**SECTION A - Multiple-choice questions**

|  |
| --- |
| **Instructions for Section A**Answer all questions in **pencil** on the answer sheet provided for multiple-choice questions. Choose the response that is correct or that best answers the question. A correct answer scores 1, an incorrect answer scores 0. Marks will not be deducted for incorrect answers.No marks will be given if more than one answer is completed for any question. |

**Question 1**

In a use case diagram, an eclipse represents:

1. An actor
2. An entity
3. A use case
4. A data flow

**Question 2**

An array is:

1. A file of records
2. A data structure o store files
3. A file that allows data to be manipulated easily
4. A data structure that allows data to be manipulated easily

**Question 3**

In the Open Systems Interconnections (OSI) model the physical layer represents how:

1. A connection to a communications medium can be started or stopped
2. A connection to the correct network is established
3. A connection to an application is established
4. A connection to the user is established

**Question 4**

For K ← 1 to 50

 NewMember ← Person[K]

EndFor

In the above Pseudocode, Person is an example of:

1. A file
2. An array
3. A record
4. An array of records

**Question 5**

User documentation is prepared in which phase of the problem solving methodology?

1. Analysis
2. Design
3. Development
4. Evaluation

*The following pseudocode applies to questions 6, 7 and 8*

Begin

 Read T

 If T>35 then

 X←’very hot’

 Else

 If T>25 then

 X←’warm’

 Else

 If T>20 then

 X←’perfect’

 Else

 If T<20 then

 X←’cold’

 End

**Question 6**

When converting the above algorithm into a program it would be best for other programmers if the program used

1. Meaningful variable names and internal documentation
2. Meaningful variable names and external documentation
3. Single letter variable names and internal documentation
4. Meaningful variable names and no documentation

**Question 7**

When the above algorithm was tested it was found that it did not provide the correct result when T=20. This was caused by a:

1. Logic error
2. Syntax error
3. Run time error
4. Compile time error

**Question 8**

Test the above algorithm with the value T=25. The variable X will contain:

1. ‘cold’
2. ‘warm’
3. ‘perfect’
4. ‘very hot’

**Question 9**

A small company uses two floors of a building for its office. A 24-port switch is used on each floor to connect all the computers on that floor. Each computer has a network card and is connected to the switch by cable.
The topology used on **each floor** is

1. Bus only
2. Star only
3. Hybrid
4. A combination of bus and star

**Question 10**

Computer systems use protocols to communicate. A protocol that allows computers to communicate with the internet is:

1. HTTP
2. TCP/IP
3. BIOS
4. IP Address

**Question 11**

Assume A=1, B=2 and C=3. The statement that causes D to become true is:

1. If A>B and C>B then D←true
2. If B>A and B>C then D←true
3. If A>B or C>B then D←true
4. If A>B and B>C then D←true

**Question 12**

The two data types that best store the family name of a person, for example Smith, and their phone number, for example (03) 9999 9999, are:

1. Number and Time/Date
2. Number and String/Text
3. String/Text and Time/Date
4. String/Text and String/Text

**Question 13**

A variable is required to store a true/false value. The best data type for this variable is:

1. Number
2. Boolean
3. String/Text
4. Time/Date

**Question 14**



Which statement best represents what is shown in the data flow diagram above?

1. A customer places an order for a spare part and an invoice is sent to them.
2. A process reads data from two data stores and creates an invoice for customers.
3. Two processes send data to a program which then stores the invoice in a file called customer.
4. Two data flows are combined in a data store to produce a new data flow which then goes to the customer.

**Question 15**

A major weakness of a bus topology is:

1. If there is a cable break the whole network is likely to fail
2. If a computer fails the whole network is likely to fail
3. It is more expensive to install than other topologies
4. It is more difficult to install than other topologies

**Question 16**

A random access file is best described as a file where records:

1. Have no fixed length
2. Must always be read in order
3. Are kept sorted by a key field
4. Can be accessed by using their record number

**Question 17**

Project management requires precise scheduling of time throughout the duration of a project. Within this schedule there will be a set of tasks that must be completed on time if the project is going to finish on time. The term a project manager would use for this set of tasks is:

1. Goal
2. Target
3. Milestone
4. Critical path

**Question 18**

Using encryption software when transmitting data means:

1. Data intercepted on route is unreadable
2. Data cannot be intercepted on route
3. Data arrives at its destination faster
4. Data is sent in smaller packets

**Question 19**

A business is installing a new LAN with a goal of improving the overall efficiency of the transfer of information within the business. To achieve this goal there were a number of system objectives, one of which could have been:

1. To eliminate the need to copy data to and from CD
2. To improve the clarity of information presented on screen and paper
3. To minimise errors in data when it was transferred from one computer to another
4. To maximise the amount of time workers could spend speaking to one another about work issues.

**Question 20**

A software company has a policy that 50% of all lines in a program must be internal documentation. The reason for this is:

1. It make the program easier to use
2. It makes the program code more efficient
3. It makes the program output easier to read
4. It makes the program easier to alter by another programmer

**SECTION B - Short answer questions**

|  |
| --- |
| **Instructions for Section B**Answer all questions in the spaces provided. |

**Question 1**

Sue Sands Super Store has an extensive database of all items that it sells. One of the files in the database uses 3.5 GB (gigabytes) of hard disk space. A new procedure requires individual records from this file to be displayed on the screen. Two alternative methods for obtaining the data have been suggested.

|  |
| --- |
| **Method 1**At the start of the program read the whole file into memory and search memory each time a record is required.**Method 2**Obtain the record from the file each time it is required. |

**a.** State one disadvantage of method 1.

It would consume most, if not all of the computer's RAM and severely degrade system performance.

1 mark

**b.** State one disadvantage of method 2.

It would take more time to access each record from secondary storage.

1 mark

**c.**

**i.**  Outline how the file could be organised to make method 2 more efficient.

Store the records stored in a random access file.

**ii.** Explain why this file organisation will be more efficient.

Because the records are fixed-length, it's possible to locate the exact starting point of a record mathematically and jump to that spot instantly and load the record. It avoids the drawback of sequential files in which you have to scroll through all the intervening records to reach the one you need.

1 + 1 = 2 marks

**Question 2**

Threats to computer systems come in many forms.
**a.**  Explain the key difference between the way viruses and trojans get into a computer system.

Viruses enter a system attached to other files (not email attachments, virus code is injected into host files, particularly executable files and travels inside the host.)

Trojans are self-contained malware that are usually downloaded and opened by users who are deceived or tricked into thinking it is something good.

2 marks

b.  Explain the key difference between the purpose of spyware and the purpose of viruses.

Spyware is designed to monitor a system and its users and return information to the spyware's controller. Monitoring can include keystrokes, including passwords and bank account data (via keyloggers), websites visited, work rate etc.

Viruses have a variety of purposes including causing data loss, annoying people, showing off, or just getting their virus to propagate widely. Viruses yield no real or direct benefit to their authors.

2 marks

**Question 3**

A program is needed to process data for stock items. The data is stored in a file on disk. The file could contain up to 500 records. Three typical records are shown in the following table.

|  |  |  |  |
| --- | --- | --- | --- |
| **Item ID** | **Item name** | **Number in stock** | **Cost ($)** |
| 1 | CD RW drive | 2 | 23.50 |
| 2 | 4 Gb RAM | 6 | 54.00 |
| 3 | 17" monitor | 2 | 223.00 |

Juan, a novice programmer, is writing the program. He has chosen a programming language that recognises the following data types: integer, floating-point decimal, text and Boolean. It will also allow the creation of data structures such as arrays and records.
Juan will use a variable named ItemCost to hold the cost of an item. He has decided to make ItemCost an integer.

a. Explain why this is not the most appropriate choice of data type.

Because integer stores only full whole numbers, not decimals, and cost (price) values are in a decimal.

2 marks

b. State the data type he should have chosen.

Floating point.

1 mark

c. Juan knows that for the program to run most efficiently it should read all the relevant data into memory first. He starts by making a large number of variables: ItemCostl, ItemCost2, ItemCost3, and so on. A colleague sees this and tells Juan that an array will be much more efficient. Explain why.

Sample answers:

- Values stored in arrays can be addressed by their index number, which allows the use of loops to traverse or address any number of values rather than by addressing discrete variables by name.

- An array only has to be defined once and can hold a large amount of items in its list. The array can be easily loaded up at the start of the application, more quickly then loading up large amounts of variables, therefore saving on time thus being more efficient.
- Writing a large number of variables is very time consuming and makes the program look too messy (for other coders). By placing the variable ItemCost in an array, he does not need to create a large number of variables because the array is made up of the variables. An array holds one data element so an array could contain ItemCost, this way it is more efficient to run (the program needs only to search the array) and it is less time consuming.

2 marks

**Question 4**

Explain the difference between user documentation and internal documentation for a software application.

*Note this was originally a 4 mark question. I have reduced it to 2 marks as they don’t need to know too much detail about external documentation at this stage*

Sample answers:

- User documentation provides software users with information on how to use the software. Internal documentation (such as comments within a program) is not seen by users; it is meant to tell programmers important things about the code, such as: what it does; why it was written as it was; who wrote it; its version; warnings to future developers etc.

- user documentation explains the use and features of a program as well as providing a guide to fix any common errors. It is external and usually printed in hard copy through online user documentation
- internal documentation refers to the comments inserted into the code by the original programmers. These comments provide guidance or explanations to maintenance programmers who may have to fix bugs in the software
Students needed to contrast both the purpose and the type of user and internal documentation. Many students did not clearly explain that internal documentation occurs in the code or did not mention the purpose of the documentation.

2 marks

**Question 5**
For each of the acceptance criteria listed below, describe a testing procedure that would show if the criterion
was being met.

|  |  |
| --- | --- |
| **Acceptance criteria** | **Testing procedure** |
| **The server must have less than two failures in a month**  | - Count the number of failures in the server's error log. - Test log failures over a year: if less than two per month, the criterion has been met. |
| **The network is able to have 150 computers logged on at the same time without crashing**  |  - Log in at least 150 of the computers and observe if the network crashes. - Undertake a controlled test of 150 computers: log in and if the system crashes, the criterion has not been met. |

2 marks

Only 21 per cent of students gained full marks on this question as many students were unable to provide an appropriate ‘testing procedure’. Some students did indicate a test but then failed to describe what data they were looking for or what information they would gain from the test. For example, for the first criterion, ‘log server failures’ was an insufficient answer as it did not go on to state that this would need to be done over a period of time, that the log would need to be checked to ensure that there were fewer than two failures a month during that period, or that one month of testing would be insufficient to ascertain whether the criterion had been meet.

**SECTION C - Case study**

|  |
| --- |
| **Instructions for Section C**Answer **all** questions in the spaces provided. Remove the case study insert and read **all** the information provided before you answer these questions. Answers must apply to the case study.  |

### Flip Flop Bakery

Flip Flop Bakery is owned by Sebastian, who is a baker. He employs two assistant bakers, three delivery drivers and four part-time sales assistants.

Flip Flop Bakery makes bread and bread rolls that it sells to the public through a shop. It also sells its bread, and a range of cakes and pastries that are bought from a wholesale cake supplier, to local restaurants, coffee lounges and clubs.

A typical working day follows this sequence.

* Sebastian and the assistant bakers arrive at the bakery and start baking at 4.00 am.
* All breads and other bread products (bread rolls and so on) are baked by 7.00 am.
* The bakers identify any ingredients that are in low supply and inform Sebastian.
* The drivers arrive at the bakery and start loading their trucks at 8.00 am.
* Cake deliveries arrive at the bakery around 8.00 am.
* The bakery drivers deliver their goods and invoices to the customers from 9.00 am.
* Drivers handwrite orders from the regular customers and return these to the bakery by 1.00 pm.
* Sebastian manually combines the orders and writes a list of the next day's baking requirements.
* Sebastian telephones his orders for the necessary ingredients and cakes from his suppliers. Orders must be
placed by 3.00 pm to ensure delivery for the next day's baking.
* The bakery closes at 5.00 pm.

#### The problems

Sebastian has identified a number of problems that, because of the growth of his business, have now become critical.

1. When drivers return to the bakery later than 1.00 pm, Sebastian cannot complete the next day's list of baking requirements.
2. Some drivers have such bad handwriting that Sebastian has trouble reading the orders.
3. Sebastian now spends too much time combining the orders and completing the next day's baking requirements list.
4. Sometimes Sebastian cannot order extra ingredients on time therefore he cannot bake all of the next day's bread and so loses orders.

#### Proposed system

To ensure the continued success of his business, Sebastian realises that he must make some changes to his ordering processes. Sebastian's daughter, Kayla, has just completed the first year of a university course in software engineering. She offers to look at how the problems might be solved.
Kayla believes that the only way to improve efficiency is to have a computerised ordering system. There would be a computer at the bakery and the delivery truck drivers would have some kind of mobile device. They would use this to collect and transfer data to the new computer.



Figure 1

Kayla knows that to help Sebastian she must first fully understand the problems, and then analyse the current system.

To start her system analysis, Kayla has drawn the diagram (Figure 1 above). It shows the movement of goods, invoices and orders associated with Flip Flop Bakery's operation, and the role of the various people involved. She now has to create a logical design for the bakery's information system by drawing a context diagram and data flow diagrams, as well as constructing a data dictionary.

**Question 1**

Kayla has started the context diagram.



a. The label for the unnamed entity should be (1 mark) State average=0.5

Customer

*... but where are the drivers? If bakers are external entities, why aren't drivers? Odd...*

b. Figure 1 in the case study shows three arrows going between the Cake supplier and the Bakery, but the context diagram only shows two. Explain this difference. 1 mark, 3 lines State average=0.2

Cakes are not **data** flow and should not appear in DFDs / context diagrams.. They are cakes. Yummy cakes.

*Is there a Cake Flow Diagram?*

c. Kayla has based her context diagram on Figure 1. Explain why she has drawn the Bakers as a separate entity. 1 mark, 3 lines State average=0.36

She's just wrong??

*I have absolutely NO idea. Throw me a bone, someone. If bakers are external entities even though they exist within the data flow of Flip Flop Bakery, why are drivers not* ***also*** *external entities?*

*Unless the drivers belong in the empty box in the question above... in which case why don't* ***customers*** *appear in the DFD?*

*I'm confused.*

A reliable teacher friend says:

*'They provide data to the 'system' rather than process the data so they would be considered entities rather the part of the system.'*

*Ahhh. Everyone - remember that!*

The Bakers provide data to the system rather than processing the data, therefore they would be considered an entity rather than part of the system.
The use of context and data flow diagrams in the written examination has been relatively consistent over the last few years; therefore teachers and students should have been well prepared for this style of question. However, this question required students to explain their understanding rather than to construct elements of a diagram, which caused problems for a number of students.
Students need to practise constructing context and data flow diagrams and also need to have an understanding of the theory behind the decision making for each element represented. The relevant key knowledge point for this question also requires students to be able to construct data dictionaries .

**Question 2**
Kayla has made three attempts to draw a data flow diagram for the processing of customer orders and determining the baking requirements. Consider the accuracy of the DFDs shown below.







a. Identify the **most correct** data flow diagram. 1 mark State average=0.45

DFD **C**



b. Explain the main error is in each of the other two data flow diagrams. 2 marks, 3.75 lines each State average=0.7

DFD **A** is incorrect because it has an input ‘low ingredient’ that is not required to process orders. Also it does not lead to any output.
DFD **B** is incorrect because it shows a data flow that is ingredients (flours, yeast, sugar) rather than data.

Kayla has also created a data dictionary, part of which is shown below.

|  |
| --- |
| **Flip Flop data dictionary - page 3**  |
| **Name** | **Description** |
| Items to be baked | List of bread and bread rolls (items) to be baked for next day's requirementsComposed of: Item name + quantity required |
| Recipe ingredients | Ingredients and quantities needed to bake a particular itemComposed of: Item name + number made + ingredients list |
| Calculate quantities needed | Use recipe ingredients and items baked to work out how much of every ingredient is needed |
| Item name  | Name of item to be baked; for example, multigrain loaf, bread roll |

c. Identify one inappropriate entry in the data dictionary and explain why it is inappropriate. 2 marks, 4 lines State average=0.65

"Calculate quantities needed" is a process or algorithm, not a datum. It belongs in an IPO chart rather than in a data dictionary.

*Impress your friends at parties: say "a datum" at every opportunity.*

The inappropriate entry in the data dictionary was 'Calculate quantities needed' as this is a process, not an element of data.
Question 2 required students to interpret elements of a data flow diagram. Part c. asked students to demonstrate their knowledge of a data dictionary, and tested new content in this study design. It is clear that this area may need to be covered in more depth. Teachers should be incorporating this into their System Analysis and Design concepts along side DFD.

Kayla has completed her analysis of the system. She has decided to start the design of the new system with the design of the software. From her analysis she knows that the system needs to work out how many days of stock are available to the bakery for each ingredient. Where enough stock is not available, the quantity required must be determined for ordering

She designs an algorithm that

* takes in the number of days of stock order that needs to be catered for (Num\_Days)
* takes in the total number of different ingredients that the bakery uses (Num\_Ingredients)
* reads a file to get the total number of days of each ingredients Sebastian currently has in stock
* for each item of stock, checks amount of stock available (Stock\_Amount)
* if insufficient stock for an ingredient calculate the amount required and store this (Qty())

PROCEDURE Calc\_Order(Num\_Days, Num\_Ingredients)

BEGIN

 Open File

 Ingredient\_ID <-- 1

 REPEAT

 Qty(Ingredient\_ID) <-- 0

 Ingredient\_ID <-- Ingredient\_ID + 1

 Until Ingredient\_ID > Num\_Ingredients

 REPEAT

 READ Stock\_Amount

 Ingredient\_ID <-- 1

 REPEAT

 Ingredient\_ID <-- Ingredient\_ID + 1

 IF Stock\_Amount < Num\_Days

 Qty(Ingredient\_ID) = Num\_Days - Stock\_Amount

 END IF

 UNTIL Ingredient\_ID = Num\_Ingredients

 UNTIL END OF FILE

 Close File

END

**Question 3**
a. What is the purpose of the **first** REPEAT - UNTIL loop? (1 mark) State average=0.1

It initialises all of the Qty array's elements to zero. It uses Ingredient\_ID as the loop's index (counter).

The REPEAT – UNTIL loop initialises (sets up) the array to store the quantity of each ingredient.

To test the algorithm Kayla decides to use some simplified test data with only 1 bread product and only 3 ingredients.

b. Describe the error in this algorithm (1 mark) State average=0.1

Ingredient\_ID is set to 1 in the first of the nested repeat loops and immediately in the second part of the nested loop it is increased by 1, so the amount of the first ingredient is never calculated.
Students needed to have a full understanding of each of the elements within the algorithm in order to answer this part correctly.

c. Suggest **one** way the algorithm could be altered to fix this error (1 mark) State average=0.15

Either of:
• change the initial value of Ingredient\_ID from 1 to 0 in the second loop
• move the Ingredient\_ID increment line to the end of the loop and change the terminal conditions from UNTIL Ingredient\_ID = Num\_Ingredients to UNTIL Ingredient\_ID > Num\_Ingredients
Question 3 challenged students. Only a very small number completed all parts successfully, and over 80 per cent scored no marks for all parts. The algorithm in the question was not overly complicated; however, the use of a two dimensional array may have provided an extra element of difficulty that prevented many students from attempting the question.

**Question 4**
Complete the following data table by inserting the correct data type from the options **integer, floating point, string, boolean, one dimensional array**. 3 marks State average=1.1

|  |  |  |
| --- | --- | --- |
| **Variable** | **Use** | **Type** |
| Ingredient\_ID | Loop Control on Ingredient ID | integer  |
| Num\_Ingredients | Total Number of Ingredients  | integer |
| Qty( ) | Quantity in days of an Ingredient  | one dimensional array  |

This question type has occurred in a variety of forms since 2002; however, only a small percentage of students were able to gain full marks.

**Question 5**
Kayla and her uncle Michael, an experienced programmer, were discussing the best file structure to use to store the orders for each day. Sebastian has 300 regular customers who have different orders for each day of the week. As well he can have 50 casual orders on any day. Twenty-five per cent of his regular customers alter their daily order from week to week. Kayla suggests using a simple **Serial** Access file while Michael argues that due to the number of changes each day, a **Random** Access file would be best.
For the bakery's application, explain the advantage of the Random file structure for the organisation of this data. (2 marks, 6 lines) State average=0.5

Data would be faster to access, modify, and manage. It would give functionality similar to a database. It would also be far easier to scale up to large sizes compared to a simple, dumb serial file.

A record in a random file can be accessed instantly by its record number, and a record is read as a single cohesive package with a known structure. All the records between the current read position and target position can simply skipped over.

Also, random file is opened for both read and write access. A serial file must be closed and reopened in the desired mode.

A serial file must be completely read and written as a single unit. One can't (for example) open it, change data item 45 and close it again. You'd have to open it in read mode. Read all the data up to item 44 and simultaneously write it to a second file that is opened in write mode. You'd read the target item (#45), change it, write it to the output file, then go on and read every remaining data item and write it to the output file! That's a serious amount of disk activity to modify a single data element!

Serial files are simply a sequential list of data items. To read an item from the list, one must start reading at the start of the file and read every data item up the one you want. If the data consists of groups of related fields (such as the buyer's name, the goods ordered and the quantity), you then have to read the items individually. If the reading gets 'out of sych' so you start reading data in the middle of a 'record' rather than at its beginning, you start getting rubbish data.

The single-unit records in a random file, on the other hand, are read and written as a single object, and the fields in the records can be referred to easily by their field names. e.g. employee.surname (where **employee** is a record and **surname** is a field in the employee record..

*Yeah - a bit longer than the 6 lines available on the paper, but I added more as an educational freebie.*

Appropriate responses included the following.
• A Serial Access file would be unsuitable because each new order is written to the end of the file, and finding a required order would be difficult as Sebastian would have to go through each in turn until he found the required order. The same would be true for changing customers’ orders. However, Random Access files allow Sebastian to locate customer orders efficiently as the files can be added or accessed in any order; this also allows orders to be changed more efficiently. This would be important since many regular customers change their orders with the bakery. • The entire Serial Access file must be read and written each time there is a change, but Random Access files can be searched and only the part that is changed needs to be rewritten. This is important as a large number of daily changes to orders must be made.
Although similar questions have appeared in past examination papers, many students were unable to provide responses that showed their understanding of Random Access files versus Serial Access files

**Question 6**
Kayla now plans to set up the new system.
She explains that buying the bakery computer and PDAs and setting them up will take about 3 days; writing, testing and debugging the programs for the bakery computer will take 20 days; writing, testing and debugging the programs for the PDAs will take 15 days; and testing the whole system and making sure it works could take 5 days.

a. Complete the Gantt chart below to show how the four tasks could be completed inside 30 days. 4 marks State average=3.05


b. The task 'test system' is dependent on both programming tasks being completed first. Show this on your Gantt chart above.1 mark State average=0.25

**Question 7**
Michael has agreed to help write the programs for the new system. He has advised Kayla that he wants the programs to have good internal documentation: they should contain comment lines and all variables and procedures should have meaningful names. Kayla feels this is a waste of time as she can program much faster with short variable names and no comments.

Explain why Kayla should follow Michael's advice. 2 marks 7 marks State average=1.15

Meaningful variable names make it easier to remember what variables are used for and help avoid using the wrong variable by mistake. Is is especially useful when you return to your code after some time has passed and you must re-learn your unintuitive names all over again. It is also valuable when other people must maintain your code.

Internal documentation reminds the programmer and co-programmers what the code is meant to do and why things were done as they were. Without such inline documentation it is easy for you or a colleague to change some code without realising the implications of what you were doing.