

PAPER CODE NO.
2CS62

EXAMINER : Prof. MJR Shave / Dr. M D Beer
DEPARTMENT : Computer Science Tel. No. 3667 / 3672



THE UNIVERSITY
of LIVERPOOL

JANUARY 1999 EXAMINATIONS

Degree of Bachelor of Arts : Year 3
Degree of Bachelor of Engineering : Year 3
Degree of Bachelor of Science : Year 3

TIME ALLOWED: TWO Hours

INSTRUCTIONS TO CANDIDATES

If you attempt to answer more than the required number of questions (in any section), the marks awarded for the excess questions will be discarded (starting with your lowest mark).

ATTEMPT THREE QUESTIONS.



THE UNIVERSITY
of LIVERPOOL

Q1 a) Draw a B-tree of *order 1* to represent the following key values in ascending order:

4, 5, 10, 12, 19, 25, 32, 33, 40, 44

By placing a ring around appropriate elements in the tree you have drawn, identify

- (i) a *key value*
- (ii) a *partition value*

What is the relationship between a specific partition value and the set of key values?

[8 marks]

b) Two further key values, 16 and 20, must now be stored. Re-draw your tree to show its format after these two additions have been made.

[3 marks]

c) What is the *maximum* number of key values which can be stored in a B-tree of order 1 which has three structural levels (i.e. including the root but not counting the level which contains the key values themselves)?

[4 marks]

d) How does a B-tree of order 2 differ from a B-tree of order 1? What is the answer to (c) for a B-tree of order 2?

[6 marks]

e) Represent the data given above (including the two additional values 16 and 20) as a B-tree of *order 2* with the least possible height.

[4 marks]



THE UNIVERSITY
of LIVERPOOL

- Q2 (a) Distinguish between a *system table*, a *stored table* and a *view* in the Ingres Database Management System. Include in your answer.
- (i) two examples of a *system table*
 - (ii) the SQL command which creates a *stored table*
 - (iii) the manner in which a *view* is implemented
- [9 marks]

- (b) Why is an efficient *query processing strategy* important in the context of a distributed database? Name *three* factors which may influence the choice of a suitable strategy.
- [4 marks]

- (c) Merseyside University operates a distributed database. One of its 'local' sites is the University Registry (site R), which holds details of new students and their accommodation requirements in the relation:

Student (S#, Sname, HomeAddress, Type, MaxRent)

where *MaxRent* is the maximum amount the student is willing to pay, and *Type* has a value such as 'house', 'flat', or 'furnished room'.

Another local site of the database is the University's Property Bureau (site B) which maintains the relation:

Property (Pcode, Plocation, Type, Rent)

The Accommodation Officer of the University works at site A and wishes to know what accommodation, in a suitable price range, is available to students and located in Wavertree. He therefore submits the following query from site A to the distributed database:

```
SELECT      S#, Pcode, Type, Rent
FROM        Student s, Property p
WHERE       s.Type = p.Type
AND         s.MaxRent = < p.Rent
AND         p.Plocation = 'Wavertree'
```

The query can be processed in whole or in part at any of the three sites R, B, A. Determine the most efficient query processing strategy, given the following statistics.

Student : 200 tuples, each of 500 bytes

Property : 1000 tuples, each of 4000 bytes

10% of available properties are in Wavertree

On average, the query shown above produces about 10 results

You should also assume that, at all sites:

Transfer rate = 2000 bits/sec.

Access delay = 0.2 secs.

All computation time is negligible

[12 marks]



THE UNIVERSITY
of LIVERPOOL

Q3 An International company wishes to install an automatic stock distribution system, which will allow stock to be drawn directly from the production control databases at each factory. The local database will then draw on local stock or schedule its manufacture and dispatch.

- a) Describe how the two-phase commit protocol can be used to implement this stock transfer system. Why is it important that messages are passed reliably, with ensured single delivery?

The various involved are distributed around the world. This means that because of time differences and local communications failures all the databases cannot be assumed to be available at all times. How does this effect the use of the two-phase commit protocol?

[10 marks]

- b) It is proposed to provide caches at each site that will queue stock requisitions until they can be passed to the next stage reliably. The following diagram illustrates the proposed architecture:

Can a two-phase commit protocol still be used in this scenario?

Describe in detail a mechanism that will ensure that all messages are passed securely from database to database, without loss or duplication.

[15 marks]

- Q4 a) What is the most significant difference between the Data Protection Act (1998) and the 1984 Act? How will this effect the processing of personal data?

[5 marks]

- c) Describe four of the eight Data Protection Principles as defined in the 1998 Act, showing the issues that need to be addressed in an organisation, such as the University, to comply with each.

[20 marks]