Материалы к практическим занятиям по Программной инженерии

1. Теоретический материал

**1. Структура единиц хранения в базах данных с фиксированным форматом**

Данные, содержащиеся в экономических информационных системах, имеют в большинстве случаев фиксированный формат.

**Свойства данных в фиксированном формате:**

1. Обычная структура подобных данных представляет собой таблицу с фиксированным количеством столбцов и переменным количеством строк. При этом строки обычно соответствуют отображаемым реальным событиям, фактам или объектам, а величины или слова, размещенные в столбцах таблицы, соответствуют свойствам, характеристикам этих событий, фактов или объектов или же идентифицируют их. Поэтому каждая строка характеризуется одним и тем же набором свойств, но различными значениями этих свойств.
2. На пересечении строки и столбца всегда находится единственное значение соответствующего свойства, характеризующее данный объект.
3. Порядок столбцов таблицы, как правило, фиксирован, но это не существенно. Каждый столбец таблицы имеет уникальное название, или имя. При этом элемент данных может иметь несколько имен, ни одно из которых не может совпадать с именем какого-либо другого элемента.
4. Любые две строки таблицы различаются, по крайней мере, одним элементом.
5. Порядок строк таблицы не существен и может меняться в процессе обработки данных.

Строки определенной подобным образом таблицы данных принято называть записями. Каждый элемент таблицы рассматривается при обработке данных как неразложимый на более мелкие элементы и называется **элементом данных,** или **реквизитом,** или **атрибутом.**

***Определение:*** для таблицы данных, обладающей перечисленными свойствами, приняты такие названия, как **нормализованный файл**, **реляционный файл**, **Р-файл**, **реляционная база данных** (в последнем случае речь чаще идет о совокупности связанных таблиц). При этом перечень имен реквизитов файла (с возможной дополнительной информацией о реквизитах – типы, длины и проч.) называется **схемой** файла, тогда как совокупность значений реквизитов, находящихся в строках (записях) файла, называется **экземпляром** файла или **экземпляром схемы**.

Таким образом, свойство 3 нормализованного файла можно понимать в том смысле, что каждый реквизит повторяется в схеме файла не более одного раза.

**Связь реляционного файла с документом**

**Реляционный файл представляет собой машинный образ некоторого внемашинного документа с жестким форматом, или же некоторой совокупности подобных документов.** Каждая строка реляционного файла может соответствовать одному документу или одной строке некоторого документа. В таком, характерном для бухгалтерских систем, файле, как журнал операций, каждая запись соответствует одной бухгалтерской проводке. Проводка – это строка данных в жестком формате с реквизитами: дата, номер проводки, счет по дебету, счет по кредиту, сумма, описание. При использовании реляционной структуры устанавливается следующая **иерархия единиц данных**:

**1 ступень**. **Элементарной единицей данных** в жестком формате считается **реквизит**. Реквизит характеризуется уникальным именем, соответствующем имени соответствующего столбца в реляционном файле. Реквизит может иметь более одного имени. Каждый реквизит имеет область допустимых значений (тип); в каждом данном файле каждый реквизит может принимать некоторое подмножество значений из своей области значений. Реквизит в системах ручной обработки данных не определяется столь строго и по существу не отличается от традиционного понятия реквизита обычного документа.

При машинной обработке данных имя реквизита (представляющее собой аналог имени переменной в алгоритмических языках) есть тот идентификатор, под которым данный реквизит известен обрабатывающим программам. В каждый данный момент существования записи реквизит принимает некоторое значение (аналог значения переменной) из допустимого множества значений, определенного типом данного реквизита. Реквизиты с числовой областью значений принято относить к категории оснований. Реквизитом-основанием является любой реквизит, значениями которого являются денежные суммы.

Реквизиты символьного типа – разного рода коды и наименования – часто называют призначными, или реквизитами-признаками. Таким образом, реквизит рассматривается как наименьшая, неразложимая единица данных.

**2 ступень**. Более крупные (составные) единицы данных обычно строятся путем композиции (конкатенации) реквизитов. Совокупность реквизитов, описывающая некоторый объект реального мира, называется **составной единицей данных** (СЕД). Составные единицы данных представляют собой, как правило, композиции реквизитов, имеющие некоторое экономическое содержание**. СЕД, имеющая законченный экономический смысл, обладает свойством документообразования**. Так, в автоматизированном бухгалтерском учете составными единицами данных могут считаться проводка, строка баланса, актив (или пассив) баланса как целое, баланс в целом. Отметим, что файл обычно не относят к составным единицам данных; его считают объектом более высокого уровня, тогда как запись в жестком формате – разновидность составной единицы данных. Под длиной записи фиксированного формата понимают сумму длин реквизитов, входящих в запись.

ПОКАЗАТЕЛЬ – это составная единица данных, содержащая количественную характеристику некоторого экономического объекта или процесса (основание) и признаки, позволяющие идентифицировать этот процесс или объект.

Бумажный экономический документ во многих случаях состоит из показателей. Так, каждая строка бухгалтерского баланса представляет собой показатель, построенный по определенным правилам из итоговых показателей (остатков) счетов бухгалтерского учета. Показатели могут быть исходными и производными, причем правила построения производного показателя должны быть строго фиксированы.

Область значений составной единицы данных представляет собой декартово произведение областей значений основных реквизитов.

**3 ступень**. Множество записей представляет собой реляционный, или нормализованный файл (см. выше). Под длиной реляционного файла понимается количество записей (строк) в этом файле.

***Пример.*** Каждая запись в экзаменационной ведомости содержит такой реквизит, как номер зачетной книжки. Этот цифровой реквизит представляет собой реквизит-признак. Отметим, что такой общий для всех строк ведомости реквизит, как номер группы, можно мысленно представить себе, как присутствующий в каждой из строк и вынесенный в заголовок лишь из соображений технического удобства. Таким образом, экзаменационная ведомость не есть нормализованный файл. Можно представить себе нормализованный файл, построенный из экзаменационной ведомости путем внесения реквизитов заголовка в каждую из строк. Подобную операцию называют нормализацией. Практически существующие документы в жестком формате обычно не нормализованы, однако процедура нормализации при проектировании машинного файла достаточно проста, хотя и несколько громоздка. Процедура нормализации для одного частного случая приводится ниже.

Из свойства 4 нормализованного файла (см. выше) следует, что во всяком нормализованном файле существует такая СЕД (в частном случае реквизит), значения которой не могут совпадать ни в какой паре записей файла. Значения СЕД (реквизитов), обладающих подобным свойством, естественно использовать в качестве идентификаторов записей. Подобную СЕД (реквизит) принято называть **реляционным ключом** файла.

**Предупреждение:** термин "ключ" в информатике чрезвычайно полисемичен[[1]](#footnote-2), не следует смешивать введенное только что понятие реляционного ключа с иными употреблениями этого термина, например, с введенным выше понятием ключевого слова.

**4 ступень**. Совокупность реляционных файлов, обычно относящихся к одной и той же предметной области, или к семантически близким предметным областям, образует ***реляционную базу данных (РБД)*.**

**2. Базы данных в свободном и смешанном формате**

Базы данных в свободном формате предназначены, главным образом, для визуальной диалоговой работы пользователя за терминалом. (В частности, система "Консультант +" и другие подобные системы).

ЗАПИСЬ в свободном формате представляет собой машинный образ абзаца или строки документа свободного формата и обычно имеет переменную или неопределенную длину (количество символов).

При использовании и жесткого, и свободного форматов из записей, как элементов, состоит ФАЙЛ. Файл обычно содержит однородную экономическую или управленческую информацию. Разделение файла на записи обусловлено структурой алгоритма обработки данных. Следует отметить ту особенность, что для файла свободного формата порядок записей в нем обычно существен, тогда как записи жесткого формата могут располагаться в файле в любом порядке, в зависимости от потребностей процедуры обработки. Поэтому файл записей жесткого формата можно рассматривать как множество записей, тогда как подобное утверждение для документов свободного формата обычно неверно.

БАЗА ДАННЫХ в свободном формате – совокупность тематически связанных файлов в свободном формате. Отметим, что обычно базы данных содержат файлы как в том, так и в другом формате. Таким образом, в конечном счете внутримашинный информационный фонд представляет собой совокупность разнородных по формату баз данных.

Следует различать понятия файла для операционной системы компьютера и для системы обработки данных. С точки зрения операционной системы файл – это именованная область, занятая некоторой совокупностью данных, обрабатываемая операционной системой (или операционной средой) как одно целое, внутренняя структура которого операционной системе недоступна. Как внутренняя структура файла, так и объединение файлов в тематические базы определяется содержанием и назначением данных и принадлежат области ответственности системы обработки данных. Обычно тип внутренней структуры файла и его назначение отображаются той частью имени файла, которую принято называть расширением. Оно служит в основном для информирования пользователя и обычно не влияет на работу операционной системы и приложений с этим файлом.

Положительным качеством баз данных в свободном формате является возможность оперативного визуального обновления данных, предназначенных для некомпьютеризованного использования. Подобного рода информация используется юристами, а также работниками налоговых служб, финансистами, бухгалтерами и др. В таких базах данных информация – это различные инструкции, законодательство, поправки к нему и проч.

**3. Реквизитный анализ документа.**

РЕКВИЗИТНЫЙ АНАЛИЗ представляет собой процесс выявления в ручном (бумажном) документе отдельных реквизитов и тех составных единиц данных, в которые эти реквизиты входят. Всякий **внутримашинный** образ документа для компьютерной обработки представляет собой результат реквизитного анализа, выполненного в процессе проектирования. При реквизитном анализе выявляются:

* реквизиты, которые подлежат вводу в машинную память,
* элементы документа, которые не относятся ни к какому реквизиту и ни к какой СЕД,
* те данные, которые не подлежат хранению в информационном фонде.

Для документов, которые предполагается хранить в свободном формате реквизитный анализ для перевода их в машинную форму не так принципиален, так как в большинстве случаев возможные автоматические операции над ними сводятся к вводу документа в систему обработки данных полностью с помощью текстового редактора или сканера и к его выводу на подходящий внешний носитель. И далее все операции обработки подобных документов выполняются в режиме взаимодействия человек-ЭВМ (в *диалоге*) пользователем за терминалом системы.

Рассмотрим хорошо известный в студенческой практике документ – экзаменационную ведомость. Первая страница документа изображена на рис. 1. На рис. 2 воспроизводится часть второй страницы ведомости. В методических целях документ разделен на части, но такое деление в реальном документе отсутствует. Метки жирным курсивом, также фактически отсутствующие в реальном документе, предназначены для анализа структуры данного документа. Значения реквизитов, которыми заполняется бланк документа, выделены нежирным курсивом.

Для деления документа на части отметим сначала, что в составе большинства внемашинных документов в фиксированном формате принято выделять:

1. **Общую часть документа** (иногда называемую заголовочной, хотя по месту расположения реквизиты общей части могут располагаться и в так называемом подножии (обычно итоговые реквизиты).
2. **Предметные строки документа** – множество однородных строк данных, в большинстве случаев содержащее неопределенное количество подобных строк.
3. **Оформительскую**, или **заверительную часть** документа – подписи, печати и штампы, удостоверяющие юридическую силу документа и ответственность за его оформление и выполнение; оформительские реквизиты не предназначены, как правило, для хранения в виде нормализованного файла; отметим, что существуют современные технологии электронной подписи и другие методы электронной авторизации юридической силы документа, введенного в память ЭВМ, которые относятся к технологиям защиты информации в ЭВМ.
4. **Реквизиты, предназначенные для улучшения** читабельности внемашинного документа и не подлежащие вводу в нормализованный файл, но, возможно, хранящиеся в форме машинной константы того или иного вида или же в позиции машинного файла специального вида, например, в поле типа MEMO.
5. **Текст в свободном формате**, который может храниться в виде машинной константы, отдельного файла или же в поле типа MEMO.

В дальнейшем будем обозначать с помощью введенной нумерации (1, 2, 3, 4, 5) типы отдельных частей документа.

Проанализируем теперь формат отдельных частей документа "Экзаменационная ведомость". В составе этого документа мы видим **СЕД** заголовочной части ***2р*** (содержащую реквизиты общего характера, тип 1, относящиеся ко всему документу, "№ экзаменационной ведомости", "№ группы", "факультет", "отделение", "специальность", "дисциплина", "семестр", "экзаменатор", "день", "месяц", "год") и ***4р*** (содержащую реквизиты схожего типа "количество отличных оценок", "количество хороших оценок", "количество удовлетворительных оценок", "количество неудовлетворительных оценок", "количество неаттестованных студентов", а также СЕД "ведомость проверена" с реквизитами "день", "месяц", "год"). Предметные строки (тип 2) в количестве, равном количеству студентов в группе, содержат для каждого из них реквизиты ***8р*** – порядковый номер, ***9р*** – фамилия и инициалы студента, ***10р*** – сдача зачета, ***12р*** – экзаменационная оценка, ***13р*** – номер зачетки. Оформительские (заверительные) реквизиты (тип 3) имеются как в общей части – подписи декана ***5з*** (в заголовке) и ***15з*** (в подножии), так и в предметной части – ***11з*** и ***14з*** – подпись преподавателя. Константы типа 4, обозначенные через ***1н*** и ***3н***, находятся в заголовочной части документа. Наконец, часть первой страницы документа является текстом в свободном формате ***6св***.

Необходимо отметить, что текст ***6св*** содержит в своем составе описание области значений (перечень допустимых значений) для реквизита "сдача зачета": {***c1*** – "зачтено", ***c0*** –"не зачтено", ***m0*** – "не явился"}, а также область допустимых значений реквизита "экзаменационная оценка": {***m5*** – "отлично", ***m4*** – "хорошо", ***m3*** – "удовлетворительно", ***m2*** – "неудов­летворительно" ***m0*** – "не явился"}. Напомним еще раз, что жирные курсивные обозначения в документе фактически не используются и внесены в рисунок в методических целях. Во всех случаях использования условных обозначений ниже в тексте данной книги следует помнить, что фактически используются реальные обозначения, а не заменяющие их условные.

|  |  |
| --- | --- |
| ***1н*** НИЖЕГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  им. Н.И. ЛОБАЧЕВСКОГО | |
| ***2р*** ЭКЗАМЕНАЦИОННАЯ ВЕДОМОСТЬ № \_ *14*\_  Группа № \_*719*\_  Факультет *экономический* *(7)* \_\_\_\_\_\_ Отделение *дневное* \_\_\_\_  Специальность \_\_*0611* \_\_\_\_\_\_ Дисциплина \_ \_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_  \_\_\_ *Общая теория систем*\_\_\_\_\_ (*ОТС)*  Семестр \_\_*2*\_\_\_\_\_  Экзаменатор \_\_*Белов В.А.*\_\_\_\_\_\_\_\_\_\_\_\_\_ Дата "\_*18*\_\_" *июня* \_\_ 20 \_*06*\_\_ г. | |
|  | ***3н*** ИТОГИ ЭКЗАМЕНА |
|  | ***4р***  Отлично \_*3*\_  Хорошо \_*4*\_  Удовлетворительно \_*7*\_  Неудовлетворительно \_*2*\_  Не аттестовано \_*3*\_  Ведомость проверена  "\_*19*\_"\_*июня*\_20*01\_*г. |
|  | ***5з***  Декан факультета \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  (подпись) |
| ***6св*** ВНИМАНИЮ ПРЕПОДАВАТЕЛЕЙ:  Перед началом экзамена студент представляет экзаменатору зачетную книжку (а при необходимости и документ, удостоверяющий личность), при этом номер зачетной книжки заносится в экзаменационную ведомость.  Успеваемость студентов оценивается следующими оценками: ***c1*** –"зачтено", ***c0*** –"не зачтено", ***m5*** – "отлично", ***m4*** – "хорошо", ***m3*** – "удовлетворительно", ***m2*** – "неудовлетворительно". Оценки "не зачтено" и "неудовлетворительно" заносятся только в экзаменационные ведомости.  Все записи в документы заносятся чернилами, а исправления должны быть точно оговорены за подписью лиц, вносящих исправления (например: исправленному на "хорошо" верить (подпись преподавателя).  После завершения приема экзамена (зачета) у студента оценка не­медленно проставляется экзаменатором в ведомость и зачетную книжку, минуя черновики, и заверяется подписью экзаменатора.  По окончании экзамена в группе экзаменатор в ведомости указывает ***m0*** – "не явился" против фамилий студентов, отсутствовавших на экзамене (без указания причин) и расписывается, а также подводит итоги экзамена.  По окончании экзамена (зачета) экзаменатор лично представляет ведомость в деканат факультета, но не позднее 10 часов утра следующего дня. | |

Рис. 1 Бланк экзаменационной ведомости - 1-я стр. разметкой для реквизитного анализа

***7сх***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| №№ | Фамилия и инициалы  студента | Сдача зачета | Подпись препод. | Экз. оценка | №  зачетки | Подпись препод. |
| ***8р*** | ***9р*** | ***10р*** | ***11з*** | ***12р*** | ***13р*** | ***14з*** |
| 1. | АБРАМЯН А. А | *зачтено* |  | *хорошо* | 951190 |  |
| 2. | АРХАНГЕЛЬСКАЯ Т.И. | *зачтено* |  | *отлично* | 951191 |  |
| 3. | БАРИНОВА А.Е. | *зачтено* |  | *удовлетв.* | 951192 |  |
|  | ------------------ | ------ | -------- | -------- | ---- | ---- |
|  | СОКОЛОВ Г.Е. | *не зачтено* |  | *не явился* | 941787 |  |
|  | СТРОИТЕЛЕВ Д.А. | *не явился* |  | *не явился* | 951208 |  |
|  | ШВЕЦОВА М.А. | *зачтено* |  | *удовлетв.* | 951209 |  |
|  | ЯСНОВ А.С. | *зачтено* |  | *отлично* | 941025 |  |

***15з*** ДЕКАH ФАКУЛЬТЕТА \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Рис. 2. Бланк экзаменационной ведомости – 2-я стр. с разметкой для реквизитного анализа

Особую роль играет часть документа, обозначенная через ***7сх***. Нетрудно видеть, что ***7сх*** – не что иное, **как *схема* предметной части документа**.

Проведем нормализацию файла "Экзаменационная ведомость" путем, как уже указывалось, **внесения реквизитов общей части в строки предметной части**. В нормализованном варианте используются присвоенные реквизитам машинные имена и условные значения соответствующих реквизитов.

Рассмотрим результаты реквизитного анализа и нормализации документа "Экзаменационная ведомость".

Результаты анализа сведены в табл. 1. (представлены на следующей странице).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| №№№ | Часть  документа | Тип части | Наименование рекв. во внемаш. документе | Перенос в норм. файл | Наим. в нормал. файле | Тип в нормализованном.  файле |
|  | ***1н*** | 4 | – | нет | – | – |
|  | ***2р*** | 1 | № экзаменационн. ведомости | да | P1 | признак |
|  | ***2р*** | 1 | № группы | да | P2 | признак |
|  | ***2р*** | 1 | факультет | да | P3 | признак |
|  | ***2р*** | 1 | отделение | да | P4 | признак |
|  | ***2р*** | 1 | специальность | да | P5 | признак |
|  | ***2р*** | 1 | дисциплина | да | P6 | признак |
|  | ***2р*** | 1 | семестр | да | P7 | признак |
|  | ***2р*** | 1 | экзаменатор | да | P8 | признак |
|  | ***2р*** | 1 | день | да | P9 | признак |
|  | ***2р*** | 1 | месяц | да | P10 | признак |
|  | ***2р*** | 1 | год | да | P11 | признак |
|  | ***3н*** | 4 | – | нет | – | – |
|  | ***4р*** | 1 | Количество  отличных оценок | – | Q12 | основание |
|  | ***4р*** | 1 | Количество  хороших оценок | – | Q13 | основание |
|  | ***4р*** | 1 | Количество удовлетворительн. оценок | – | Q14 | основание |
|  | ***4р*** | 1 | Количество  неудовл. оценок | – | Q15 | основание |
|  | ***4р*** | 1 | Количество неаттестованных студентов | – | Q16 | основание |
|  | ***4р*** | 1 | день | – | P17 | признак |
|  | ***4р*** | 1 | месяц | – | P18 | признак |
|  | ***4р*** | 1 | год | – | P19 | признак |
|  | ***5з*** | 3 | подпись декана | нет | – | – |
|  | ***11з*** | 3 | подпись  преподавателя | нет | – | – |
|  | ***14з*** | 3 | подпись  преподавателя | нет | – | – |
|  | ***15з*** | 3 | подпись декана | нет | – | – |
|  | ***6св*** | 5 | ВНИМАНИЮ ПРЕПОДАВАТЕЛЕЙ: | нет | – | – |

Таблица 1.Результат реквизитного анализа документа "Экзаменационная ведомость" – таблица реквизитов (начало)

**Дополнительное задание**

**Описать процесс использований найденного документа. Чтобы не придумывать этот процесс, найти описание этого процесса в сети интернет и изложить этот процесс в отчёте. В отчёте указать ссылку на этот источник.**

**Ознакомиться с системами описания бизнес-процессов типа BPML.**

**Контрольные вопросы по теме лабораторной работы**

1. Охарактеризуйте соотношение понятий информации и данных. Что мы ищем в Интернете – информацию или данные? Ответ обоснуйте.
2. Охарактеризуйте понятия «информация», «данные», «Экономическая информация», их основные свойства и область применения.
3. На какие виды разделяется экономическая информация по функциям управления?
4. Что включает учётная информация?
5. Перечислите особенности нормативно-справочной и отчётно-статистической информации.
6. Как классифицируются информационные системы?
7. Охарактеризуйте понятие «пользователь информационной системы», приведите классификацию пользователей.
8. Как классифицируются информационные системы по сфере применения?
9. Дайте характеристику и приведите примеры документов в свободном и фиксированном форматах.
10. Дайте определение понятия «экономический документ».
11. Чем отличаются между собой БД свободного и жёсткого форматов?
12. В чём заключается взаимосвязь между понятиями «атрибут» и «СЕИ»?
13. Что понимается под термином «экономическая информационная система». Опишите общие свойства систем на примере экономических информационных систем (ЭИС).
14. Какие принципы создания экономических информационных систем Вы знаете? Охарактеризуйте каждый из них.
15. Дайте примеры объектов предметной области экономических систем;
16. Какие факторы определяют различие между атрибутом-признаком и атрибутом-основанием в конкретном документе;
17. Как определяется значение СЕИ?
18. Назовите свойства и типы реквизитов?
19. Дайте определение понятий «имя реквизита» и «множество значений» (домен) реквизита;
20. Охарактеризуйте допустимые операции над показателями;
21. Перечислите аспекты, по которым производится классификация показателей;
22. Как организована иерархия экономических показателей?
23. Перечислите основные разновидности документов, применяемых в системах управления субъектами хозяйствования;
24. Приведите примеры типичных экономических документов и проведите анализ их структуры;
25. Приведите примеры типичных экономических документов и проведите их реквизитный анализ;
26. Перечислите и дайте трактовку возможным операциям над СЕИ экономических документов;
27. Понятие унифицированной системы документации.
28. Использование баз данных для отображения СЕИ.

1. Полисемия – свойство термина обозначать различные понятия в одной или различных предметных областях, напр., вал как деталь механического устройства и вал как валовая продукция или валовая выручка предприятия. Явление полисемии внешне сходно с омонимией, т. е. с графическим совпадением слов, несущих различный смысл, напр., простой (прилагательное) и простой – вынужденный перерыв в работе. [↑](#footnote-ref-2)