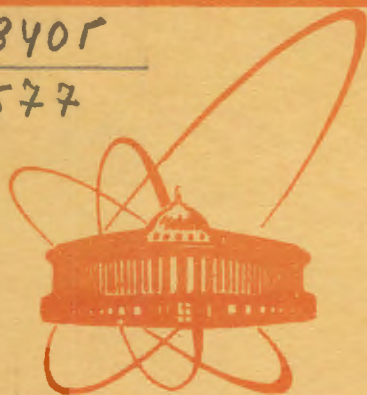


Ц840Г
Г-577



сообщения
объединенного
института
ядерных
исследований
дубна

3950 / 2-79

1/10-79

P10 - 12445

Н.Н.Говорун, С.М.Мальцева, В.И.Никитина,
Г.Н.Тентюкова

ЛОГИЧЕСКАЯ И ФИЗИЧЕСКАЯ ОРГАНИЗАЦИЯ
БАЗЫ ДАННЫХ ИСК

1979

P10 - 12445

Н.Н.Говорун, С.М.Мальцева, В.И.Никитина,
Г.Н.Тентюкова

ЛОГИЧЕСКАЯ И ФИЗИЧЕСКАЯ ОРГАНИЗАЦИЯ
БАЗЫ ДАННЫХ ИСК

Объединенный институт
ядерных исследований
БИБЛИОТЕКА

Говорун Н.Н. и др.

P10 - 12445

Логическая и физическая организация базы данных ИСК

Рассмотрены принципы организации базы данных системы ИСК с точки зрения прикладного программиста и пользователя системы /логическая организация/ и способ организации данных в среде хранения /физическая организация/.

Система управления базой данных ИСК предназначена для работы с документами картотечного типа и реализована на ЭВМ БЭСМ-6. В системе имеется два типа структур: древовидная и сетевая. Обсуждаются структура документа, являющегося основной единицей базы данных, а также соотношения между записями разных логических файлов. В системе применен способ организации файла документов методом формирования частично инвертированных списков. Система ИСК ориентирована на применение в информационно-поисковых системах, реализована в задаче кадрового учета АСУ ОИЯИ и может быть использована в других задачах АСУ, имеющих документы картотечного типа.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1979

Govorun N.N. et al.

P10 - 12445

The Logical and Physical Structure of Data Base ISK

Concepts of data base organization in ISK system from the application programmer point of view (logical structure) and data arrangement in storage medium (physical structure) are considered.

Data base management system ISK is intended for card type documents handling and realized with the BESM-6 computer. Two types of data structures are defined: hierarchical and networking. The structure of document, as a main unit of database, and relationships between records of different logical files are discussed. The file organization implemented in system is partly inverted file. The system ISK is oriented to be used in information retrieval systems. Personnel system in JINR is realized on it. Also it may be used in other management problems with card type documents.

The investigation has been performed at the Laboratory of Computing Techniques and Automation, JINR.

Communication of the Joint Institute for Nuclear Research. Dubna 1979

© 1979 Объединенный институт ядерных исследований Дубна

В настоящее время уделяется много внимания вопросам конструирования автоматизированных баз данных различного назначения^{/1,2/}. Системы управления базами данных входят в качестве основной компоненты в конкретные системы обработки данных, в том числе и в АСУ.

Система управления базой данных ИСК^{*}) предназначена для работы с документами картотечного типа и реализована на ЭВМ БЭСМ-6. Она создана на основе информационной системы "Кадры"^{/3,4/} как некоторое ее расширение.

I. Логическая организация базы данных ИСК

Основной структурной единицей базы является документ, содержащий набор сведений об объекте системы и состоящий из набора информационных полей. Совокупность документов для однородных объектов образует логический файл. Совокупность логических файлов образует базу данных.

В системе определены два типа связей между данными:
- первый тип определен внутри документа и его составных частей,
- второй тип - связь между логическими файлами.

Соответственно в системе реализованы два типа структур: древовидная и сетевая.

Документ образует неоднородную древовидную структуру переменной глубины. В рассматриваемом варианте допускается максимально два уровня. Допускаются простые (один к одному) и сложные (один ко многим) связи между элементами внутри документа.

В документе могут быть определены группы или повторяющиеся группы, состоящие из элементов фиксированной или переменной длины. Каждое подразделение данных (логический файл, документ, груп-

^{*} ИСК - информационно-поисковая система для документов картотечного типа.

па, элемент) имеет уникальное имя. Для элементов данных должны быть определены:

- 1) тип представления (строка символов, дата или число),
- 2) длина поля,
- 3) признак блокировки доступа,
- 4) признак активности,
- 5) признак возможности быть не задействованным.

Для элементов, имеющих числовое представление, дополнительно задается список возможных значений или диапазон изменения.

Характеристики объектов, для которых составлены документы, будем называть в дальнейшем признаками.

Отношения между записями разных логических файлов образуют неоднородную сетевую структуру. Каждый документ одного логического файла может быть связан с любым документом любого другого логического файла. На рис. 2 приведен пример связи между документами с номерами от 1 до 9 логических файлов F1, F2, F3. Рис. 3 изображает пример связи между двумя логическими файлами, имеющими имена "Заявки" и "Авторы", которые введены в систему учета изобретений ОИЯИ.

Связь между записями двух разных логических файлов устанавливается на основании описания данных одного файла.

2. Физическая организация базы данных ИСК

Физическая организация — это организация данных на запоминающих устройствах, размещение данных в среде хранения.

При выборе способа физической организации определяющими факторами являются обеспечение минимальных затрат памяти и малого времени на ответ.

В системе ИСК все документы заносятся в один физический файл, называемый файлом документов, который может быть размещен на магнитном диске или магнитной ленте. Доступ к записям этого файла может быть последовательным или произвольным в зависимости от характера работы. Каждый документ является отдельной записью. Адреса и имена документов образуют поименованные таблицы адресов и имен, используемые для поиска документов, относящихся к заданному логическому файлу.

В системе применен способ организации файла документов методом формирования частично инвертированных списков. Это означает, что для некоторых значений признаков, являющихся характе-

ристиками объектов, составляются списки номеров документов, содержащих эти значения. Списки могут быть составлены также для интервалов значений признаков. В качестве значений признаков для составления инвертированных списков выбираются наиболее часто используемые. Для хранения всех инвертированных списков (линеек) использован один физический файл- файл линеек. Адреса списков в этом файле содержатся в таблицах справочников ключей, в которые также включены другие сведения о значениях признаков. Строка таблицы содержит, в частности, мнемоническое имя значения признака, его числовое представление (или интервал числовых представлений), длину списка. Каждый инвертированный список является отдельной записью этого файла.

Связи между логическими файлами задаются в виде упорядоченных поименованных списков указателей переменной длины.

Таблицы связей, а также таблицы адресов и имен документов, справочники ключей хранятся в одном физическом файле, называемом файлом таблиц. Большинство таблиц имеет значительные объемы, и они наращиваются по мере расширения фонда документов. Таблицы заносятся во внешнюю память на магнитный диск страницами в очередную свободную зону диска, определяемую по каталогу, так что физически каждая такая таблица может оказаться разделенной на участки, занимающие несколько зон. Записью файла таблиц является участок таблицы или вся таблица длиной, не превышающей 1024 слова. Записи могут быть вызваны в основную память по ключам записи или последовательно по именам таблиц.

Распределение памяти на диске для указанных трех физических файлов (документов, линеек и таблиц) выполняется автоматически в соответствии с информацией, задаваемой системе.

Программы обмена с внешней памятью описаны в работе /5/.

- Для экономии внешней памяти применены различные способы:
- блокирование памяти, т.е. объединение нескольких логических записей в одну физическую,
 - максимально возможное сжатие данных,
 - замена восьмибитного представления символов на шестибитное для хранения текстов,
 - замена некоторых текстов на числовые коды.

Система опробована на группе документов патентного отдела ОИЯИ. Программное обеспечение системы широко использует возможности операционной системы "Дубна" /6/.

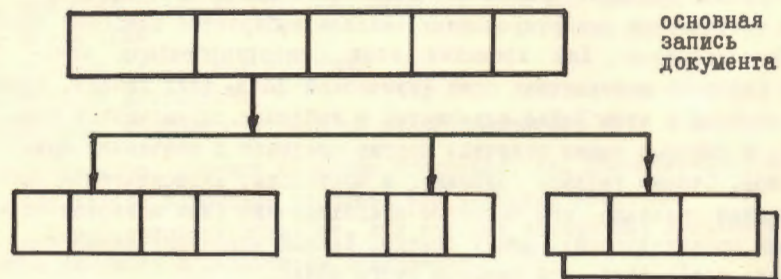
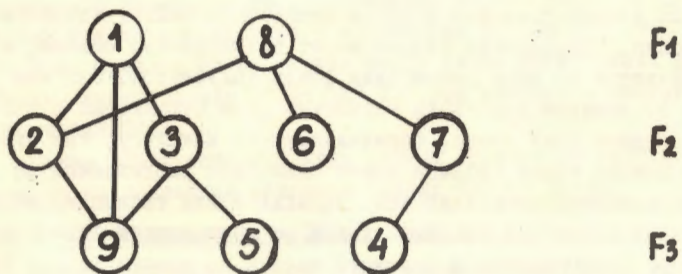


Рис. 1



F1

F2

F3

Рис. 2

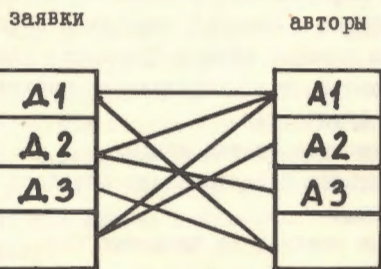


Рис. 3

Литература

1. Дж. Мартин. Организация баз данных в вычислительных системах. М., "Мир", 1978.
2. Информационные системы общего назначения (аналитический обзор систем управления базами данных). М., "Статистика", 1975.
3. Говорун Н.Н. и др. Программное обеспечение системы "КАДРЫ". ОИЯИ, IO-IO95I, Дубна, 1977.
4. Говорун Н.Н. и др. Организация информационных массивов системы "КАДРЫ". ОИЯИ, IO-IIIO5I, Дубна, 1977.
5. Мазный Г.Л. Вариант программ для обмена с МЛ, МБ и МД строками и страницами на ФОРТРАНе. ОИЯИ, II-9845, Дубна, 1976.
6. Мазный Г.Л. Программирование на БЭСМ-6 в системе "Дубна". М., "Наука", 1978.

Рукопись поступила в издательский отдел
7 мая 1979 года.