# **File Organization And Access**

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

# Question 1

	File	File organisation
	Text file	
		Sequential
File	e for recording the temperature every hour	
		Random
Maste	er file for paying each employee every month	
		Serial
	Customer user name and password file	
		[4]
(b) (i)	The weather station records how the outside temp. The system will read the temperature once every. The temperature readings are automatically store.	hour, over a period of 100 days.
(b) (i)	The system will read the temperature once every	hour, over a period of 100 days.  d in a file. No other data are stored.
b) (i)	The system will read the temperature once every  The temperature readings are automatically store  Explain why the weather station has decided to us	hour, over a period of 100 days.  d in a file. No other data are stored.  se serial organisation for the file.
b) (i) (ii)	The system will read the temperature once every  The temperature readings are automatically store  Explain why the weather station has decided to us	hour, over a period of 100 days.  d in a file. No other data are stored.  se serial organisation for the file.
	The system will read the temperature once every The temperature readings are automatically store Explain why the weather station has decided to us	hour, over a period of 100 days.  d in a file. No other data are stored.  se serial organisation for the file.
	The system will read the temperature once every The temperature readings are automatically store Explain why the weather station has decided to use  Serial files can be accessed using sequential acc Explain how sequential access could be used for	hour, over a period of 100 days.  d in a file. No other data are stored.  se serial organisation for the file.
	The system will read the temperature once every The temperature readings are automatically store Explain why the weather station has decided to use  Serial files can be accessed using sequential acc Explain how sequential access could be used for	hour, over a period of 100 days.  d in a file. No other data are stored.  se serial organisation for the file.  ess.  the temperature readings file.

(iii)	Nam	e <b>and</b> describe a method of fil	le orgar	nisation other t	han serial or sequential.	
	Meth	nod				
	Des	cription				
	200	, , , , , , , , , , , , , , , , , , ,				
						[4]
Ques	tion	3				
_						
<b>4</b> A	bank h	as 95 000 customers. Each customers.	omer ha	is a unique ID.		
		customer uses an Automated Te ed. The balance is stored in a file				balance
•	the	customer ID (6-digit number in th	ne range	100000 to 999	999)	
•		ncrypted PIN current balance				
т	he file (	can store a maximum of 100 000	records			
(a	a) Give	e a reason why a random organis	sation w	ould be approp	riate for this file.	
	•••••					
						[1]
	<b>(b)</b> Ar	algorithm for inserting a new re	cord in t	his file uses the	following hash function:	
		RecordKey ←				
		_				
	VVI	nere RecordKey is the record po				
	(i)	Complete the table to show to customer IDs.	the valu	es generated b	by the hash function for t	he given
		Custome	erID	RecordKey		
		80213	39	2139		
		7000	04			
		68999	98			
		10213	39			
						[1]
	(ii)	State the range of possible va	lues for	RecordKey.		
		Minimum value of RecordKey	7 <b>:</b>			
		Maximum value of RecordKe	у			[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 .....
   5. Customer record is located in file.
   6. .....
   7. If match then transaction can proceed.
```

### 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	

**(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
Α	Account number and meter readings for the current month.	Each time a customer submits their reading, a new record is added to the file.
В	Customer's personal data.	At the end of the month to create a statement that shows the electricity supplied and the total cost.
С	Usernames and encrypted passwords.	When customers log in to their accounts to submit meter readings.

[4]

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.

Just	tify your choice.
(i)	File A organisation
	Justification
	[3
(ii)	File B organisation
	Justification
	[3
(iii)	File C organisation
	Justification
	[3]
Ques	tion 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 File access method

File organisation method		File access method
serial		direct
	1	
sequential		sequential
random		

- **(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.

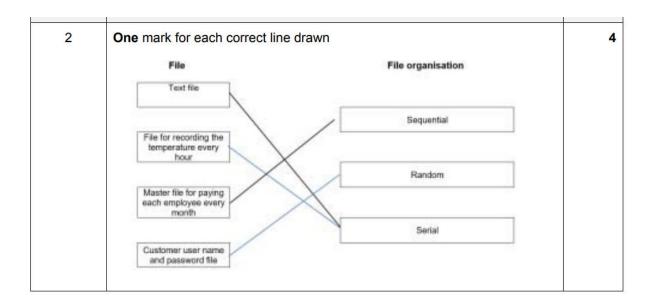
For each of the files A, B and C, state an appropriate method of organisation. Justify your choice.

(i)	File A organisation	
.,	Justification	
		[3]
(ii)	File B organisation	
	Justification	
		[3]
(iii)	File C organisation	
	Justification	
		[3]

# Question 6

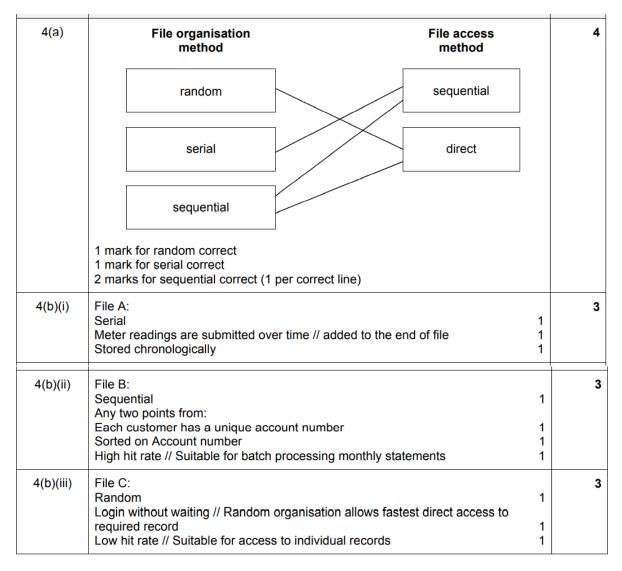
(ii)	The programmer decides to store all the data in a file. Initially, data from 27 locations will be stored. More rainfall locations will be added over time and will never exceed 100.
	The programmer has to choose between two types of file organisation. The two types are serial and sequential.
	Give two reasons for choosing serial file organisation.
	[2]
Questio	n 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 <b>three</b> steps which show how a new weather station record is added to the file.
	1
	2
	3
	[3]

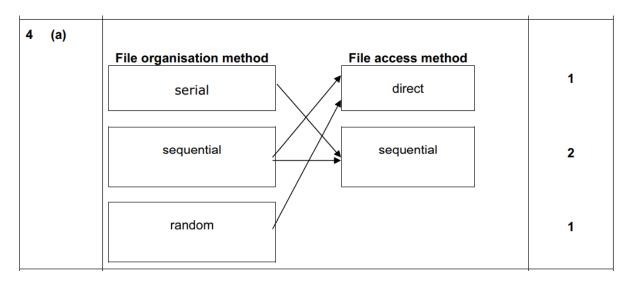
### Answer 1



	+	
5(b)(i)	1 mark per bullet point to max 2	2
	<ul> <li>So the readings are stored in chronological 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>	
5(b)(ii)	1 mark per bullet point max 2	2
	<ul> <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>	
5(b)(iii)	1 mark for Random	4
	1 mark per bullet point for description to max 3	
	Record locations are calculated	
	using a hashing algorithm on a key field	
	If a record cannot be stored / found at that location	
	<ul> <li> then subsequent locations are searched // closed hash</li> <li> or an overflow area is searched // open hash</li> </ul>	

4(a)	Example: Speed of access Just used as a lo No need for any 1 mark for any v	ook-up file serial or seque	ential processing	1
	1			<u> </u>
4(b)(i)	CustomerID	RecordKey		1
	802139	2139		
	700004	4		
	689998	89998		
	102139	2139		
4(b)(ii)	Minimum value: Maximum value		1 1	2
4(b)(iii)	RecordKey Success  // Find p REPEAT IF rec THE  ELS	Custome FALSE cosition for cord at position Insert new Success ← THEN Record ELSE Record ENDIF	ey = <b>99999</b> cdKey ← <b>0</b> cdKey ← <b>RecordKey</b> + 1	4
4(c)(i)		PINs are transi	PIN cannot be used mitted and compared	Max 2
4(c)(ii)	6. PIN is ch	enters PIN PIN is encr ID is hashe record is lecked agair	rypted	3





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

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

(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	