

ПРОГРАММА ГЕНЕРАТОРА ОТЧЕТОВ И ДРУГИЕ УТИЛИТЫ ИЕРАРХИЧЕСКОЙ БАЗЫ ДАННЫХ

В Институте общей физики АН СССР проводятся работы по созданию математического обеспечения для систем управления базами данных. Описанную в настоящей работе базу данных следует отнести к классу реляционных, то есть данные, хранящиеся в базе, представляются пользователю в виде двумерных таблиц (отношений) [1]. Описываемая база данных является дальнейшим развитием линейной БД; термин "иерархическая" указывает на способ физической организации данных в виде цепочки иерархически подчиненных файлов.

ФИЗИЧЕСКАЯ ОРГАНИЗАЦИЯ БД

Иерархический файл (ИФ) базы данных физически организован как набор файлов операционной системы — информационных и служебных (описания структуры ИФ). ИФ может состоять из нескольких информационных тел (до 5), которые могут располагаться на нескольких различных устройствах прямого доступа. Информационное тело может содержать до 32767 физических записей (фиксированной длины); размер записи ограничен размером дискового блока (512 байт). Одно и то же информационное тело может использоваться как часть разных ИФ, а сами ИФ могут образовывать сколь угодно сложную сеть, покрывающую информационные тела. Таким образом ИФ группирует в себе информацию, необходимую для конкретного приложения.

Связи между информационными телами поддерживаются не с помощью указателей, а с помощью самих данных, хранимых в файле. В старших файлах выделяются поля связи (возможно, состоящие из нескольких фрагментов), в подчиненных файлах выделены поля, составляющие первичный ключ. Для какой-либо записи старшего файла "сыновними" являются те записи подчиненного файла, для которых начало первичного ключа совпадает со значением поля связи этой "родительской" записи. Такой механизм связи информационных тел обеспечивает большую гибкость в управлении данными и поддерживается разработанными механизмами доступа к ИФ.

Механизм доступа к ИФ позволяет, в частности, обращаться к записи без указания ее точного ключа. К каждому информационному файлу, независимо от других тел, может осуществляться прямой или индексный доступ. Индексные (инвертированные) файлы строятся с использованием указателей на записи основного информационного файла.

Описание ИФ состоит из описания цепочки подчиненных файлов, образующих данный ИФ. Для каждого информационного тела имеется описание используемой подсхемы (имена, типы и размеры полей записи) и описание физической структуры тела (имя информационного тела, длина записи в нем, положение первичного ключа и т. п.).

Обработка информации, хранящейся в базе данных, обеспечивается набором утилит, описанных ниже. Все утилиты работают со стандартными описаниями ИФ. Имеется языковой интерфейс (Фортран-5) к БД для написания прикладных программ, но необходимость в них отпадает благодаря использованию программы генератора отчетов HRPG.

УТИЛИТА HEDIT — РЕДАКТОР БАЗЫ ДАННЫХ

Утилита HEDIT предназначена для занесения информации в файлы иерархической базы данных. Основные функции: обновление данных (изменение содержимого записей), добавление новых записей, удаление и копирование записей, просмотр файла базы данных.

Доступ к записям БД осуществляется по номеру записи или по ключу. Адресация по номеру записи может быть абсолютной или относительной (от текущей записи БД). Ключ при адресации записи может иметь один из следующих типов: точный, приближенный с поиском вперед или приближенный с поиском назад по базе.

При редактировании записей БД допускаются следующие направления перемещения по иерархическому файлу: вперед (к следующей записи), назад (к предыдущей записи), вверх (к родительской записи), вверх-вперед (к следующей, родительской записи), вниз (к первой, сыновней записи по какой-либо связи).

HEDIT является экраным редактором, команды редактирования задаются нажатием специальных клавиш или контрольными символами. Кроме обычных клавиш управления курсором используются клавиши для быстрого перемещения курсора в пределах записи по так называемым точкам табуляции (обычно они совпадают с границами полей в записи), клавиши вставки и удаления символа в пределах поля записи, клавиши удаления информации от текущей позиции курсора до конца поля.

При переходе в режим установки точек табуляции оператор получает возможность установить точки табуляции по своему усмотрению, в том числе он может удалять некоторые точки табуляции или добавлять новые. Одной из команд редактора является установка точек табуляции на начала полей записи.

В настоящее время редактор HEDIT реализован для дисплеев типа "Видеотон-340" и Dasher D200. Предполагается адаптирование редактора для дисплеев типа "Электроника-15ИЭ" и расширение функций экранного редактирования записей с помощью функциональных клавиш (на дисплеях двух последних типов).

УТИЛИТА HRPГ – ГЕНЕРАТОР ОТЧЕТОВ

Утилита HRPГ – это мощное и гибкое средство, позволяющее получать различные отчеты по информации, содержащейся в файлах иерархической БД. Форма и содержание получаемого отчета задается прикладным программистом в виде программы для утилиты HRPГ. Утилита закидывает HRPГ-программу из файла в оперативную память и фактически интерпретирует ее.

Пользователю нет необходимости самому программировать открытие иерархического файла БД и операции ввода-вывода с ним – эти механизмы встроены в саму утилиту HRPГ. Пользователю необходимо запрограммировать лишь вычисления и формат выдачи отчета.

Возможности, предоставляемые HRPГ:

- печать нескольких копий отчета;
- направление отчета в любой файл, а не только на принтер;

- обработка области ИФ, а также обработка с печатью лишь части отчета (используется при перепечатке части страниц отчета).

Программирование на HRPГ предполагает использование механизма контрольных полей. Программист может объявить несколько сегментов иерархической записи (до восьми) контрольным полем; HRPГ допускает неограниченную степень вложенности контрольных полей. Если при чтении очередной записи из ИФ значения некоторого контрольного поля сменятся, утилита автоматически выводит в отчет требуемые подзаголовки и вы-

полняет запрограммированные вычисления для этого и всех вложенных в него контрольных полей (подсчет итогов по контрольным полям).

С помощью HRPГ-программы можно запрограммировать:

- четыре арифметические действия с арифметическими переменными, сравнение арифметических переменных с 0 (арифметические операции выполняются над операндами с двойной точностью – 12 десятичных знаков);

- операции со строковыми переменными и константами, в том числе выделение подстроки;

- ввод-вывод арифметических и строковых переменных из тела записи ИФ или с консоли оператора;

- декодирование закодированных полей записи;

- задание текущей даты и времени суток;

- задание глубины обработки ИФ (то есть размера цепочки информационных тел).

В дальнейшем предполагается реализовать возможность отбора для обработки лишь части записей из базы данных, удовлетворяющих определенным условиям (функция языка запросов).

HSORT – ПРОГРАММА СОРТИРОВКИ ИЕРАРХИЧЕСКИХ ФАЙЛОВ БД

Программа позволяет осуществлять сортировку ИФ по нескольким сегментам. Сегментом может служить поле, часть поля записи БД или фрагмент целой записи, указываемый с помощью номера первого байта в записи и длины фрагмента в байтах.

HSORT позволяет осуществлять сортировку по убыванию или возрастанию значений сегментов в соответствии с произвольным, задаваемым пользователем "алфавитом сортировки".

ИФ состоит, как правило, из нескольких информационных тел, составляющих цепочку. Сыновние файлы такой цепочки должны быть упорядочены особым образом для обеспечения правильного доступа к ним. Программа HSORT обеспечивает "правильный порядок" записей во всех сыновних информационных файлах с учетом заказа пользователя на сортировку.

HSORT поддерживает целостность описаний и состояние информации в файлах и в некоторых случаях пересортировывает даже те информационные тела, сортировку которых пользователь не затребовал явно. Кроме этого, на стадии сортировки файла происходит также физическое исключение записей, удаленных с помощью редактора HEDIT.

ПЕРСПЕКТИВЫ РАЗВИТИЯ ИЕРАРХИЧЕСКОЙ БД

В настоящее время производится реализация программы HVERIFY. В ее функции будет входить

проверка состояния информации в ИФ, включая обнаружение дисковых сбоев, приведших к порче информации, и проверку семантической правильности информации в полях различных типов (например, проверка того, что в поле ФИО содержатся русские буквы фамилии, пробел и инициалы, разделенные точками). Эта утилита призвана полностью решить проблему целостности данных в БД.

В дальнейшем планируется разработка программы NQUERY, реализующей язык запросов в базу данных. Запрос будет представлять собой логическое выражение, где термами служат арифметические и строковые выражения, соединенные операторами сравнения. Термами арифметических и строковых выражений служат значения полей из запи-

сей иерархической БД и константы соответствующего типа.

Например:

```
ФИО = "ЛЮБИН А. А." & ЦЕНА-ИЗНОС <
< = 600.00
```

Описанная в докладе иерархическая база данных реализована (на языке Фортран-5) и используется при автоматизации общих отделов Института общей физики АН СССР [2]. База данных эксплуатируется на ЭВМ Eclipse S/200 в операционной системе RDOS фирмы Data General. Общий объем информации, содержащейся в базе, оценивается в 20 Мбайт.

ЛИТЕРАТУРА

1. Мартин Дж. Организация баз данных в вычислительных системах. — Мир, 1980.
 2. Кузнецов М. М., Славинский И. Н. ВР — база данных для автоматизации общих отделов (Руководство оператора) : Препринт. — М.: ИОФАН, 1986.
-