

Databases

What the Specification Says

Describe flat files and relational databases, explaining the differences between them;

Design a simple relational database to the third normal form (3NF), using entity-relationship (E-R) diagrams and decomposition;

Define and explain the purpose of primary, secondary and foreign keys;

Describe the structure of a DBMS including the function and purpose of the data dictionary, data description language (DDL) and data manipulation language (DML);

Use SQL to define tables and views, insert, select and delete data and to produce reports.

Notes

Databases

- A file is a collection of sets of similar data, called records.
- An item is what each item of data within a record is called.
- The items are stored in fields.
- Records all have the same sort of contents, but relates to a different object, they are usually represented as a row in a table.
- A database is a collection of data arranged into related tables. How it is arranged depends on its normal form.

Flat Databases

Originally all data was held in files, which consisted of a large number of records, each containing a larger still number of fields. Each field has its own data type and stores a single item of data.

This led to very large files that were difficult to process, and quite inflexible not making the best use of computer resources. The disadvantages of it are:

- Separation and isolation of data.
- Duplication of data.
- Data dependence.
- Incompatibility of files.
- Fixed queries and the proliferation of application of programs.

A solution to this is to use a relational database.

Relational Databases

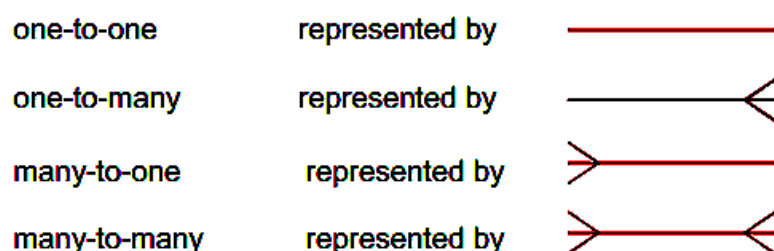
- Instead of individual unrelated files, the data is stored in tables which are related to each other.
- Each table has a key field by which their values in that table are identified.
- The records (or entities) in the tables can be related to entities in other tables by sharing keys as attributes within the entities
- There is less data duplication, as relationships are used to link matching fields.

Arranging Data

- A normal form is the name given to how data in a database is arranged
- First normal form (1NF) is where each table has no repeating groups
- Second normal form (2NF) is where the values of the attributes are all dependant on the primary key
- Third normal form (3NF) is where no attributes are predictable because of one of the other attributes

Entity Relationship Diagrams

These are used to illustrate the relationships between entities. If 3NF is used, there will be no many-to-many relationships, and one-to-one relationships can be combined. Many-to-many relationships can be replaced by a link entity.



Keys

- The primary or key field is a unique field used to identify a record
- A foreign key is a primary key in one table used as an attribute in another to link tables though providing relationships
- A secondary key is a field in a table that can be used to access the data in different ways. It is used to search for a group of records.
- The candidate keys are all the possible primary keys, i.e. all the unique fields

Database Management System (DBMS)

- Contains the data definition language (DDL)
- Contains the data manipulation language (DML)
- Allows data to be amended and controlled
- Includes a data dictionary which is a file of descriptions of the data and structure.

Data Definition Language (DDL)

- Part of the DBMS
- Used by the database designer to define the tables of the database
- Allows the designer to specify data types and data structures as well as any constraints on the database
- Cannot be used to manipulated the data
- Just creates tables and structures that hold information about the data that will be put in the tables, such as validation checks.
-

Data Manipulation Language

- Allows data to be manipulated
- Users of the database will have different rights

Views of Data

The DBMS can be made to present various views of the data held in the database.

- Internal level, this is the view of the entire database store in the system. It is hidden from the user by the DBMS.
- Conceptual level, gives a single usable view of all the data on the database.
- External level, where data is arranged according to user requirements and rights. Different users will get different views of the data.

Key Words

- **File** – a is a collection of sets of similar data called **records**
- **Table** - another name for a file
- **Tuple** - another name for a record
- **Item** - an item of data within a recorded
- **Field** - the area where items are stored
- **Attribute** - another name for a field
- **Database** - a series of related files, called **tables**
- **Primary** or **key field** - a unique field in a record
- **Foreign key** - the field which contains a link to another field
- **Secondary key** - a field in a table that can be used to access the data in different ways
- **Normal Form** - the name given to how data can be arranged in a system
- **1NF** – where there are no attributes that have multiple data in them
- **2NF** – where the values of the attributes are all independent of on the primary key
- **3NF** – where no attributes are predictable because of any other attributes
- **Link Entity** – additional table used to join many-to-many relationships in 3NF
- **DBMS** – database management system, the piece of software that allows databases to be easily viewed and edited and ensures that the rules remain unbroken. Includes DML, DDL and the data dictionary.
- **DDL** – data description language, used to define tables in the database as well as the data types and data structures to be used, and any validation checks required.
- **Schema** – The database design produced by the DDL.
- **DML** – data manipulation language, allows different users to carry out operations on the data. (e.g. amend, delete, insert data). Different users will have different permissions.
- **Data Dictionary** - a file maintained by the DBMS which includes the descriptions of the data and structure of the storage of the data.
- **Internal Level** – a view of the entire database as it is stored in the system. The data is organised as it is stored, usually hidden from the user by the DBMS.
- **Conceptual Level** – gives a single usable view of all the data on in the database.
- **External Level** – where the data is arranged according to the user requirements and rights. Different users will get different views of the data.
- **SQL** – structured query language, a language that allows the user to set up their own queries on the database.

Past Exam Questions and Answers

What is the use of the primary key?

unique identifier

What is the use of the foreign key?

primary key from one table...
...used as an attribute in another...
...to link tables/provide relationship between tables

What is the use of the secondary key

used to search for a group of records

What is a data dictionary?

a file containing descriptions of data in database
used by database managers...
...then altering database structure
Uses metadata to define the tables

What are the benefits of a relational database compared with a flat file database?

avoid data duplication/save storage
data consistency
data integrity
easier to change data
easier to change data format
data can be added easily
data security/easier to control access to data.

Explain why a foreign key is also a primary key, but a primary key need not be a foreign key

foreign key links tables (to represent many to one relationship)...
...so that only one record is accessed/to avoid duplicate data
eg primary key from B used as foreign key in C from (a)
primary key is in a table that may contain data not required in another table
eg primary key from C is not used in B and hence cannot be a foreign key

What's the difference between char and varchar data types

CHAR is fixed length
VARCHAR is variable length

Why are different views of data made available to users in a database?

so users can access the data they need
users do not need specialist knowledge
to protect data
to prevent unauthorised access.

What are the consequences of a many-to-many relationship?

not allowed/not in 3NF
needs another table between Student & Subject...
...to avoid duplication of data/to change to 3NF

What is a report, and what are the key features?

presentation of selected data...
...usually in the form of a table/specific layout
may be defined in advance...
...so the user does not need to set it up
Features of report definition: a query, a display order