

Основные понятия баз данных

Базы данных и система управления баз данных

В истории развития вычислительной техники наблюдалось два основных направления ее применения.

Первое направление связано с выполнением больших численных расчетов, которые трудно или невозможно произвести вручную.

Второе направление – это использование вычислительной техники для создания, хранения и обработки больших массивов данных.

Такие задачи решают информационные системы (**ИС**).

К ним относятся: поисковые, справочные, автоматизированные системы управления, экспертные системы.

Одной из характерной разновидностью **ИС** является **банк данных**. В банке данных хранится информация об определённой предметной области в специальном представлении. При этом сами данные образуют **базу данных**, а банк, наряду с базой, содержит программные средства обработки данных.

База данных (БД) – это организованная совокупность данных, предназначенная для их хранения, накопления и обработки с помощью компьютера.

Свойства БД

Структурированность – БД должны иметь определенную структуру.

Взаимосвязанность – данные должны быть связаны между собой в базе.

Независимость от прикладных программ. Базы данных конструируются таким образом, чтобы с ними можно было работать в различных программных средах и на различных компьютерных платформах. Обычно файлы БД имеют расширение *.db (date baze).

Характеристики БД

Полнота – чем полнее БД, тем вероятнее, что она содержит нужную информацию (однако не должно быть избыточной информации);

Правильная организация – чем лучше структурирована БД, тем легче найти в ней информацию;

Актуальность – БД в каждый момент времени должна соответствовать состоянию отображаемого объекта;

Удобство в использовании – БД должна быть проста и удобна в использовании и иметь развитые методы доступа к любой части информации.

Программные приложения предназначены для обработки и управления *БД* удобной для пользователя форме, называется ***системой управления базой данных (СУБД)***.

Любая ***СУБД*** должна выполнять три основные функции:

- ввод данных;
- запросы по данным;
- составление отчетов по данным.

Все существующие современные СУБД удовлетворяют, как правило, следующим требованиям:

Возможности манипулирования данными (ввод, выбор, вставка, обновление, удаление и др.). Важными показателями при этом являются производительность **СУБД**, стоимость хранения и использования данных, простота обращения к базе данных и проч.

Возможность поиска и формирования запросов. С помощью запросов пользователь может оперативно получать различного рода информацию, которая хранится в базе данных.

Обеспечение целостности (согласованности) данных. При использовании данных многими пользователями важно обеспечить корректность операций, при которых не может быть нарушена согласованность данных.

Обеспечение защиты и секретности.
Кроме защиты от некорректных действий пользователей, важно обеспечить защиту данных от несанкционированного доступа и аппаратных сбоев. Проникновение в базу лиц, имеющих на это права, может иметь следствием разрушение данных. Секретность базы данных позволяет определять круг лиц, имеющих доступ к информации, и порядок доступа.

Среди **СУБД**, которые устанавливаются в небольших организациях и ориентированы на работу с конечными пользователями, наиболее популярны **Access**, **FoxPro**, **Paradox**. К более сложным системам относятся распределенные **СУБД**, которые предназначены для работы с большими базами данных, распределенными на нескольких серверах (серверы могут находиться в различных регионах). Мощными **СУБД** такого типа являются **Oracle**, **Sybase**, **Informix**.

Структура простейших баз данных

В основе любой базы данных лежит модель данных, иначе говоря, ***информационная структура***.

Термин «модель данных», впервые введенный в 1970-х годах, основоположником БД Дж. Коддом, в современной трактовке отображает совокупность правил порождения структур данных в базах данных, последовательности их изменения.

Модель данных является средством интерпретации содержимого базы данных и реализации операции по обработке и управлению данными.

Для организации базы данных используются три типа информационных структур (модели данных):

- иерархическую;
- сетевую;
- реляционную;

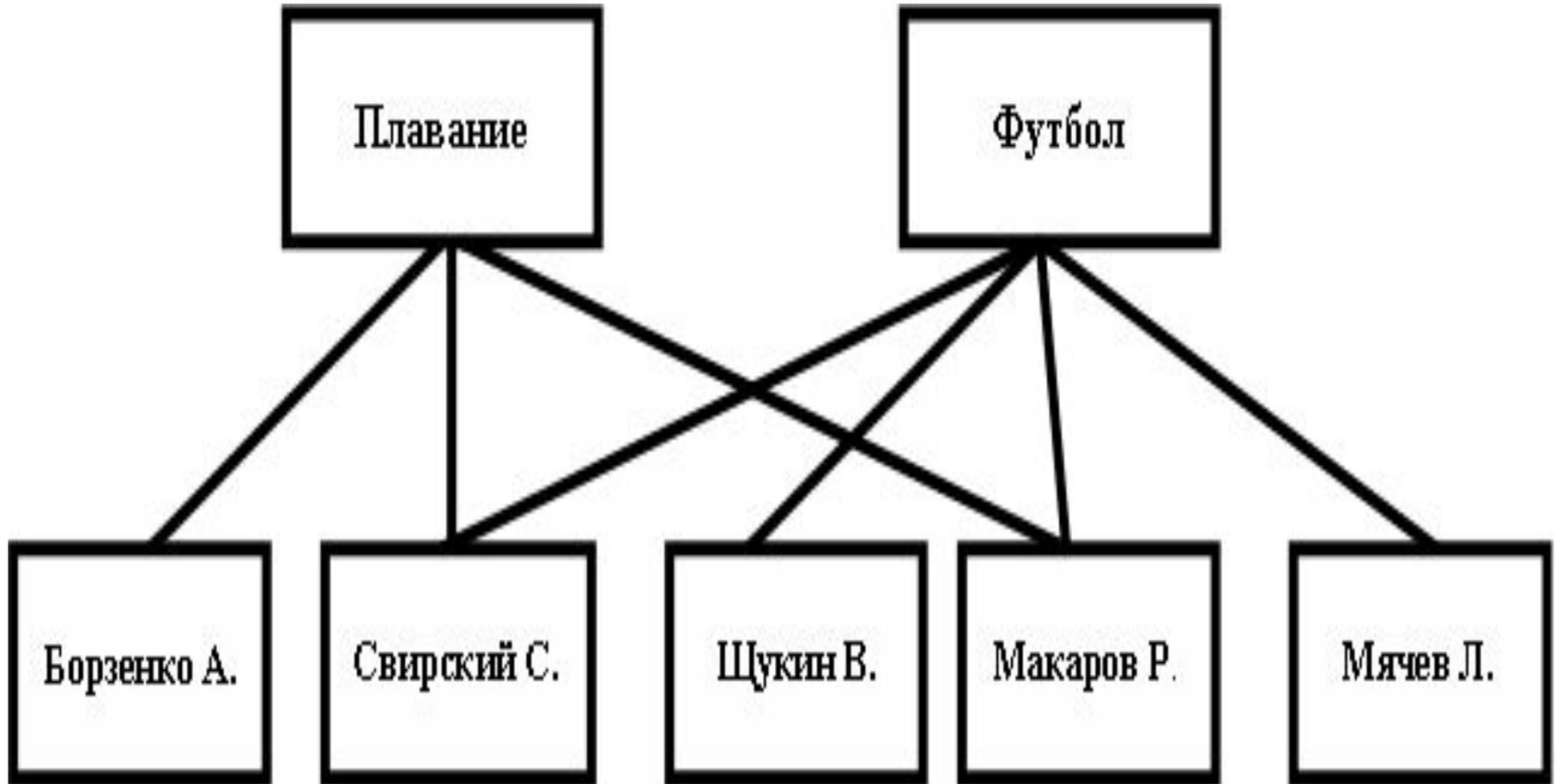
В *иерархической* БД существует упорядоченность элементов в записи. Для неё характерны подчиненность объектов нижнего уровня объектам верхнего уровня. Иерархическую БД образует, например, каталог файлов, хранимых на диске, а дерево каталогов – наглядное демонстрация его структуры.

Пример иерархической модели данных



Сетевая БД отличается большой гибкостью, так как в ней существует возможность установить дополнительно к вертикальным иерархическим связям горизонтальные связи. Это облегчает процесс поиска нужных элементов данных, так как уже не требует обязательного прохождения нескольких иерархических ступеней.

Пример сетевой модели данных



Реляционными (от англ. relation – отношение) является БД, содержащая информацию, организованную в виде прямоугольных таблиц.

Реляционные БД (РБД) характеризуются простой структурой данных, удобным для пользователя табличным представлением. Реляционная таблица представляет собой двухмерный массив и обладает следующими свойствами:

- каждый элемент таблицы – один элемент данных;

- все столбцы в таблице однородные, т.е. элементы в столбце имеют одинаковый тип (числовой, символьный и т.д.) и длину;
- каждый столбец имеет уникальное имя;
- одинаковые строки отсутствуют;
- порядок следования строк и столбцов может быть произвольным.

Отношения в реляционной БД представлены виде таблиц, строки которых соответствуют **записям**, а столбцы – **полям**.

Запись – полный набор данных об определенном объекте, т.е. это информация расположенная в строках.

Поле – элемент таблицы, который содержит данные определенного рода, расположенных в столбцах.

Таблица представляет отношение СТУДЕНТ

Отношение
СТУДЕНТ
(вся) таблица

Схема отношения
(строка
заголовков)

Атрибуты (поля)
(отдельные
заголовки)

Строка
(запись,
кортеж)

ФАМИЛИЯ	ИМЯ	ОТЧЕСТВО	ФАКУЛЬТЕТ	КУРС
Иванов	Иван	Иванович	Физической культуры	1
Петров	Петр	Петрович	Филологический	5
Сидоров	Александр	Александрович	Физ.-мат.	2

Поле, каждое значение которого
однозначно определяет
соответствующую запись,
называется **простым ключом**
(**ключевым полем**).

СТУДЕНТ

Код студента	ФАМИЛИЯ	ИМЯ	ОТЧЕСТВО	ФАКУЛЬТЕТ	КУРС
00312	Иванов	Иван	Иванович	Физической культуры	1
00313	Петров	Петр	Петрович	Филологический	5
04303	Сидоров	Александр	Александрович	Физ.-мат.	2

СТУДЕНЧЕСКИЙ СПОРТИВНЫЙ КЛУБ

Код студента	Баскетбол	Футбол	Хоккей	Бокс
00312	+	-	-	-
00313	+	-	-	+
04303	-	+	-	-

Ключевое поле

Объекты базы данных

Таблицы. Это основная категория объектов в реляционной **СУБД**, Работа с таблицей выполняется в двух основных режимах: в режиме *конструктора* в режиме *таблицы*.

Запросы. Объекты этого типа служат для получения данных из таблицы или нескольких таблиц. Отбор нужных сведений производится на основе формулируемых критериев.

Формы. Этот тип объектов используется в основном для удобного вывода данных. Форма представляет собой как бы бланк, который следует заполнить. Заполнение такого бланка под силу пользователю даже невысокой квалификации.

Отчёты. Объекты-отчеты предназначены для отображения данных в виде удобном для просмотра. На основе отчета может быть создан документ, который будет распечатан или включен в документ другого приложения.

Макросы. Макросами называются «макрокоманды», которые запускаются простым нажатием нескольких клавиш и могут выполнять такие действия, как открытие таблиц и форм, выполнение опций меню, управление окнами и др. Пользователь может создавать свои макросы для последовательностей часто применяемых им операций.

Модули. Этот тип объектов представляет собой программные модули, написанные на языке VBA. Модули – это обычно процедуры для обработки событий или выполнения вычислений. Разбиение на модули облегчает процесс составления и отладки программы.

Свойства полей базы данных

Поля базы данных не просто определяют структуру базы – они еще определяют групповые свойства данных, записываемых в ячейки, принадлежащие каждому из полей. Ниже перечислены основные свойства полей таблиц баз данных на примере СУБД *Microsoft Access*.

Имя поля – определяет, как следует обращаться к данным этого поля при автоматических операциях с базой (по умолчанию имена полей используются в качестве заголовков столбцов таблиц).

Тип поля – определяет тип данных, которые могут содержаться в данном поле.

Размер поля – определяет предельную длину (в символах) данных, которые могут размещаться в данном поле.

Формат поля – определяет способ форматирования данных в ячейках, принадлежащих полю.

Маска ввода – определяет форму, в которой вводятся данные в поле (средство автоматизации ввода данных).

Подпись – определяет заголовок столбца таблицы для данного поля (если подпись не указана, то в качестве заголовка столбца используется свойство Имя поля).

Значение по умолчанию – то значение, которое вводится в ячейки поля автоматически (средство автоматизации ввода данных).

Условие на значение – ограничение, используемое для проверки правильности ввода данных (средство автоматизации ввода, которое используется, как правило, для данных, имеющих числовой тип, денежный тип или тип даты).

Сообщение об ошибке – текстовое сообщение, которое выдается автоматически при попытке ввода в поле ошибочных данных (проверка ошибочности выполняется автоматически, если задано свойство Условие на значение).

Обязательное поле – свойство, определяющее обязательность заполнения данного поля при наполнении базы;

Пустые строки – свойство, разрешающее ввод пустых строковых данных (от свойства Обязательное поле отличается тем, что относится не ко всем типам данных, а лишь к некоторым, например к текстовым).

Индексированное поле – если поле обладает этим свойством, все операции, связанные с поиском или сортировкой записей по значению, хранящемуся в данном поле, существенно ускоряются. Кроме того, для индексированных полей можно сделать так, что значения в записях будут проверяться по этому полю на наличие повторов, что позволяет автоматически исключить дублирование данных.

Типы данных

Тип данных определяется значениями, которое предполагается вводить в поле (столбец), например, текст или число.

Текстовой – для хранения одной строки текста;

Поле MEMO – для хранения произвольной информации;

Числовой – для чисел;

Дата/время – для даты и времени;

Денежный – для представления денежных величин;

Счетчик – этот тип предназначен для ввода числа, которое автоматами увеличивается на единицу при добавлении в таблицу новой запаси. Данные этого типа занимают 4 байта;

Логический – предназначен для хранения логического значения **Да** или **Нет**, **Истина** или **Ложь**. Такого рода поле занимает 1 бит;

Объекты OLE – для хранения объектов из других приложений, поддерживающих технологию OLE;

Гиперссылки – служит для записи в таблицу гиперссылок (пути URL).

Безопасность баз данных

Базы данных – это тоже файлы, но работа с ними отличается от работы с файлами других типов, создаваемых прочими приложениями. Для баз данных предъявляются особые требования с точки зрения безопасности, поэтому в них реализован другой подход к сохранению данных.

Базы данных — это особые структуры. Информация, которая в них содержится, очень часто имеет общественную ценность. Нередко с одной и той же базой (например, с базой регистрации автомобилей) работают тысячи людей по всей стране. От информации, которая содержится в некоторых базах, может зависеть благополучие множества людей.

Поэтому целостность содержимого базы не может и не должна зависеть ни от конкретных действий некоего пользователя, забывшего сохранить файл перед выключением компьютера, ни от перебоев в электросети.

Проблема безопасности баз данных решается тем, что в СУБД для сохранения информации используется двойной подход. В части операций, как обычно, участвует операционная система компьютера, но некоторые операции сохранения происходят в обход операционной системы.

Операции изменения структуры базы данных, создания новых таблиц или иных объектов происходят при сохранении файла базы данных. Об этих операциях СУБД предупреждает пользователя. Это, так сказать, глобальные операции. Их никогда не проводят с базой данных, находящейся в коммерческой эксплуатации, – только с ее копией. В этом случае любые сбои в работе вычислительных систем не страшны.

С другой стороны, операции по изменению содержания данных, не затрагивающие структуру базы, максимально автоматизированы и выполняются без предупреждения. Если, работая с таблицей данных, мы что-то в ней меняем в составе данных, то изменения сохраняются немедленно и автоматически.

Обычно, решив отказаться от изменений в документе, его просто закрывают без сохранения и вновь открывают предыдущую копию. Этот прием работает почти во всех приложениях, но только не в СУБД. Все изменения, вносимые в таблицы базы, сохраняются на диске без нашего ведома, поэтому попытка закрыть базу «без сохранения» ничего не даст, так как все уже сохранено.

Таким образом, редактируя таблицы баз данных, создавая новые записи и удаляя старые, мы как бы работаем с жестким диском напрямую, минуя операционную систему.