

# File Organization And Access

## Question 1

- 2 The diagram shows four files and three methods of file organisation.

Draw **one** line to match each file with its most appropriate method of file organisation.

File	File organisation
Text file	Sequential
File for recording the temperature every hour	Random
Master file for paying each employee every month	Serial
Customer user name and password file	

[4]

## Question 2

- (b) (i) The weather station records how the outside temperature changes over a period of time. The system will read the temperature once every hour, over a period of 100 days.

The temperature readings are automatically stored in a file. No other data are stored.

Explain why the weather station has decided to use serial organisation for the file.

.....

.....

.....

..... [2]

- (ii) Serial files can be accessed using sequential access.

Explain how sequential access could be used for the temperature readings file.

.....

.....

.....

..... [2]

(iii) Name **and** describe a method of file organisation other than serial or sequential.

Method .....

Description .....

.....  
 .....  
 .....  
 .....

[4]

### Question 3

4 A bank has 95 000 customers. Each customer has a unique ID.

When a customer uses an Automated Teller Machine (ATM) to obtain cash, their current balance is checked. The balance is stored in a file which has the following fields:

- the customer ID (6-digit number in the range 100000 to 999999)
- an encrypted PIN
- the current balance

The file can store a maximum of 100 000 records.

(a) Give a reason why a random organisation would be appropriate for this file.

.....  
 .....[1]

(b) An algorithm for inserting a new record in this file uses the following hash function:

$$\text{RecordKey} \leftarrow \text{CustomerID} \bmod 100000$$

where *RecordKey* is the record position in the file.

(i) Complete the table to show the values generated by the hash function for the given customer IDs.

CustomerID	RecordKey
802139	2139
700004	
689998	
102139	

[1]

(ii) State the range of possible values for *RecordKey*.

Minimum value of *RecordKey*: .....

Maximum value of *RecordKey*: .....

[2]

- (iii) A procedure is written to insert a new record into the file.

Complete the algorithm for this procedure.

```
PROCEDURE InsertRecord(CustomerID : INTEGER)
    RecordKey ← CustomerID MOD 100000
    Success ← FALSE
    // Find position for new record and insert it
    REPEAT
        IF record at position RecordKey is .....
            THEN
                Insert new record at position RecordKey
                Success ← TRUE
            ELSE
                IF RecordKey = .....
                    THEN
                        RecordKey ← .....
                    ELSE
                        RecordKey ← ..... + 1
                ENDIF
            ENDIF
        UNTIL Success = TRUE
    ENDPROCEDURE
```

[4]

- (c) (i) Explain why an encrypted version of the PIN is stored in the file.

.....  
.....  
.....  
.....[2]

- (ii) A customer attempts to withdraw cash from an ATM. An algorithm is used to check if the customer has entered the correct PIN.

Complete the algorithm.

1. Customer ID is read from card.
2. Customer enters PIN.
3. Customer PIN is .....
4. ....
5. Customer record is located in file.
6. ....
7. If match then transaction can proceed.

[3]

## Question 4

- 4 (a) Three file organisation methods and two file access methods are shown below.

Draw lines to link each file organisation method to its appropriate file access method(s).

File organisation method	File access method
random	sequential
serial	direct
sequential	

[4]

- (b) An energy company supplies electricity to a large number of customers. Each customer has a meter that records the amount of electricity used. Customers submit meter readings using their online account.

The company's computer system stores data about its customers.

This data includes:

- account number
- personal data (name, address, telephone number)
- meter readings
- username and encrypted password.

The computer system uses three files:

File	Content	Use
A	Account number and meter readings for the current month.	Each time a customer submits their reading, a new record is added to the file.
B	Customer's personal data.	At the end of the month to create a statement that shows the electricity supplied and the total cost.
C	Username and encrypted passwords.	When customers log in to their accounts to submit meter readings.

For each of the files A, B and C, state an appropriate file organisation method for the use given in the table.

All three file organisation methods must be different.

Justify your choice.

(i) File A organisation .....

Justification .....

.....

.....

.....[3]

(ii) File B organisation .....

Justification .....

.....

.....

.....[3]

(iii) File C organisation .....

Justification .....

.....

.....

.....[3]

## Question 5

4 (a) Three file organisation methods and two file access methods are shown below.

Draw lines to link each file organisation method to its appropriate file access method or methods.

### File organisation method

serial

sequential

random

### File access method

direct

sequential

[4]

**(b)** A bank has a very large number of customers. The bank stores data for each customer. This includes:

- unique customer number
- personal data (name, address, telephone number)
- transactions

The bank computer system makes use of three files:

- A – a file that stores customer personal data. This file is used at the end of each month for the production of the monthly statement.
- B – a file that stores encrypted personal identification numbers (PINs) for customer bank cards. This file is accessed when the customer attempts to withdraw cash at a cash machine (ATM).
- C – a file that stores all customer transaction records for the current month. Every time the customer makes a transaction, a new record is created.

For each of the files A, B and C, state an appropriate method of organisation. Justify your choice.

**(i)** File A organisation .....

Justification .....

.....

.....

.....[3]

**(ii)** File B organisation .....

Justification .....

.....

.....

.....[3]

**(iii)** File C organisation .....

Justification .....

.....

.....

.....[3]

## Question 6

- (ii) The programmer decides to store all the data in a file. Initially, data from 27 locations will be stored. More rainfall locations will be added over time and will never exceed 100.

The programmer has to choose between two types of file organisation. The two types are serial and sequential.

Give **two** reasons for choosing serial file organisation.

.....

.....

.....

.....[2]

## Question 7

- (ii) The programmer decides to store all the data in a file. The number of weather stations could grow to reach 20000, but not all stations will be present at first.

The programmer decides on random organisation for the file.

Describe **three** steps which show how a new weather station record is added to the file.

1 .....

.....

2 .....

.....

3 .....

.....[3]

# Answer

## Answer 1

2	<p><b>One mark for each correct line drawn</b></p>	4
---	--	---

## Answer 2

5(b)(i)	<p>1 mark per bullet point to <b>max 2</b></p> <ul style="list-style-type: none"> <li>• So the readings are stored in <b>chronological</b> order</li> <li>• Easy to add / append each new reading to the end of the file // no further processing is required</li> <li>• Allows the readings to be read in the order that they were taken</li> <li>• Readings do not need to be given further identification as to date / time // no key field needs to be added</li> </ul>	2
5(b)(ii)	<p>1 mark per bullet point <b>max 2</b></p> <ul style="list-style-type: none"> <li>• Earliest temperature reading is accessed first</li> <li>• and each successive temperature reading is read (in date / time order)</li> <li>• until the final reading has been accessed</li> </ul>	2
5(b)(iii)	<p>1 mark for Random</p> <p>1 mark per bullet point for description to <b>max 3</b></p> <ul style="list-style-type: none"> <li>• Record locations are calculated</li> <li>• ... using a hashing algorithm <b>on a key field</b></li> <li>• If a record cannot be stored / found at that location</li> <li>• ... then subsequent locations are searched // closed hash</li> <li>• ... or an overflow area is searched // open hash</li> </ul>	4



### Answer 3

4(a)	Example: Speed of access Just used as a look-up file No need for any serial or sequential processing 1 mark for any valid point	1										
4(b)(i)	<table><tr><td>CustomerID</td><td>RecordKey</td></tr><tr><td>802139</td><td>2139</td></tr><tr><td>700004</td><td>4</td></tr><tr><td>689998</td><td>89998</td></tr><tr><td>102139</td><td>2139</td></tr></table>	CustomerID	RecordKey	802139	2139	700004	4	689998	89998	102139	2139	1
CustomerID	RecordKey											
802139	2139											
700004	4											
689998	89998											
102139	2139											
4(b)(ii)	Minimum value: 0 Maximum value: 99999	1 1 2										
4(b)(iii)	<pre>PROCEDURE InsertRecord(CustomerID : INTEGER)   RecordKey ← CustomerID MOD 100000   Success ← FALSE   // Find position for new record and insert it   REPEAT     IF record at position RecordKey is <u>empty</u>       THEN         Insert new record at position RecordKey         Success ← TRUE       ELSE         IF RecordKey = <u>99999</u>           THEN             RecordKey ← <u>0</u>           ELSE             RecordKey ← <u>RecordKey + 1</u>           ENDIF         ENDIF       UNTIL Success = TRUE   ENDPROCEDURE</pre>	4										
4(c)(i)	For security If file is hacked then encrypted PIN cannot be used Only encrypted PINs are transmitted and compared 1 mark for any valid point	Max 2										
4(c)(ii)	1. Customer ID is read from card 2. Customer enters PIN 3. Customer PIN is <u>encrypted</u> 4. <u>Customer ID is hashed</u> 5. Customer record is located in file 6. <u>PIN is checked against PIN in record</u> 7. If match then transaction can proceed	3										

## Answer 4

4(a)	<p><b>File organisation method</b></p> <pre> graph LR     subgraph "File organisation method"         R[random]         S[serial]         SQ[sequential]     end     subgraph "File access method"         SE[sequential]         D[direct]     end     R --&gt; SE     R --&gt; D     S --&gt; SE     SQ --&gt; SE     SQ --&gt; D           </pre> <p>1 mark for random correct 1 mark for serial correct 2 marks for sequential correct (1 per correct line)</p>	4
4(b)(i)	<p>File A: Serial Meter readings are submitted over time // added to the end of file Stored chronologically</p> <p>1 1 1</p>	3
4(b)(ii)	<p>File B: Sequential Any two points from: Each customer has a unique account number Sorted on Account number High hit rate // Suitable for batch processing monthly statements</p> <p>1 1 1 1</p>	3
4(b)(iii)	<p>File C: Random Login without waiting // Random organisation allows fastest direct access to required record Low hit rate // Suitable for access to individual records</p> <p>1 1 1</p>	3

## Answer 5

4 (a)	<p><b>File organisation method</b></p> <pre> graph LR     subgraph "File organisation method"         S[serial]         SQ[sequential]         R[random]     end     subgraph "File access method"         D[direct]         SE[sequential]     end     S --&gt; D     SQ --&gt; D     SQ --&gt; SE     R --&gt; SE           </pre>	<p>1</p> <p>2</p> <p>1</p>
-------	--	----------------------------

(b) (i)	Sequential	1
	As all customers get statement ... // high hit rate	1
	Suitable for batch processing of the records // the records will be processed one after the other	1
	File organised using customer's unique ID (as primary key field)	1
	//	
	Serial	1
	As all customers get statement ... // high hit rate	1
	Suitable for batch processing of the records // the records will be processed one after the other	1
	Order not important	1
		<b>Max 3</b>
(ii)	Random	1
	Real-time transaction processing	1
	Requires fastest access to data	1
	No need to search through records	1
(iii)	Serial	1
	Each new record is appended	1
	Transactions are recorded in chronological order	1
	File re-organisation not required for each new record // no need for the records to be sorted	
		<b>Max 3</b>

## Answer 6

(ii)	• no need to re-sort data every time new data is added	1
	• only a small file so searching will require little processing	1
	• new records can easily be appended	1
		<b>[max 2]</b>

## Answer 7

(ii)	StationID is hashed to produce home location	1
	If home location is free insert record	1
	Else use overflow method to find free location	1