

сообщения
Объединенного
института
ядерных
исследований
Дубна

48 84071

P10-84-439

4574/84

А.И.Островной

МЕТОДИКА ПРИМЕНЕНИЯ БАЗЫ ДАННЫХ
В СИСТЕМАХ АВТОМАТИЗАЦИИ ЭКСПЕРИМЕНТОВ
И ЕЕ РЕАЛИЗАЦИЯ

1984



ВВЕДЕНИЕ

Одним из важнейших направлений развития в производстве прикладного программного обеспечения является его интеграция^{/1/}. Она призвана облегчить передачу данных между различными программами или комплексами программ, свести к минимуму ручные операции ввода результатов работы одной программы в качестве исходных данных для работы другой.

В области создания программного обеспечения для автоматизации научных экспериментов, в частности, ядерно-физических, результатом интеграции должно явиться устранение разрыва между программами накопления и программами обработки в режиме off-line. Например, в спектрометрии программы накопления обеспечивают сбор, накопление и сохранение основной спектрометрической информации и части дополнительной, необходимой для обработки накопленных данных. Основная спектрометрическая информация - это большие массивы однородных данных, поступивших от детектора. Дополнительная информация характеризует условия получения и качество основной. Она включает данные о времени измерения, экспериментальных условиях, мощности реактора и др. Часть дополнительной информации не фиксируется аппаратурой и экспериментатор записывает ее в журнале измерений.

Вся накопленная информация /и основная, и дополнительная/ поступает на обработку, выполняемую в режиме off-line. И если основную информацию передать на обработку достаточно просто /обычно это файлы на магнитных носителях/, то дополнительную, вследствие ее разнородности и различных способов хранения, физикам зачастую приходится вводить вручную /набивать на перфокартах или вводить в интерактивном режиме/.

В данном сообщении предлагается методика применения базы данных в программном обеспечении систем автоматизации экспериментов. База данных используется в качестве основы для интеграции программного обеспечения эксперимента, обеспечивает унифицированное представление дополнительной информации и позволяет автоматизировать ее накопление. В работе описаны также программные средства, реализующие эту методику. Они составляют пакет подпрограмм на языке Паскаль^{/2/} на ЭВМ семейства СМ-3, СМ-4 под управлением операционной системы РАФОС^{/3/}.

Описанные методика и программные средства использованы в системе автоматизации экспериментов на нейтронном дифрактометре^{/4/} в Лаборатории нейтронной физики ОИЯИ. Они могут применяться так-

же при создании программного обеспечения систем автоматизации экспериментов в других областях.

1. МЕТОДИКА ИСПОЛЬЗОВАНИЯ БАЗЫ ДАННЫХ В СИСТЕМАХ АВТОМАТИЗАЦИИ

Предлагаемая методика состоит в том, что база данных используется как центральный элемент, позволяющий объединить программы накопления и программы обработки в единый комплекс. База данных обеспечивает функции интерфейса между этими программами. Во время проведения измерений она используется для накопления дополнительной информации, характеризующей файлы с основной экспериментальной информацией. По окончании измерений база данных позволяет проводить автоматическую обработку накопленных данных, извлекая для каждого файла с основной экспериментальной информацией дополнительные параметры, нужные при обработке. Поиск и чтение информации из базы данных можно выполнять в программе обработки, вызывая соответствующие подпрограммы, либо создав специальную программу для извлечения информации из базы данных и передачи ее программам обработки.

База данных может также использоваться для хранения истории работы системы автоматизации, информации о действиях оператора и сбоях оборудования, а в дальнейшем - для анализа качества программно-аппаратной системы, поиска причин случайно возникающих сбоев и т.п.

Для организации базы данных и работы с ней, в принципе, можно было бы использовать имеющиеся системы RMS-11 или DBMS-11, но они, в силу своей универсальности, занимают много ресурсов ЭВМ, требуют использования многопользовательских и многопрограммных операционных систем RSX-11M или IAS. Это делает невозможным их применение для реализации предложенной методики. К тому же она и не требует использования сложных структур данных и развитых средств для работы с базами данных. Нужны компактные программные средства, которые можно включать в программы автоматизации конкретных экспериментов, созданных на базе мини- и микроЭВМ.

Предложенный способ использования базы данных и задачи, для решения которых она должна использоваться, ограничивают возможность применения традиционных методов организации базы данных. Например, нельзя требовать организации службы администратора базы данных^{/5/}, нельзя создавать единую систему управления базой данных. Наиболее предпочтительным способом реализации средств для образования базы данных и работы с ней является пакет подпрограмм, ориентированных на описанный способ применения. Проблемная ориентация пакета позволит сделать его компактным и использовать в системах автоматизации на мини- и микроЭВМ.

2. ОРГАНИЗАЦИЯ БАЗЫ ДАННЫХ

Предлагаемая база данных, как и всякая другая, имеет логическую и физическую организацию. Логическая организация представляет собой набор структур данных, которые может построить пользователь. В ее основу положена иерархическая модель баз данных^{/5,6/}. Она позволяет строить иерархические /древовидные/ структуры из наборов данных /сегментов/. Каждый сегмент этой структуры может иметь один исходный и несколько порожденных сегментов. Например, на рис.1 показана схема базы данных "цикл измерения". В прямоугольниках указаны имена сегментов. Сегмент "цикл измерений" имеет несколько порожденных сегментов, стоящих ниже по иерархии: один - "параметры цикла" и несколько подобных сегментов "эксперимент" /им присвоены порядковые номера/. Всякий "эксперимент" имеет один порожденный сегмент "параметры эксперимента" и несколько порожденных сегментов "измерение". Физический смысл этой иерархии состоит в отношении подчинения. "Цикл измерений" содержит несколько "экспериментов", каждый из которых в свою очередь содержит несколько "измерений".

Физическая организация базы данных представляет собой совокупность способов размещения данных в памяти ЭВМ. Описываемая база данных имеет форму одного или нескольких файлов в рамках базовой операционной системы. Связь файлов закладывается путем записи их имен в поля данных.

Каждый из файлов в первых четырех байтах содержит адрес конца информации. Физической единицей обмена в базе данных является слово /2 байта/, в отличие от файловой системы машин типа СМ-3, СМ-4 и PDP-11, где единицей обмена является блок /512 байт/. Следом за адресом конца информации последовательно располагаются сегменты. Их представление в памяти ЭВМ делится на две части. Первая выполняет служебные функции и содержит имя и номер

сегмента, ссылки на исходный, порожденный и подобный сегменты. Вторая часть содержит имена и значения полей данных.

Каждая ссылка является адресом соответствующего сегмента. На рис.2 и рис.3 показаны физическая организация данных внутри файла и структура отдельного сегмента. Отметим, что сегменты, порожденные от

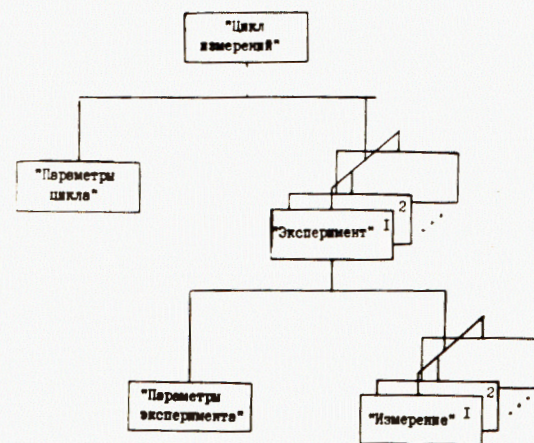


Рис.1. Схема иерархической базы данных "цикл измерений".

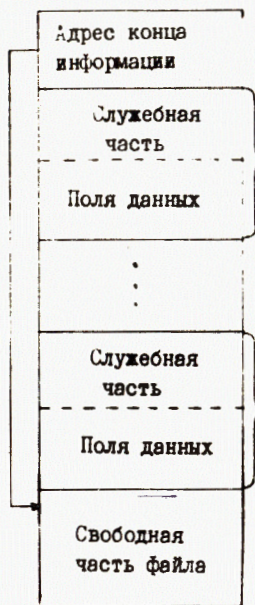


Рис.2. Физическая организация файла базы данных.

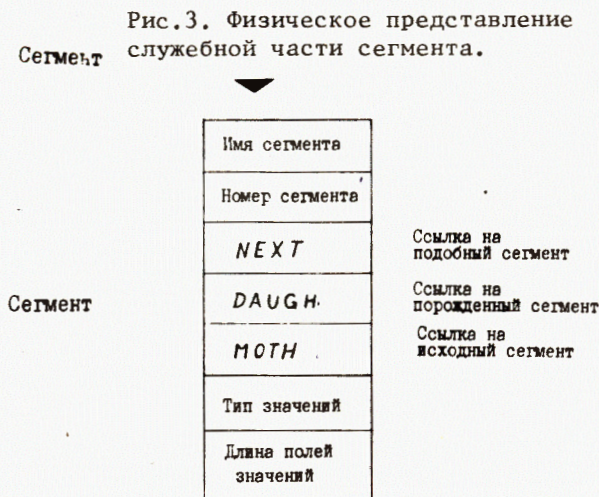


Рис.3. Физическое представление служебной части сегмента.

хических взаимосвязей сегментов в базе данных, а также роль всех трех ссылок.

В структуре предлагаемой базы данных нет специально выделенных полей для поиска нужных наборов информации - ключевых полей. Создано несколько процедур поиска по разным полям. На их основе можно построить любые другие процедуры, которые обеспечат необходимые операции поиска.

Описанная организация базы данных достаточно проста и позволяет решить поставленные задачи. Особенности физической организации базы данных "спрятаны" внутри пакета процедур, реализующих работу с базой данных. Знать их полезно для оценки эффективности применения такой базы данных и созданных программных средств в различных задачах. Для использования базы данных достаточно знать ее логическую организацию и набор предлагаемых процедур.

3. ОПИСАНИЕ ПАКЕТА ПРОЦЕДУР

Пакет процедур, обеспечивающих использование базы данных описанной структуры назван HFS /Hierarchical File Service/. Этот пакет легко может быть дополнен новыми программами, созданными на основе имеющихся.

Следует отметить особенность реализации пакета HFS. Процедуры HFS используют файлы прямого доступа и соответствующие процедуры языка Паскаль^{/2/} (REWRITE, RESET, SEEK, DEPOSIT, CLOSE RANDOMFILE).

HFS включает 20 процедур общим объемом около 600 строк на языке Паскаль. Их можно разделить на 4 группы. Первая включает основной набор пакета, являющийся основой построения остальных процедур. Вторая группа содержит рабочие процедуры чтения и записи сегментов, третья - процедуры поиска, а четвертая - некоторые важные сервисные процедуры. Ниже приводится список процедур и короткое описание их назначения.

Первая часть пакета HFS

HFSRD - чтение массива информации в режиме последовательного доступа /длина массива указывается в списке параметров/.

HFSWR - запись массива указанной длины в режиме последовательного доступа.

HFSBREAK - запись остатков информации из служебного буфера пакета HFS /операция выполняется после использования процедуры HFSWR перед закрытием файла/.

HFSUPD - модификация указанной части файла.

HFSADD - добавление очередной записи /служебной части сегмента или полей значений/ в файл базы данных.

HFSFADR - определение адреса начала свободной части файла.

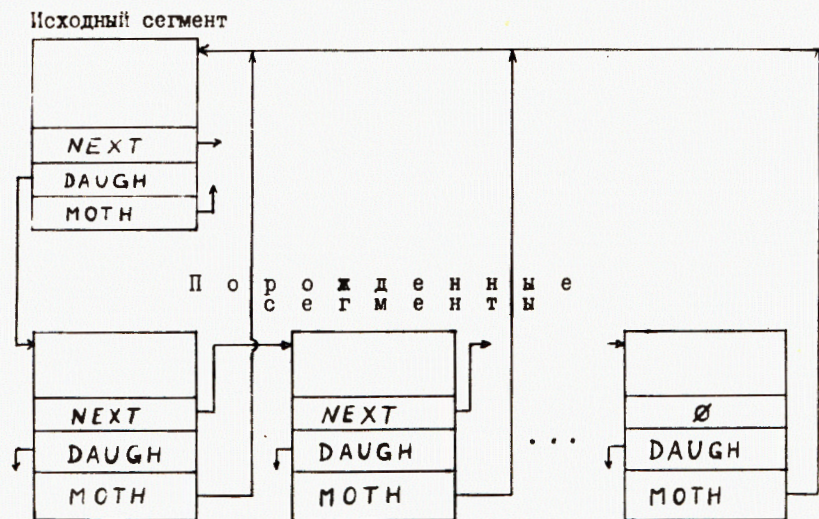


Рис.4. Физическая организация связей сегментов в базе данных.

одного исходного сегмента, составляют линейный список, организованный с помощью ссылок NEXT. На рис.4 показана схема иерар-

Вторая часть пакета HFS

HFSGETS - чтение служебной части сегмента /адрес сегмента задан в списке параметров/.

HFSPUTS - запись служебной части сегмента по указанному адресу /запись выполняется в режиме модификации файла, используя процедуру HFSUPD/.

HFSVGET - чтение полей данных сегмента в единый массив /в списке параметров указывается адрес сегмента/.

HFSVPUT - запись полей данных сегмента.

HFSAUPD - установка ссылок на подобный и порожденный сегменты.

HFSADDS - запись служебной части порожденного сегмента.

HFSNADDS - запись служебной части порожденного сегмента и присвоение ему очередного номера среди подобных.

Третья часть пакета HFS

Все процедуры поиска в результате своей работы определяют адрес искомого сегмента. Если нужный сегмент не обнаружен, то записывается нулевое значение адреса.

HFSDDL - поиск последнего в линейном списке порожденного сегмента.

HFSDDL - поиск последнего экземпляра порожденного сегмента с указанным именем.

HFSDDNU - поиск порожденного сегмента с указанным именем и номером экземпляра.

Четвертая часть пакета HFS

HFSKILL - уничтожение порожденного сегмента с указанным именем и номером экземпляра. Процедура не выполняет физического уничтожения сегмента; ссылки изменяются таким образом, что указанный сегмент убирается из схемы базы данных. Связи всех остальных сегментов сохраняются.

HFSSCHEME - вывод на печать или алфавитно-цифровой дисплей схемы любой базы данных, созданной с помощью пакета HFS.

NEXTSEGM - поиск очередного сегмента и чтение его служебной части. Процедура реализует универсальный алгоритм, позволяющий по очереди просмотреть все сегменты базы данных.

GARBADGE - "сборка мусора" в файле данных. Процедура путем физической перезаписи выбрасывает все образы уничтоженных с помощью HFSKILL сегментов.

ЗАКЛЮЧЕНИЕ

Методика использования базы данных может служить основой для создания комплексов, объединяющих используемые физиками программы - от программ накопления информации и управления

экспериментом до программ окончательной математической обработки. Она также обеспечивает независимость прикладных программ от организации данных, позволяет совершенствовать одни программы без изменения других, используя как интерфейс базу данных и единый набор подпрограмм для работы с ней. Создание программного обеспечения эксперимента в виде единого комплекса облегчает работу физика с написанными для его эксперимента программами, уменьшает количество ошибок в процессе диалога "экспериментатор - ЭВМ", позволяет освободить физика от необходимости детального знания организации данных, полученных во время накопления или передаваемых от одной фазы обработки к другой.

К отличительным особенностям разработанной базы данных относятся следующие:

1. Принятая организация базы данных проста и ориентирована на применение ее в системах автоматизации, построенных на базе мини- и микроЭВМ.

2. Создание и использование базы данных не требует специальной программы-администратора.

3. Ключевые слова для организации поиска не отмечены в структуре самой базы данных, а определяются выбором необходимой процедуры поиска.

4. Средства управления базой данных реализованы в виде пакета процедур. Они просты, невелики и могут использоваться в программном обеспечении систем автоматизации экспериментов на мини- и микроЭВМ.

Описанный набор процедур ориентирован на использование его разработчиками программного обеспечения систем автоматизации экспериментов. Достаточно низкий уровень операций, реализуемых процедурами, позволяет разработчику рационально компоновать прикладную программу, чтобы она удовлетворяла требованиям к скорости ее работы и объему занимаемой памяти.

Предложенную методику можно использовать и в других областях автоматизации человеческой деятельности. Очевидно, что реализация методики может быть осуществлена на различных языках программирования и в различных операционных системах.

Автор считает приятной обязанностью выразить благодарность Г.П.Жукову за помощь и поддержку данной работы, А.М.Балагурову и Г.М.Мироновой за помощь в применении описанных методики и программных средств в системе автоматизации экспериментов на нейтронном дифрактометре.

ЛИТЕРАТУРА

1. Gates W. The Future of Software Design Byte, 1983, vol.8, No.8, p.401-403.
2. Грогано П. Программирование на языке Паскаль. "Мир", М., 1982.

3. Малые ЭВМ и их применение. Под ред. Б.Н. Наумова. "Статистика", М., 1980.
4. Frank I.M. Pacher P. Physica, 1983, 120B, p.37-44.
5. Мартин Дж. Организация баз данных в вычислительных системах. "Мир", М., 1978.
6. Капп Д., Лебен Дж. Техника программирования для IMS. Методология использования DL/1. "Финансы и статистика", М., 1983.

Рукопись поступила в издательский отдел
25 июня 1984 года

Островной А.И.

P10-84-439

Методика применения базы данных
в системе автоматизации экспериментов и ее реализация

Предложена методика применения базы данных в системах автоматизации экспериментов, созданных на основе мини- и микроЭВМ. Эта методика обеспечивает унифицированное представление экспериментальной информации и единообразный доступ к ней со стороны большого числа программ, работающих как в on-line, так и в off-line режиме. Способ реализации предложенной методики в виде пакета подпрограмм позволяет использовать ее в системах автоматизации на мини- и микроЭВМ. В работе описан пакет подпрограмм на языке Паскаль для ЭВМ типа СМ-3, СМ-4.

Работа выполнена в Лаборатории нейтронной физики ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1984

Перевод О.С. Виноградовой

Ostrovnoj A.I.

P10-84-439

Methods of Data Base Application
in Experiment Automation Systems and Its Realization

Methods of data base application in experiment automation systems created on the base of mini- and microcomputers are suggested. These methods support a unified format of experimental data and unified access to it for any on-line and off-line routines. Realization of this method performed as a set of procedures allows one to use them in automation system based on mini- and microcomputers. Procedures written in Pascal language worked out on the PDP-family computers are described.

The investigation has been performed at the Laboratory of Neutron Physics, JINR.

Communication of the Joint Institute for Nuclear Research. Dubna 1984