկայական աշխարհը տվյալների բազայում, ստանալ ճկուն և ավտոմատ կարգավորվող տրամաբանական կառուցվածք: Նշված հատկությունները ապահովվում են տվյալների բազայից զիտելիքների և տվյալների ենթարազանների առանձնացման և վերջիններիս միջև հատուկ ինֆորմացիայի կապերի ստեղծման շնորհիվ:

Л И Т Е Р А Т У Р А

1. Михаловский С. А. Описание внешнего пространства данных автоматизированных информационных систем. «Управляющие системы и машины», 1979, № 6, с. 22—23.
2. Ерещи-Ерещенко А. А., Михаловский С. А. Уточнение описания некоторых процедур семантического анализа описания исходных данных при проектировании автоматизированных информационно-справочных систем. Сб. «Вопросы проектирования банков данных», Киев, «Наукова думка», 1977, с. 43—55.
3. Михаловский В. С. и др. Интегрированная система организации данных для решения задач анализа транспортных сетей. Киев, Изд. ИК АН УССР, 1977, 30 с.
4. Маруля Дж. Организация баз данных и вычислительных систем. Москва, «Мир», 1978.
5. Tamis R. Le Viet C. et al. A file organization suitable for relational database operations, lect. Notes Romput. Sci., 1979, 75, p. 193—227.