A2 MODULE 5 (ICT5) 14.3 DATABASE MANAGEMENT CONCEPTS

- *Explain the purpose of a database management system (DBMS). (Chapter 55, 57)*
- *Explain the role of the database administrator. (Chapter 57)*
- *Explain what is meant by data consistency, data integrity, data redundancy and data independence.*
- *Explain the concept of entity relationships and data normalisation. (Chapter 56)*
- *Explain the concept of a client/server database.*
- *Recall the relevant advantages of a client/server database over a non-client/server database.*

server based	*Peer to peer.*

Problems when using a flat-file approach (see the attached bank example) p.300

1. **Data redundancy** is found- the presence of duplicate data in multiple data files e.g. Customer Name, NI number, Address is present in several different files on several different systems. Errors are generated, time entering data is wasted, computer resources are needlessly taken up and updates can be an enormous problem.
2. **Data inconsistency** will become a problem e.g. information is **duplicated** in each system and may be **updated** in most systems but **not** necessarily in all – the savings account and loan account may have different addresses for a particular customer.
3. **Lack of data independence. (Data dependence** refers to the close relationship between data stored in files and the specific software programs needed to update and maintain those files). In the bank imagine cases where any change in data format or structure requires a change in all the programs that access the data. e.g. effort required to change from 3-digit STD to 4-digit STD may be very difficult.
4. **Data lacks integrity ,** i.e. this is the quality by which information from the sysstem can be trusted. The problems already stated show that information can be out of date, can have different values in different parts of the system, can be inaccurate etc.
5. **Lack of flexibility i**.e. creating reports, e.g. in this system a summary of account balances in each separate account would require separate reports for each open account or cutting and pasting into a wordprocessor file – not a satisfactory state of affairs.
6. **Data not shareable**. Similarly if different departments want to access data this would be very awkward.

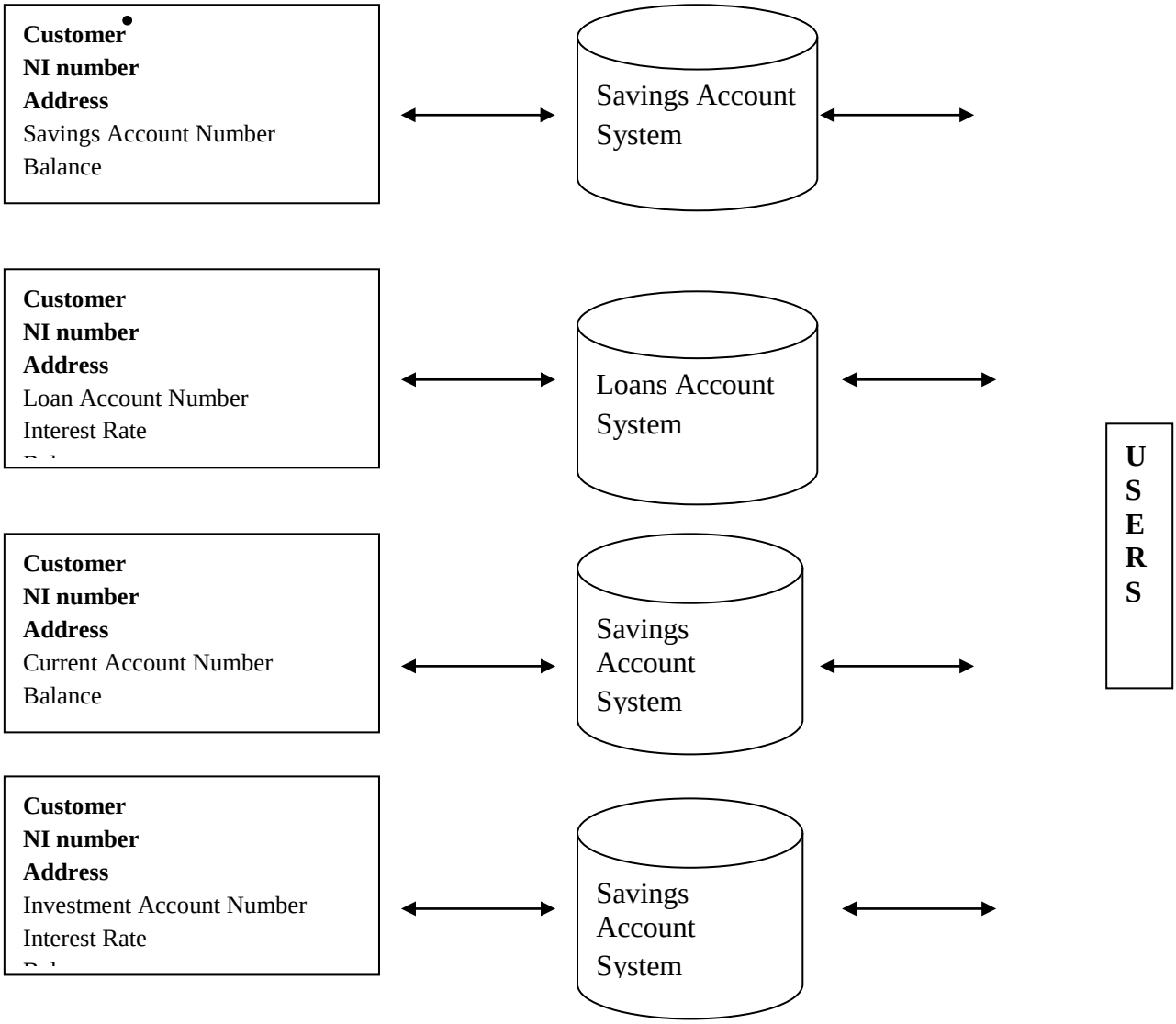
Note the issues that remain given in P.300

- **Unproductive maintenance** i.e. updates to information and changes to file structure can be unweildy.
- **Lack of security** i.e. there are no restrictions on who can see what.

A Bank using a Flatfile

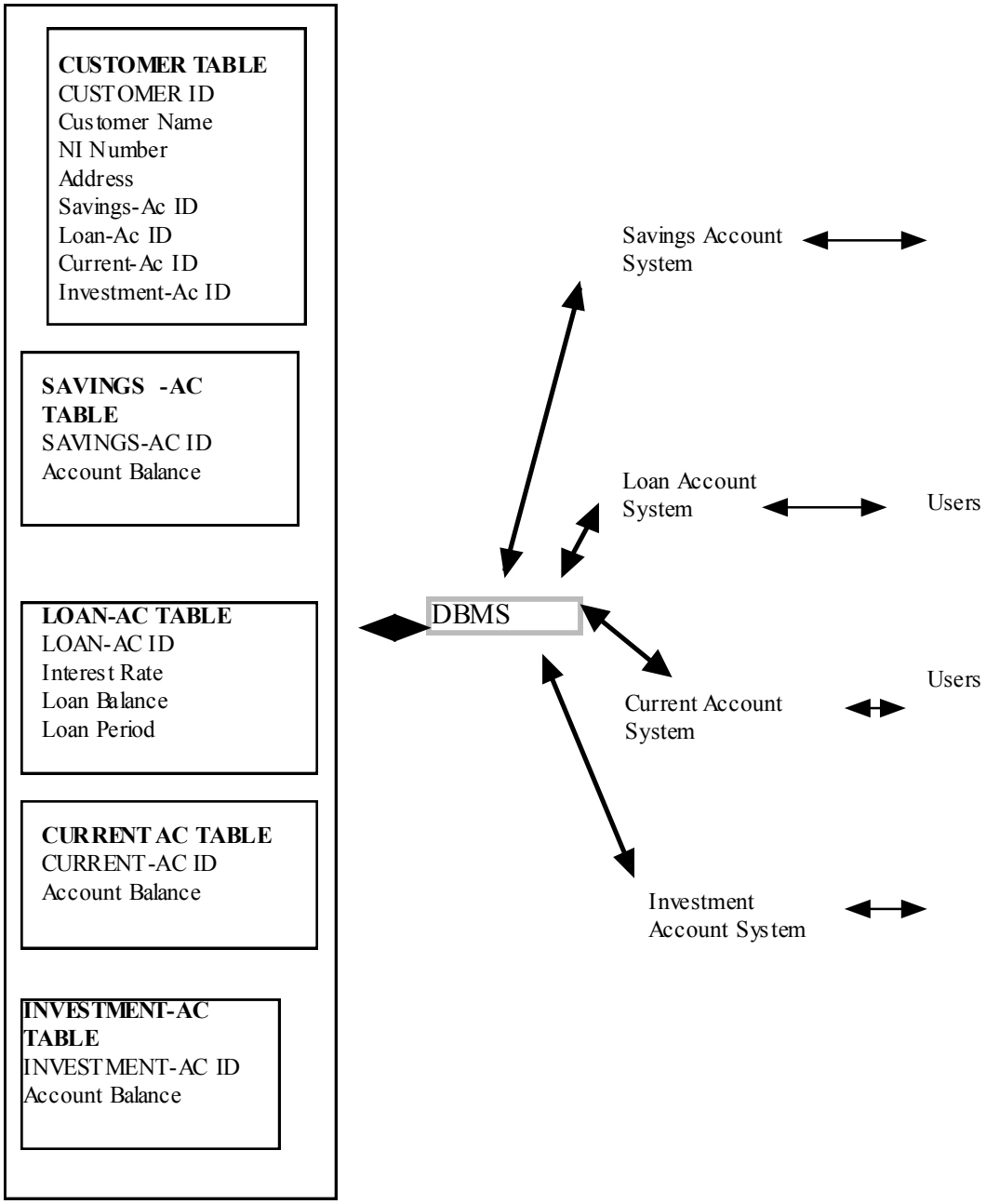
- Different Information Systems have grown up over time, often with their own software programs and ways of doing things.
- Information is duplicated throughout the system; updates may occur at different times, some errors remain undetected.

Costs to run the system are driven up because of clerical time needed to input data and check for errors, it is very difficult to combine information.



A DBMS APPROACH TO THE SAME PROBLEM

- data is combined into a single customer database (pooled data)
- data available to multiple applications and users



The Database Approach p.300 http://databasepronews.com/2003/0522.html

- **Flat-File or card index Databases -** contain one file. The records can only be accessed in a limited number of ways, and the number of fields in a record will be limited.
- **A Relational Database** is designed to handle data in TABLE form and a single database is likely to contain a number of separate but related, tables.

Database management systems (DBMS) P301 and 311

- A DBMS is application software that lets you create, organise, update, store, and retrieve data from a single database or several databases. Lets you transform or map data from one model to another, or between the central model and stored database.
- Teachers access their classes electronically but have a specific view.

The DBMS (see p.311)

• **Builds the database** – creates and maintains the database dictionary.

• **Manages the database** - has to ensure problems do not arise if two people simultaneously access a record and try to update it.

• Governs interactions between application programs, input data and the database itself i.e. allows users to **store, retrieve and update** as easily as possible without having to be aware of the internal structure of the database.

• **Provides an interface** between users and the database

• **Backup and Recovery** in event of a system failure.

• **Security** - password allocation and access rights to particular layouts.

Data Models p.301

- Although a single collection of data it will appear different when viewed from different parts- by different users and applications. A data model is the logical structure of the data as it appears at a particular level of the database system e.g. **mail** order retailer consider stock control and accounting.

Entities

- When a new system is to be designed, the systems analyst will identify the **entities (or tables) that exist within it.**
- **An entity is** a person, place or thing on which information is maintained - e.g. employee in a personnel file, Product, Supplier etc. (p.301).
- Examples of entities in a computerised mail order system might be customer, order, invoice, delivery note, stock item, supplier etc.

Attributes

- These tables are organised into fields **(attributes)** which are characteristics of entities. (p.302)
- **ATTRIBUTES** e.g. a Personnel database an entity Employee may have attributes employeeCode, Name, JobCode, JobTitle, DepartmentCode etc. These attributes determine the fields in the Employee table. Some attributes may be kept in a different table e.g. JobTitle is likely to have same value in numerous records

Relationships

- One of the jobs of the system designer is to **map the relationships between entities**. This helps to ensure that the system will have the necessary data structure and that redundancy will be avoided.
- Relationships between entities can be one to many, many to many and one to one.
- Records within tables are so organised that they can be accessed in any way, i.e. different entities can be combined in a number of ways because of the **relationships** established between them

Types of relationships (P.302)

In a RDBMS relations are of three types (read p.279). I will give school contexts

1 ONE-TO-ONE

each student develops a unique student history and each student record can only relate to one student.

Staff		**Student**

2 ONE-TO-MANY

Each course has many different students but the student-takes table concerns one student.

Staff		**Student**

3 MANY-TO MANY

Staff		**Student**

Students can be taught by several staff and staff will teach many students
Courses will have many students and students can take many courses
The same course can be taught by several staff (A-Level IT) and staff can teach several courses

A relational database management system (RDBMS)

- There are other DBMS approaches (Object oriented etc.)
- The easiest to understand and the **most flexible**- still the Industry standard approach..
- The only data structure in an RDBMS is the relation itself..
- In making queries information from different tables are combined as the tables share a common data element. See p.302.
- Each record within a well-designed table will have a **KEY FIELD,** which is unique to
- that record.
- Relationships between tables are shown using **LINK FIELDS,** there may be several
- relationships from one table to other tables. It would be highly unlikely to have more
- than one link between the same two tables.
- Using these relations can generate **REPORTS** using data from several tables.

Advantages of using an **RDBMS**. http://www.exel.co.uk/selection_process.asp

1. Single centralised store of data for all applications in the organisation that can then be pooled.

2. Independent of application program - many different applications can use data from common shared database(s)

3. Data is consistent: when an attribute in a table is updated, its up-to-date value is available to all users of the RDBMS, in whatever report they use and in exactly the same form.

4. Less data redundancy- because there is only one copy of each attribute kept-duplication should be eliminated altogether in a well-designed RDBMS

5. Ease of reporting/flexibility -easy to set up new relationships and new entities. New tables and reports can be set up as and when required.

6. Easier security- all access to data is via a centralised system, a uniform system of security monitoring can be implemented.

Link Tables

It is often easier and more convenient to deal with Many-to-Many relationships by splitting them into two one-to-many relations by creating **a Link Table** such as student-takes, (see p.305).
In any examination and in ICT6 you must explain this method.

Standard Form

TABLES for a School system can be written in standard form.

STUDENT (STUDID. Surname, Forename, Address, Sex)

STUDENT TAKES(STUDID, CRSECODE. StaffNo)

COURSE (CRSECODE. CrseTitle)

STAFF (STAFFNO. StaffName)

Key fields are underlined and capitals used.

Composite Keys

The STUDENT_TAKES table which is the link table needs **a composite** key to uniquely identify the course taken by a particular student.

Students take several courses so this does not uniquely identify a record or transaction.

Courses are taken by many students so more information is needed.

But student number and course ID will give us the record we are seeking.

You must explain this in questions and your projects.

Reporting

Draw up the following Reports for the tables above (colour in a different colour where information is used from different tables)

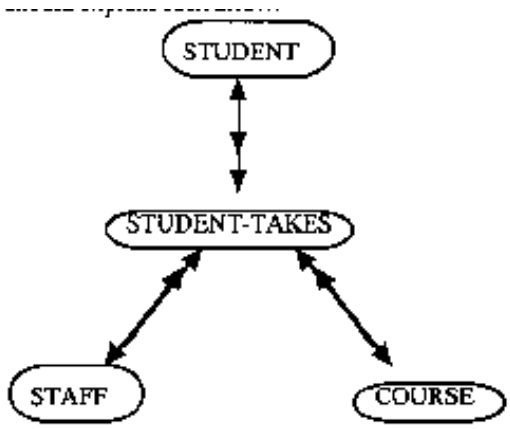
Report 1 A Student Record Card including personal information, course **and** staff details.

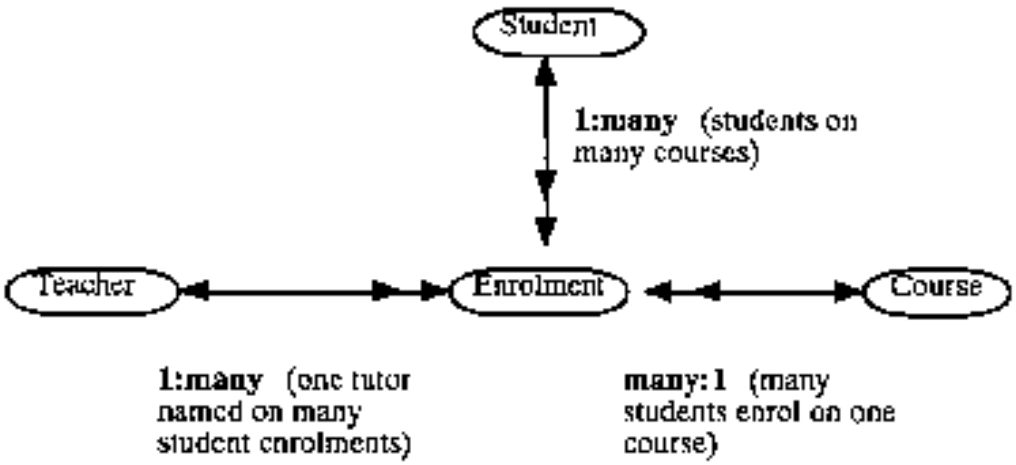
Report 2 A Class list showing student and staff details

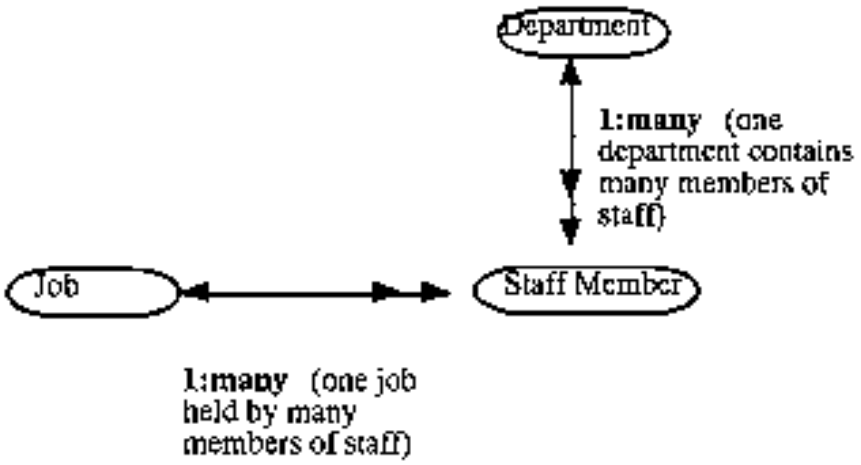
Report 3 A list showing courses taught by each staff member.

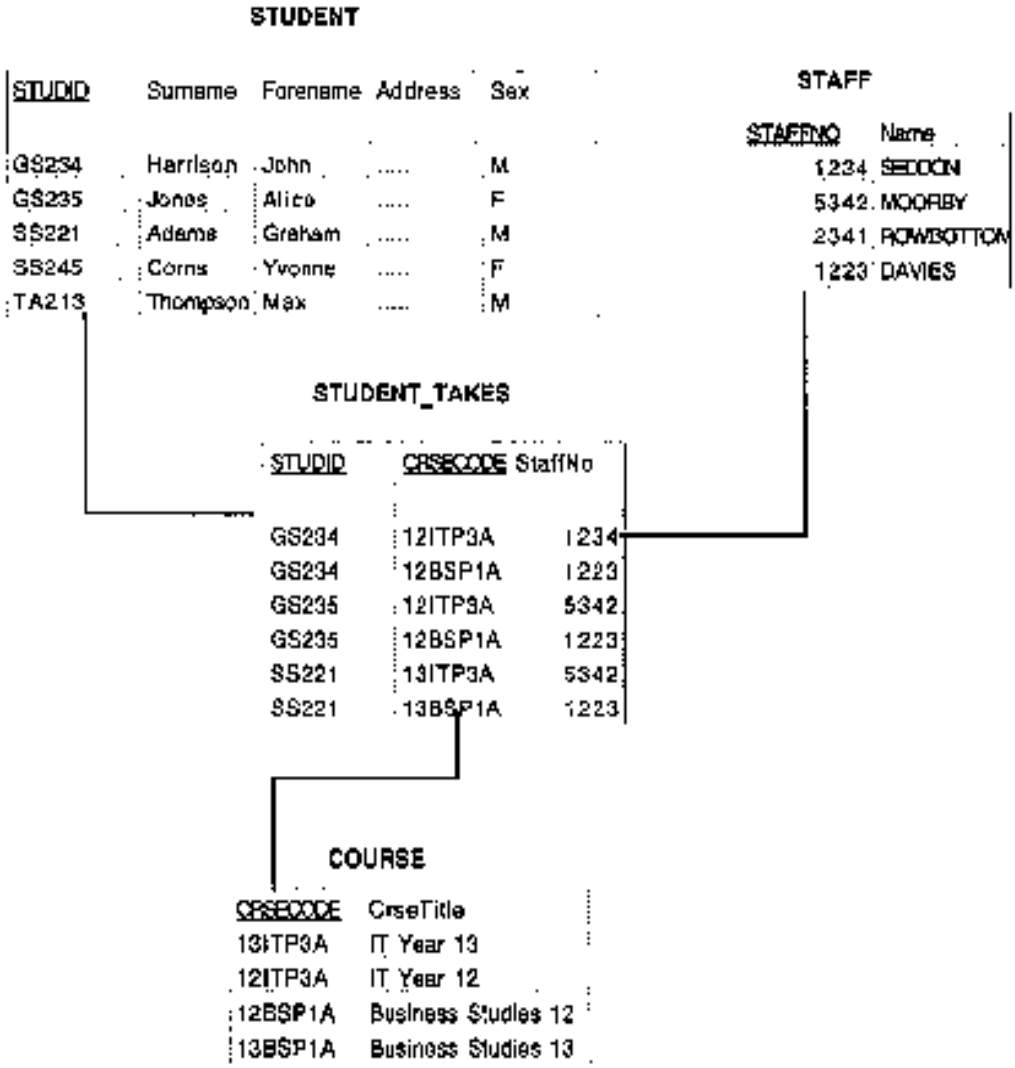
AN ENTITY RELATIONSHIP DIAGRAM (ERD) p.302

As below is an important aid to designing a RDBMS and explaining your logic. You should explain each arrow.









- **LINK FIELDS** are clearly indicated in this example in order to establish relationships

- **Link fields** are indicated in this example in order to establish relationships between tables.
• CRSECODE appears in the Course table as a **primary key** but in the STUDENT_TAKES table as a **foreign key** (p.304). Label primary and foreign keys on the diagram.
- The composite key in thestudent-takes table has been described.

Index

Special index files created by database software - a list in numerical or alphabetical order for a file or database of the key field o I'each record and its associated storage location. Speed up querying the master file - do not need to read the entire database to find a particular item. Indexed fields other than the primary key field are known as **secondary keys.**

NORMALISATION (p.305 to 308)

Purpose of Normalisation **http://www.gslis.utexas.edu/~l384k11w/normover.html**

The aim of all well-designed RDBMS. A step-by-step set of rules by which data is put in its simplest forms i.e. distinct entities and attributes. Thus in a fully normalised RDBMS there will

- Reduced data redundancy i.e. no unnecessarily duplicated data
- All attributes will be independent
- Database structure is flexible i.e. it will be possible to add new data and entities without a wholesale reorganisation of the database structure.
- Data is consistent throughout the database
- Complex queries required by the user should be possible

Example: Student and Course
In a flatfile this would be represented as

STUDENT(Student number, student name, DOB, Sex, course number, course name, lecturer number, lecturer name)

Because this is a many-to-many relationship many course entries would need to be entered . **We do not want repeating attributes** such as course1, course2, course3 (see p.306).

1NF (First Normal Form) No repeating attributes

To be in the first normal form, a table should contain no repeating attributes. This avoids data inconsistency.

Divide the information into two entities in standard notation, course number becomes part of the composite key in the student table.

STUDENT(Student number, student name, DOB, Sex, course number)

COURSE (course number , course name, lecturer number, lecturer name)

2NF (Second Normal Form) Partial key dependence test

No partial dependencies are allowed. This involves removing many-to-many relationships.
In addition to being in first normal form (1NF), each non-key data item in the record is fully dependent on the primary key; i.e. the primary key alone could be thought of as determining the other data in the record. (This is called Second Normal Form). The term fully dependent here requires that each non-key item must depend on the whole of the primary key. In some cases – particularly when the primary key is concatenated – there may be non-key data items that are dependent on only part of the primary key. In this case the record structure is not in 2NF.

Student name is only dependent on Student number and not on course number)
A third table is needed to **link** between the two entities.

STUDENT(Student number, student name, DOB, Sex)

STUDENT-TAKES(Student number, course number)

COURSE (course number , course name, lecturer number, lecturer name)

3NF (Third Normal Form) Non-key dependence test

No non-key dependencies are allowed. No field should be dependant on another field (unless the field is a key field). This prevents any duplication of attributes. In addition to being in second normal form (2NF), each non-key item in the record is fully independent of any other non-key data item in the record.

Even though lecturer no. is not a key there are issues here – it is not dependent on course number which is the key in this table.

We therefore need to create a new relation and table**.**

STUDENT(Student number, student name, DOB, Sex)

STUDENT-TAKES(Student number, course number)

COURSE (course number , course name, lecturer number)

LECTURER(lecturer number, lecturer name)

It is important to remember that the purpose of normalisation is to produce record structures that are independent of the applications that may need to access them. The applications will access virtual records that will draw related data from a number of different tables using the foreign keys as pointers to put the data together.

Database Management (Chapter 57)

The role of the database administrator (p.310)

http://www.dfas.mil/technology/pal/ssps/slc/views/prtdbaonly.htm

A DBA is a full time manager responsible for maintaining DBMS and ensuring accuracy and the integrity of data.

- What data should go into the database
- What relationships should exist between different data items
- Who has permission to read the database information (allocate passwords)
- Who has the authority to update the database.
- Be involved in the design of the database and make changes if necessary
- Keep users informed of any changes
- Maintain the data dictionary
- Provide training to users

The contents of the data dictionary p.310

http://www.witnessminer.com/eample_2_data_dictionary.htm

This is a database about the database. It includes information about:
Tables
Data lengths and field types
Data validation restrictions
Descriptions of fields
Relationships

Defining a database structure

primary/key fields –

Every entity in a database must contain at least one field that uniquely identifies that record so it can be retrieved (accessed), updated or sorted, e.g. employee number, N1 number, ISBN no., student no., course code etc.

Types of fields- these will include

Alphanumeric	numeric
logical e.g. y/n, m/f	date
Input Masks	Container etc.

Field lengths –

logical and date fields usually have predetermined length typically 1 and 8 respectively. Usually numeric needs to have a maximum length specified.

Coding data

without obscuring meaning saving space queries can be shorter data entry will be quicker

Querying a Database p.311

Query is database jargon for searching a file for records that meet certain criteria.
Secondary Keys. Apart from the primary keys used to identify records uniquely in an RDBMS can also query on secondary keys e.g. Prod ID

QUERY BY EXAMPLE "QBE"
is characterised by filling inquiry conditions into
empty slots under displayed field information, examples

AND requests	"Johnson" and "Personnel" to find all the records for people named Johnson who work in Personnel
OR requests	to find all records that contain either New York or San Francisco in the City field make two requests using New Request or Duplicate Request then Find (Use Paste Special to use Index)
Wildcard searches e.g.	To find Gray and Grey type Gr@y in the field. To find McKineh and McKennah type McK*h in the field
Numbers, dates, times and finding a **range of values** e.g.	12:30...17:30 finds records **that** contain times between 12:30 to 17:30 in that field
Records that don't match the request	e.g. type description of what you want to omit then click **Omit**
finding some records and omitting others Sorting information Operators usually include =, **o,** **<,** **>,** <=, **>=,** AND, OK, .NOT	

STRUCTURED QUERY LANGUAGE SQL p.312

(A variation of command mode) is the database access standard for relational database management systems (RDBMS) - enabling to exchange files with other databases.
Developed by IBM in the mid-1970s for mainframe and minicomputers.
The basic structure of a SQL query for retrieving data is

SELECT <columns>	identifies the columns or data fields to retrieve
FROM <tables>	specifies the tables or files from which to retrieve this information.
[WHERE <condition>]	restricts the information output to only those records or rows matching a specified condition.

QBE is automatically converted by the software to a SQL request.- can be retrieved from any source with standard SQL exchange facility. This way most users never deal with SQL even though database in effect handles SQL language requests.

Examples using SQL

Selection by Criteria

If a Sales rep wants the Name and Address ofCUST 1234 in the example given above the **SQL** statements would take the **form** SELECT Name, Address FROM CUSTOMER WHERE Number =1234

Updating the Database

SQL can also change values in the database: for example **to** give customer 1234 a 10 per cent discount, the following statements may be used
UPDATE SALES
SETPRICE-PRICE*0.90
WHERE CUST NO-1234

Client-Server networks p315

Networks with powerful servers - in which different pieces of hardware work on the same processing problem. The network tries to use the full processing power of each computer in the network, including the **file-server** and the user workstations or **clients**.

Servers are required to: -

Manage the activities of the network

Can store application programs and data files

Can distribute programs or data files to other computers on the network as they request them.

A client-server solution has enabled companies to downsize from the sole use of mainframes. **Advantages** include:-

It is less expensive and more flexible than mainframe.

Power of client and servers are used jointly (p.163)

Administration e.g. backup and file-sharing can be dealt with by server

Printing can be queued

http://penguin.dcs.bbk.ac.uk/academic/networks/application-layer/client-server/index.php

A Client-Server network requires:-

- A file-server
- Network cards which provides a physical interface
- A networking OS such as **Windows NT**

Specialist servers might be used such as a

- File server
- Print server
- Communications server.
- Database servers for example provide access to central database files for all workstations or "clients" on the network.

Peer-to-Peer networks

A peer is an **equal** i.e. each node in such a network is equal to any other node. If the node is a computer it can also act as a stand-alone device, and is also able to share data back and forth with other nodes at any time. If the node is a fax, printer, modem, tape-backup drive etc. it can be used and shared by any other node. It is therefore not necessary to have file server since all workstations contribute to control of the network. Performance is slowed down and file management is complicated. Windows for Workgroups (up to 10, LANtastic 6.0)

See example of Sage accounting software p.312

EXAMINATION QUESTIONS

Please note this topic was formerly in IT02 until 2001.

1995.11 (17 Marks)

A college information system currently uses three sets of files for its student records, staff records and finance records, each being run separately. The systems manager is keen to introduce a Database Management System (DBMS) based on a relational database, claiming that this will have major benefits for the college.

a) By the use of an appropriate example explain what is meant by the term 'relational database'. (5)

b) Describe three advantages of a DBMS approach in contrast with the use **of** independent files. (6)

c) During the design of the system a decision is made to restrict the access of different users in different ways. Describe three different restrictions that may be imposed upon different users (6)

A SUGGESTED FORMAT FOR AN **ANSWER**

a) data arranged into 2-D tables (1
At least one of the other tables must contain at least one common attribute (1) Data from more than one table can be extracted and combined together (1)
Normalisation (1) here or in part (b) + (1) if fully explained example (1) for showing two tables with a common attribute (1) for showing the link...this must be coded
additional alternatives
a collection of data organised in such a way that it may be accessed by a variety of applications programs (1)
a number of interrelated files (1)
tables are not sorted into any order (1)
one to many and many to many relationships (1)
OR
a number of files linked together
data is shared between the files
plus a detailed and consistent example

***WRITE OUT** the three ENTITIES **given in the question** in standard form*
i.e. STUDENT(STUD1D, Surname, Forename,.....) etc. The examiners want to see tables with common attributes to enable links.
DRAW UP the TABLES
***ENTER** sample data into the tables*
***LABEL** the key field*
***WRITE OUT** something along the lines 'this is the key field, it uniquely identifies this record and is coded)*
***DRAW** the links between tables and common attributes*
***WRITE OUT** something along the lines "these are two reports that can be produced by combining the __ table and the __ table. " with description*
***WRITE OUT** something along the lines "to be fully efficient the tables must be further normalised i.e. put into their simplest forms without duplication"*

(b) the sharing of data between applications (2)
the maintenance of consistent data (2)
the logical independence of data (2)
data control and security (2)
the ease with which new application (reports or code) can be generated (2)

less data duplication (2)

(c) different restrictions within the same department (2) read-only access (1) vs edit access (1) location and terminal (2)

different restrictions between departments (2)

control is required so that only authorised users can access (1)

example: student record clerical officer cannot access personnel details (1) passwords (1) plus levels of passwords (1) access levels explained (1) any mention of DP Act (1)

NOT voice, fingerprint or retina recognition

1994.10 and 1996.10 (19 marks)

A Company's sport centre uses a database management system to operate a membership and fixture system. Normally, members register for at least three sports, although they can play any of the sports offered by the centre. Fixtures against many other organisations are arranged in a wide range of sports involving a large number of teams.

- name THREE database files you would expect to find in the system (3)
- For each of the database files you have named, list the fields required to enable this system to be maintained with minimum redundancy. (6)
- draw a diagram to show the relationship between the database files named in part (a) **(3)**
- describe THREE reports which the system might be required to produce (3)
- The manager of the centre intends to send out personalised letters to each of the members. This is done using the mail-merge facility offered by a word processor in conjunction with the database. Explain how this is achieved. (4)

(a)

Database tile 1:Membership details (1)

Database file 2: Coded sports lookup table (1)

Database file 3: Fixtures database (1)

Database file 4: Other Organisations ***(1)***

Database file 5: Results (1)

(b) Database file 1:Membership details ... personal data(l) coded (1) sports registered for coded fields (1) to indicate selected teams (1)

Database file 2: Coded sports lookup table.... sportcode (1), description

Database file 3: Fixtures database.... fixtures (1) storing membership id (1) of those selected for this team this week (1)

(c) Looking for relationship diagram with link fields - (1) for each named in a

(d) Any valid reports but 'non-trivial use'

• simple reports - complete membership list, team list, squash league table (1 max. for any number)

• non-trivial - name and phone number of those selected to play this week, names ***and*** *addresses of those not responded to subscription notice, names of those who play sport X and sport Y for a given use (1 each)*

(e) Standard text letter (1)

• fieldnames in letter (l)

• database ***of records... one*** *letter per* ***record (1)***

• substitute fields from records in database (1)

1993.2 (4 marks)

A bookshop uses an interactive computer database system to store stock details and answer customer enquiries regarding the availability of books. An enquiry for a particular book shows that there should be a single copy remaining in stock. After searching the shelves and stock room the book cannot be found. State and explain **TWO**

possible reasons for this discrepancy.

- ***data*** *input must be qualified by an example appropriate to this information system e.g. wrong ISBN (1)*
- *none for typing error or vague, unspec. errors*
- *initial stock was wrong (1)*
- *updates are batch (1) each evening whereas enquiries are interactive (1). Thus ageing of*
- *information (1)*

1991.5 (6 marks)

You are asked to advise on the purchase of a database management package. Describe the types of information you would need to know about the application before formulating your request.

type of data	*type of query/access routes*
predictability of queries	*anticipated skill of users*
sensitivity of data	*variability of data*
volatility of data	*volume of data*
output formats and volumes	*output 'destinations' e.g. export toanother package*

1997.7 (15 marks)

A college library uses a relational database management system to operate a membership and loans system. Staff and students can borrow as many books as they wish at any given time.

(a) Name **three** database tables that you would expect to find in this system. In each case, identify the columns and keys required to enable this system to be maintained with minimum redundancy. *(6)*

(b)Draw an entity relationship diagram to show the links between the database tables named in part (a). *(3)*

(c)Describe the capabilities of the relational database management system that might be used to identify and output details of overdue loans. *(6)*

(a) *Table names (3 x 1) 1 mark for each table names as below*

Books: *{ BookNo. Book Title, book category} (Key=l or 2 relevant others=l)*

Members {MemberNo. Member Name, Phone number,...} (Key=l or 2 relevant others=l but not age)

Loans {BookNo, MemberNo, LoanDate or return date...} Book & Member No's ***and*** *loan date (1),*

Book number = ISBN (OK)

(b) *Any valid Link between members and loans and any valid link between books and loans (1) Identifying the role of key fields in the links (1)*

Correct use of Lmany, many:many and 1:1 concept (1) LINKS MUST WORK

(c) To gain full marks a candidate must refer to QBE or SQL & REPORTING:

METHOD OF QUERY ***max.*** *(3)* ***QBE:*** *Named (1) and described (1) and example (1). Typically:*

Use of query by example to select subset of records from database by giving values to certain fields to specify selection ***criteria.***

For an example illustrating this (1). A typical example might be ..

OR *Field MemberNo Surname* ***Forename***

Criteria 99 WE**

All members with numbers starting 99 and Surnames starting 'WE' will be selected. ***SQL:*** *Named (1) and described (1). Typically:*

A subset of records could be selected from database by using a programming language ***(1)*** *known as a* ***SQL.***

For an example illustrating this (1). A typical example might be .. select surname, member no from members where surname is like 'WE'*
AND
Report Facility: max. (3)
particular fields selected for output (1) ,order of records can be specified (1) uses the output from a query(l) idea of linked/multiple tables/queries extensive formatting capability (l),grouping and subtotals can be defined (1) example of a report given (1)

1998.10 (13 marks)

A company makes use of a computerised Hat 1'ilc information storage ami retrieval system. **The** company is experiencing problems due to the use of this flat file system.

(a) describe three benefits that the company would gain by using a relational database as opposed to a Hat file system. (6)

(b) The company currently has three files in use; customer, stock and orders. During conversion to a relational database system these files would need to be normalised. Explain clearly what you understand by the term normalisation. (2)

(c) Examples from the three files are shown below. Normalise these files, explaining any assumptions or additions you make to the files. (5)

CUSTOMER FILE

Surname	Forename	Street	Town	City	Post Code
Smith	James	11 The Avenue	Bemersley	Ruston	RS125VF
Penfold	Jayne	67 Bathpool Road	Outclough	Wignall	WG5 6TY

ORDERS FILE

Surname	**Forename**	**Post Code**	**Order Date**	**Item ordered**	**Quantity Bought**	**Price**	**Total Cost**	**Paid**
Smith	James	RS12 5VF	6/5/98	Magic Duster	2	£10.9 9	£21.98	Yes
Penfold	Jayne	WG5 6TY	1/6/98	Banana Rack	1	£12.5 0	£12.50	No ,
Smith	James	RS12 5VF	12/5/9 8	Windsor Doormat	1	£29.9 5	£29.95	Yes
Smith	James	RS12 5VF	12/5/9 8	Easee Food Grater	1	£11.9 9	£11.99	Yes
Penfold	Jayne	WG5 6TY	1/6/98	Windsor Doormat	1	£29.9 5	£29.95	No

Item Name	Price	Quantity in Stock
Windsor Doormat	£29.95	11
Magic Duster	£10.99	34
Electric Potato Peeler	£39.00	0
Easee Food Grater	£11.99	9
Banana Rack	£12.50	1

(a) Data independence - structure does not affect the programs which access the data i.e. set up time for new applications is reduced

Quality of management info. is improved i.e. Info. is more valuable as it is based on a single, comprehensive collection of data
Increased productivity i.e. ad hoc reports can he generatedto meet particular needs
Consistency of data - less data duplication so errors due to discrepancies are reduced
Input preparation reduced to "sirigle input" principle
Control over redundancy i.e. updating less time-consuming as data duplication is minimised
Integrity of data i.e. DBMS can specify constraints when data is added
Greater security of data
Centralised control of data

More information available to users due to centralisation of data
c. Process of removing redundant elements
ensuring database has complete consistency
Allow description of the process
(c) Possible solution;

CUSTOMER FILE

Customer Code	***Surname***	***Forename***	***Street***	***Town***	***City***	***Post Code***
S17	*Smith*	*James*	*11 The Avenue*	*Bemersley*	*Ruston*	*RS12 5VF*
P90	*Penfold*	*Jayne*	*67 Bathpool Road*	*Outclough*	*Wignall*	*WG5 6TY*

INVOICE FILE

Customer Code	***Order Date***	***Invoice Number***	***Total Cost***	***Paid***
S17	*6/5/98*	*010*	*£21.98*	*Yes*
S17	*12/5/98*	*O11*	*£41.94*	*Yes*
P90	*1/6/98*	*012*	*£42.45*	*No*

ORDERS FILE

Invoice Number	***Item Ordered***	***Quantity Bought***	***Price***	***Total Cost***
010	*MD02*	*2*	*£10.99*	*£21.98*
012	*BR01*	*1*	*£12.50*	*£12.50*
Oil	*WD01*	*1*	*£29.95*	*£29.95*
Oil	*FG01*	*1*	*£11.99*	*£11.99*
012	*WD01*	*1*	*£29.95*	*£29.95*

STOCKPILE

Item Code	***Item Name***	***Price***	***Quantity in Stock***
WD01	*Winsor Doormat*	*£29.95*	*11*
MD02	*Magic Duster*	*£10.99*	*34*
PP01	*Electric Potato Peeler*	*£39.00*	*0*
FG01	*Easee Food Grater*	*£11.99*	*9*
BR01	*Banana Rack*	*£12.50*	*1*

Allocate one mark each lor the following to a maximum of live;
Creation of the invoice file (1) and reason given (1) with correct fields (1)
Insertion of customer number (1)
Insertion of Item Code (1)
***Correct removal of name and post code from** orders/invoices (1)*
Removal of item description from orders/invoices (1)

Price remaining in orders with reason (1) Max 5

1999.2 (4 marks) Give four responsibilities of a database administrator.

1999.5 (6 marks)

Database Management Systems provide facilities to extract data from a stored database.

(a) Name **two** common methods of setting up a query. (2)

(b) State **one** advantage and **one** disadvantage of the two query methods stated in part (a).

2002.5

The secretary of a local tennis club is constructing a database to store data on members' personal details and records of attendance. He has been told that a relational database management system can assist him. Having found an article on relational database construction, he does not understand some of the terminology it contains. He asks you for advice.

(a) Explain the following terms:

(i) Normalisation;
(ii) Data independence;
(iii) Data consistency;
(iv) Data integrity. (8 marks)

(b) The secretary constructs his database structure and asks you to examine his work before he enters any data. You notice that he has not included any validation.
With the aid of an example, explain why data validation is important. (3 marks)

(c) Give **three** reasons why he should consult with other members of the tennis club committee before finalising the design of the database system. (3 marks)

a.

i. Process of breaking down complex data structures into simpler forms. (1) + expansion / example (1)

ii. Changes in the structure of the data only affects those programs/ functions that are reliant on that part of the structure (1) + expansion / example (1) OR

Data structure is separate from the programs that access it (1) + expansion / example (1)

iii. Data is only stored once, and this is the sole source of that data. (1) + expansion / example (1)

iv. Correctness / how trustworthy the data is. (1) + expansion / example (1)

b. To check that entered data is sensible (1) + relevant example of validation in context (2,1, 0), e.g. Date of Birth field (1) using a range check (1)

c.

- *to ensure that the data they require is recorded on the system (1)*
- *to find out what training/ documentation may be needed by other members in order to make use of the system (1)*
- *to ensure that the system can create the relevant outputs that different members require (1)*
- *or any other sensible reasons (1 per reason up to a maximum of three)*

June 2002.7

As an ICT manager in a medium sized company, you have been asked to create a job specification for a database administrator.

a. Describe **three** responsibilities you would include in this specification. (6 marks)

b. The database that this person will be in charge of is a client/server database.
Describe **two** advantages of using this type of database over a non-client/server database. (4 marks)

a.

- *Structure of the database (1) e.g. changes to structure in order to alleviate problems (1)*

- *Keep users informed of changes made to database (1) for example change in field name or field size / introduction or deletion of queries or reports (1)*
- *Maintenance of the data dictionary (1) including such factors as setting conventions for naming of tables, fields etc (1)*
- *Controlling / implementing access rights to the database (1) e.g. so that inexperienced users who need to see data cannot inadvertently delete / change it (1)*
- *Allocating passwords to users (1) so that one person has overall responsibility for who has any access to the database and can track this (1)*
- *Provide training and support to users (1) so that new staff are aware of how to use systems, and all staff able to make efficient use of the system (1)*
- *Backup / restore (1) + expansion (1)*

Credit any other points that are reasonable to include as functions of a DBA. Do not credit any reference to personal features e.g. 'must be trustworthy'

b.

- *expensive resource is made available to a large user base (1) so this is more cost effective (1)*
- *consistency of the data is maintained (1) as only one copy of the data is held on the server, rather than copies held on workstations (1)*
- *processing is done at the server (1) so the client does not need to be so powerful (1)*
- *communication between client and server is minimal (1) only requests and results are communicated, rather than entire databases (1)*
- *Department specific report formats or queries can be held on workstations (1) meaning that less room is taken up on the server/these are less likely to be accessed by the 'wrong' people (1)*
- *Greater control over the data (1) + expansion (1)*

Spring 2003.1

(a) Explain what is meant by a relational database management system. (3 marks)

(b) A successful relational database will have undergone normalisation.

Describe what is meant by normalisation. (2 marks)

a.

A collection of programs/ layer of software (1) that

- *is between the user and the data structure (1)*
- *allows manipulation of data/ use of query functions (1)*
- *allows definition of data dictionary (1)*
- *where the data is stored in separate tables that are related through linked fields (1)*
- *can dynamically generate new tables from old (1)*

b.

Process of breaking down complex data structures into simpler forms. (1) + expansion/example (1)

June 2003.*9*

A charitable organisation needs to co-ordinate all the data that it holds at several locations across the world. They have discovered that there is often conflicting data held in different sites, and time is wasted in reconciling them. A consultant has recommended that they use a relational database management system. The organisation has accepted this advice and is now advertising for the post of database administrator. In the advertisement it states that each applicant should write a supporting letter. You have decided to apply for this post.

Write a letter in support of your application, paying particular attention to:

- the role of a database administrator;
- how a relational database management system can help with this problem;
- the advantages of a client/server solution to this problem.

The Quality of Written Communication will be assessed in your answer. *(20 marks)*