Validation and Verification

Validation and verification are two ways to check that the data entered into a computer is correct. Data entered incorrectly is of little use.

Data verification

Verification is performed to ensure that the data entered exactly matches the original source. Verification means checking the input data with the original data to make sure that there have been no transcription errors (transcription means copying the data). The standard way to do this is to input the data twice to the computer system. The computer then checks the two data values (which should be the same) and, if they are different, the computer knows that one of the inputs is wrong. E.g. entering password twice during sig-up. Verification methods include:

At the time of entry	At the time of transmission
Double entry	Parity check
Screen/visual check	Checksum.
	ARQ
	Echo Check

Validation is an automatic computer check to ensure that the data entered is sensible and reasonable. It does not check the accuracy of data.

For example, a secondary school student is likely to be aged between 11 and 16. The computer can be programmed only to accept numbers between 11 and 16. This is a **range** check.

However, this does not guarantee that the number typed in is correct. For example, a student's age might be 14, but if 11 are entered it will be valid but incorrect.

A validation check is a rule that is built into a database to check that the data entered is:
- Sensible
- Reasonable
- Within acceptable boundaries
- Complete

It does NOT mean that the data is actually correct, that requires verification.

There are a number of different validation rules that can be used in a database:

Type Checks - Field data types provide a basic method of validation. Field data types are assigned to fields during the creation of the database table and data types such as Numeric, Boolean, Date/Time and Image restrict what can be entered. If a user tries to enter text in a date field or alphabetic characters in a numeric field, their entry will be rejected.

Range checks - these are used to limit the range of data a user can enter. The 'day' part of a date must be in the range 1 to 31. An exam grade should be in the range 'A'...'G' or 'U'.

Presence checks - these simply check that an entry has been made in a particular field i.e. a null value (empty field) is not permitted. Usually, not every field in a record needs to be filled in, however there are likely to be some that must have a value and the presence check means that the system will not allow the record to be saved until an entry is made. An application for a passport must have the applicant's surname.

Length Checks - All alphanumeric data has a length. A single character has a length of 1 and a string of text such as "Hello World" has a length of 11 (spaces are counted in text strings). A length check ensures that such data is either an exact length or does not exceed a specified number of characters. Mobile phone numbers are stored as text and should be 11 characters in length.

Lookup - A lookup check takes the value entered and compares it against a list of values in a separate table. It can then return confirmation of the value entered or a second list based on the value. One use of lookups restricts users to pre-defined input using drop-down lists. A user is forced to use a list box to select from a predetermined list of valid values.

Check digits - this type of check is used with numbers. An extra 'check digit' is calculated from the numbers to be entered and added to the end. The numbers can then be checked at any stage by recalculating the check digit from the other numbers and seeing if it matches the one entered. One example where a check digit is used is in the 10 digit ISBN number which uniquely identifies books.

The last number of the ISBN is actually the check digit for the other numbers, for example - the ISBN 0192761501.

Following two Methods are used to calculate check digit

Modulo-11 Method:

(i)   The position of each digit is first considered:

10	9	8	7	6	5	4	3	2	1	← Digit Position
0	2	2	1	4	3	2	5	6	?	← Number

(ii)   Each digit in the number is then multiplied by it's digit position and the totals are added together:
i.e. $(0 \times 10) + (2 \times 9) + (2 \times 8) + (1 \times 7) + (4 \times 6) + (3 \times 5) + (2 \times 4) + (5 \times 3) + (6 \times 2)$
= 0+18+16+7+24+15+8+15+12
= 115 total

(iii)   The total is then divided by 11 (modulo 11) and the remainder, if any, is subtracted from 11. The answer then gives the check digit.
i.e. 115/11 = 10 remainder 5
i.e. 11 – 5 = 6 (check digit)
hence, the final number is: 0-221-43256-6

(iv)   If Check digit = 10 then it will be represented by X (a representation of 10 in ROMAN numbers)

Modulo-10 Method:

Modulo-10 method is used in check digit calculation in ISBN 13, where the 13th digit of the ISBN code is calculated using the following algorithm.

Steps	Example
1. Find sum of digits at odd position	
2. Find sum of digits at even position and multiply result by 3	

978-3-12-732320-?

Sum of digits at odd position												
9+8+1+7+2+2=29												
9	7	8	3	1	2	7	3	2	3	2	0	?
Sum of digits at even position x 3												
3 (7+3+2+3+3+)=54												

3. Add both sums	29+54=83
4. Find Mod10	83 MOD 10=3
5. If remainder=0 then	
Check digit=0	
Else	
Check digit=10-Remainder	Check digit 10-3=7

ISBN 978-0747595823

Summary of validation

Validation type	How it works	Example usage
Range check	Checks that a value falls within the specified range	Number of hours worked must be less than 50 and more than 0
Length check	Checks the data isn't too short or too long. Values must be a specific length.	A password which needs to be six letters long
Limit Check	Similar to Range Check but the rule involves only one limit.	>=0 means reject negative numbers. Date of birth must not be later than a date.
Type Check	Checks that the data entered is of a given data type,	Number of brothers or sisters would be an integer (whole number).
Character Check	Checks that when a string of characters is entered it does not contain any invalid characters or symbols,	A name would not contain characters such as %, and a telephone number would only contain digits or (,), and+.
Format Check	Checks the data is in the right format. Values must conform to a specific pattern, for example, two letters followed by six digits followed by a single letter	A National Insurance number is in the form LL 99 99 99 L where L is any letter and 9 is any number
Presence check	Checks that data has been entered into a field	In most databases a key field cannot be left blank
Check digit	The last one in a code are used to check the other digits are correct	Bar code readers in supermarkets use check digits

length check – e.g. only 30 characters in name field
character check – e.g. name doesn't contain numeric chars
range check – e.g. day of month in date is between 1 and 31
format check – e.g. date in the form xx/yy/zz
check digit – e.g. end digit on bar code to check if it is valid
type check – e.g. integer, real
(presence check = 0)

Test Data

Test data is the data that is used in testing of a software system.

In order to test a software application we need to enter some data for testing most of the features. Any such specifically identified data which is used in tests is known as test data.

There are following three types of test data:

- Normal Data
- Abnormal Data
- Extreme Data
- Boundary Data

1. Normal Data

This is the data a computer system should work on.Testing needs to be done to prove that the solution works correctly. In order to do this a set of test data should be used together with the result(s) that are expected from that data. The type of test data used to do this is called **NORMAL DATA**, this should be used to work through the solution to find the actual result(s) and see if these are the same as the expected result(s).

For example, here is a set of normal test data for an algorithm to record the percentage marks from 10 end-of-term examinations for a student and find their average mark:

Normal test data: 50, 50, 50, 50, 50, 50, 50, 50, 50, 50

Expected result: 50

2. Abnormal/Erroneous Data

This is data that should cause the system to tell the user that there is a problem with data entered into the system. Testing also needs to be done to prove that the solution does not give incorrect results. In order to do this, test data should be used that will be rejected as the values are not suitable. This type of test data is called **ERRONEOUS** or **ABNORMAL TESTDATA**; it should be rejected by the solution.

For example erroneous/abnormal data for an algorithm to record the percentage marks from 10 end-of-term examinations for a student and find their average mark could be:

Erroneous/abnormal data:–12, eleven

Expected results: these values should be rejected

3. Extreme Data

When testing algorithms with numerical values, sometimes only a given range of values should be allowed. For example, percentage marks should only be in the range 0 to 100. The algorithm should be tested with **EXTREME DATA**, which, in this case, are the largest and smallest marks that should be accepted. Extreme data are the largest and smallest values that normal data can take.

Extreme data: 0, 100

Expected results: these values should be accepted

4. Boundary Data

This is used to establish where the largest and smallest values occur. For example, for percentage marks in the range 0 to 100, the algorithm should be tested with the following boundary data; at each boundary two values are required, one value is accepted and the other value is rejected.

Boundary data for 0 is –1, 0

Expected results: –1 is rejected, 0 is accepted

Rogue Values

A value that stops input is called Rogue Value.

A sequence of inputs may continue until a specific value is input. This value is called a **rogue value** and must be a value which would not normally arise.

A rogue value lets the computer know that a sequence of input values has come to an end.

Example: A number of marks are to be input (terminated by a rogue value of -1). How many of them are over 50?

```
Counter ←0
INPUT Marks
REPEAT
        IF Marks >50 THEN Above50 ←Above50 +1
        INPUT Marks
UNTIL Marks=-1
OUTPUT Above50
```

Example: The flowchart below inputs the weight of a number of parcels in kilograms. Parcels weighing more than 25 kilograms are rejected. **A value of –1 stops the input (a rogue value).** The following information is output: the total weight of the parcels accepted and number of parcels rejected.
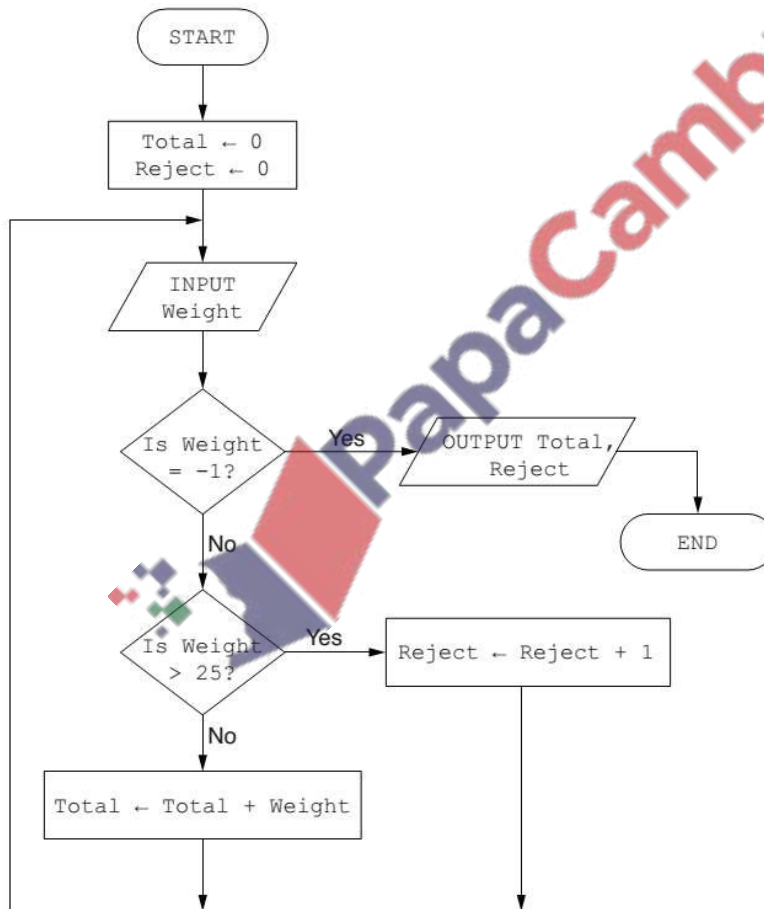
Q 8.1) Activity of data validation and verification:

1) What is an automatic computer check to make sure data entered is sensible and reasonable known as?

a) Double entry   b) Verification   c) Validation

2) What validation type would make sure a post code was entered in the correct format?
   a) Length check   b) Format Check   c) Presence check

3) What validation type would you use to check that numbers fell within a certain range?
   a) Range check   b) Presence Check   c) Check digit

4) What validation type checks that a field is not left blank?
   a) Format check   b) Length check   c) Presence check

5) What validation type uses the last one or two digits to check the other digits are correct?
   a) Length check   b) Format check   c) Check digit

6) What validation type checks a minimum number of characters have been entered?
   a) Length check   b) Format check   c) Range check

7) Data is to be entered into a computer in the format YYMMDD. Which of the following is not a valid date?
   a) 310921   b) 211113   c) 21st June 2004

8) Which of the following statements is false?
a) Validation can check that the data is sensible
b) Validation can check that the data falls between certain allowable boundaries
c) Validation can check that the data is correct

9) Which of the following is NOT a method of verification?
a) Double entry - typing the data in twice and getting the computer to check the second version against the first
b) Using presence, range and length checks to make sure that no mistakes happen
c) Printing out what you have typed in and comparing it against the source data
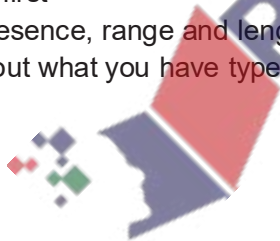
Write down the name of validation check in front of each description that how it works:

Validation type	How it works
	Checks that a value falls within the specified range
	Checks the data isn't too short or too long. Values must be a specific length.
	Similar to Range Check but the rule involves only one limit.
	Checks that the data entered is of a given data type,
	Checks that when a string of characters is entered it does not contain any invalid characters or symbols,
	Checks the data is in the right format. Values must conform to a specific pattern, for example, two letters followed by six digits followed by a single letter
	Checks that data has been entered into a field
	The last one in a code are used to check the other digits are correct

9.11 What is check Digit

....

....

....[2]

Summer 2012 P12

State two different validation checks and give an example of their use. Each example should be different.

Check 1: ....

Use: ....

Check 2: ....

Use: .... [4]

Q 9.5) Describe Test Data and its four types with the help of examples

Test Data

....

....

.... [1]

Test Data Type 1: ....

.... [1]

Test Data Type 2: ....

.... [1]

Test Data Type 3: ....

.... [1]

Test Data Type 4: ....

.... [1]

Summer 2013 P12

A company requests new customers who register online to give the following details:

• Name

• address

• Type of credit/debit card

• Payment card number

All details must be entered.

(a) (i) Describe one suitable different validation check for each field.

Name: ……………………………………………………………...

Address: …………………………………………………………….......................................

type of credit/debit card: ……………………………………………………….....................…..

payment card number: ………………………………………….......................……………… [4]

Summer 2014 pq11

A hospital holds records of its patients in a database. Four of the fields are:
• date of visit (dd/mm/yyyy)   • patient's height (m)
• 8-digit patient ID   • contact telephone number
The presence check is one possible type of validation check on the data. For each field, give another validation check that can be performed. Give an example of data which would fail your named validation check. A different validation check needs to be given for each field.

field name	name of validation check	example of data which would fail the validation check
date of visit		
patient's height		
patient ID		
Contact telephone number		

Marking scheme

Field Name	Name of validation check	Example of data which would fail the validation check
Date of visit	Format/ type/character check	e.g. 2012/12/04 e.g. 3rd March 2012
Patient's height	range check/limit check	can't be < 0 or > 2.5m. e.g. –5, five e.g. 8, -3,
Patient ID	type check / length check / range check	(can't be < 0 or > 99999999) e.g. 3142ABCD e.g. 2131451, 136498207 e.g. –3, 851341625
Contact telephone number	length check, type/character check, format check	e.g. 0773141621834 e.g. 7H215GD e.g. 01223/123456/8901234

Winter 2015 P23

4 A routine checks the age and height of children who are allowed to enter a play area. The children must be less than 5 years of age and under 1 metre in height.

(a) The first set of test data used is age 3 and height 0.82 metres.

State what type of test data this is.

....

Give a reason for using this test data.

....

....[2]

(b) Provide **two** additional sets of test data. For each, give

• the type of each set of test data

• the reason why it is used

Each type of test data and reason for use must be different.

Set 1 ....

Type ....

Reason ....

....

Set 2 ....

Type ....

Reason ....

.... [6]

Winter 2015 P21 & 22

4 A routine checks the weight of melons to be sold in a supermarket. Melons weighing under 0.5 kilograms are rejected and melons weighing over 2 kilograms are also rejected.

Give an example of each type of test data for this routine

Normal ....

Extreme ....

Abnormal ....[3]

Examiners' Comments Question 4
Most candidates could identify at least one correct example of test data. Examples of normal and abnormal test data were usually correct. Some candidates' examples of extreme test data were incorrect. A correct example of extreme test data would be 0.5 or 2.0.

Winter 2016 P21-23

4 Four validation checks and four descriptions are shown below.

Draw a line to link each validation check to the correct description. [3]

Validation check	Description
Presence check	Numbers between two given values are accepted
Range check	Data is of a particular specified type
Type check	Data contains an exact number of characters
Length check	Ensures that some data have entered

Summer 16 P11, 13

9 Check digits are used to ensure the accuracy of entered data. ()

A 7-digit number has an extra digit on the right, called the check digit.

digit position: 1 2 3 4 5 6 7 8

digit: – – – – – – – –

↑

check digit

The check digit is calculated as follows:

•each digit in the number is multiplied by its digit position

•the seven results are then added together

•this total is divided by 11

•the remainder gives the check digit (if the remainder = 10, the check digit is X)

(a) Calculate the check digit for the following number. Show all your working.

4 2 4 1 5 0 8 …

....

....

....

Check digit .... [2]

(b) An operator has just keyed in the following number:

3 2 4 0 0 4 5 X

Circle below correct if the check digit is correct OR incorrect if the check digit is incorrect.

Correct   incorrect

Explain your

answer..

....

....

.... [3]

March 2017 P21 (India)

3 There is a program that stores the following data: [8]

• EmployeeID, an employee ID which must be two letters followed by 4 numbers, e.g. TY4587

• Manager, whether the employee is a manager or not

• AnnualHoliday, number of whole days' annual holiday

• PayGrade, the employee's pay grade which must be a single letter A–F

Complete the following table to identify:

• The most appropriate data type for each variable

• An appropriate validation check for each variable. You must use a different validation check for each variable.

Variable	Data type	Appropriate validation check
EmployeeID		
Manager		
AnnualHoliday		
PayGrade		

3 (a) Explain the difference between a validation check and a verification check.

....

....

....[2]

(b) Describe, using an example, how data could be verified on data entry.

....

....

....[2]

(c) Give two examples of validation of data.

....

....

....[2]

Summer 2017 P22

2 (a) Write an algorithm to input three different numbers, and then output the largest number. Use **either** pseudo code **or** a flowchart.

....

....

....

....

....

....

....

....

....

....

....

....

....[4]
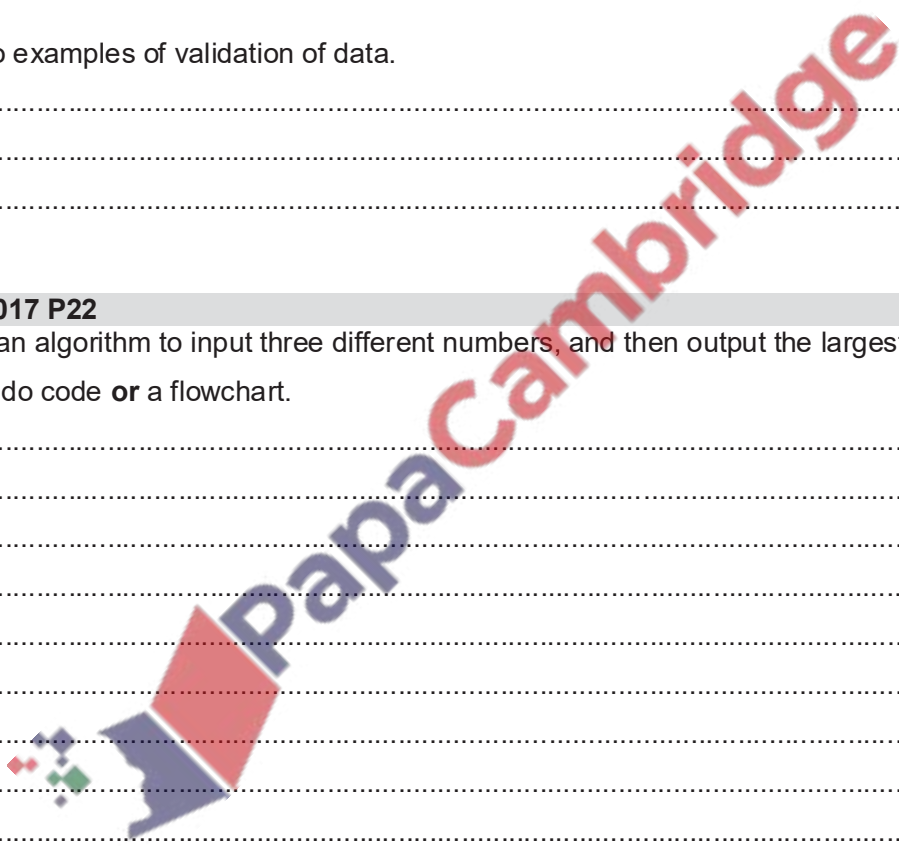
(b) Give **two** sets of test data to use with your algorithm in **part (a)** and explain why you chose each set.

Test data set 1 ....

Reason ....

....

....

Test data set 2 ....

Reason ....

....

....[4]

Summer 17 P21

4 For each of the **four** statements in the table, place a tick in the correct column to show whether it is an example of **validation** or **verification**.   [4]

Statements	Validation	Verification
To automatically check the accuracy of a bar code		
To check if the data input is sensible		
To check if the data input matches the data that has been supplied		
To automatically check that all required data fields have been completed		

March 2018 P22 (India)

A program checks if the weight of a baby is at least 2 kilograms.
Give, with reasons, **two** different values of test data that could be used for the baby's weight.
Each reason must be different.

Value 1 ....

Reason ....

....

Value 2 ....

Reason ....

....[4]

Summer 2018 P21

4 A programmer has written a routine to check that prices are below $10.00. These values are used as test data.

                                    10.00   9.99   ten

Explain why each value was chosen.

10.00 ....

....

....

9.99 ....

....

....

ten ....

....

....[3]

Summer 2018 P22

4 A programmer has written a routine to store the name, email address and password of a contributor to a website's discussion group.

(a) The programmer has chosen to verify the name, email address and password.

Explain why verification was chosen and describe how the programmer would verify this data.

....

....

....

....

....

....

....

....[4]

(b) The programmer has also decided to validate the email address and the password. Describe validation checks that could be used.

Email address...

....

....

Password ....

....

.... [2]

Summer 18 P22

5 A program checks that the weight of a basket of fruit is over 1.00 kilograms and under1.10 kilograms. Weights are recorded to an accuracy of two decimal places and any weight not in this form has already been rejected.

Give **three** weights as test data and for each weight state a reason for choosing it. All your reasons must be different.

Weight 1...

Reason...

....

Weight 2...

Reason...

....

Weight 3...

Reason...

.... [6]
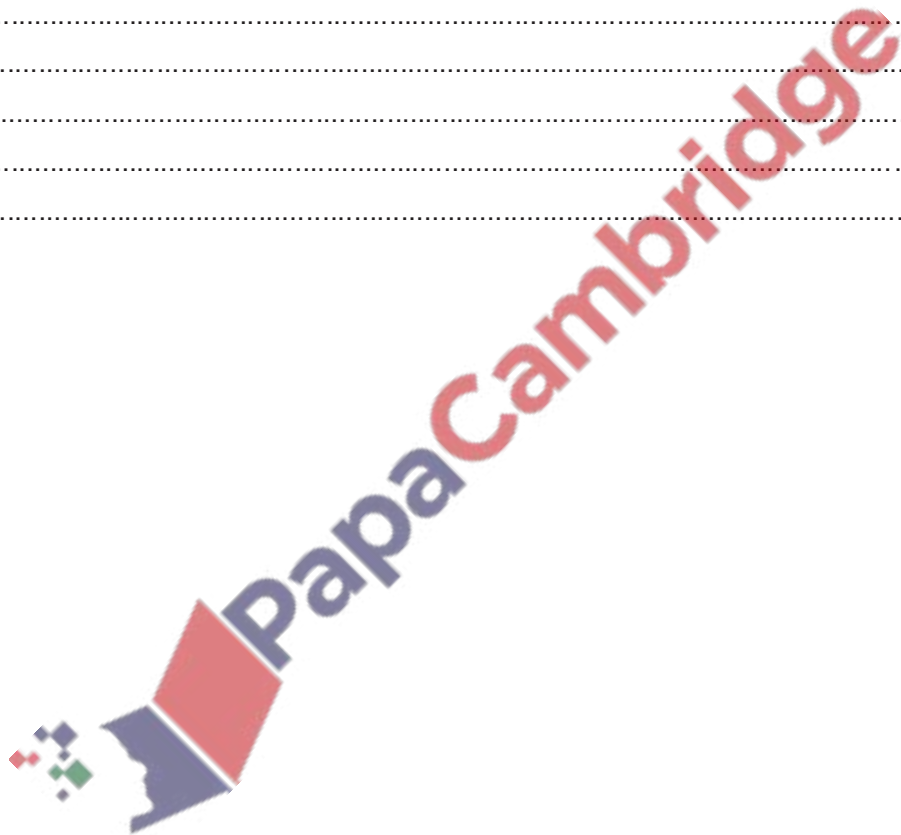
2 (a) Write an algorithm, using pseudo code, to input three different numbers, multiply the two larger numbers together and output the result. Use the variables: Number1, Number2 and Number3 for your numbers and Answer for your result.

....
....
....
....
....
....
....
....
....
....
....
....
....
....
.... [5]

(b) Give **two** sets of test data to use with your algorithm in **part (a)** and explain why you chose each set.

Test data set 1 ....

Reason ....

....

Test data set 2 ....

Reason ....

....[4]
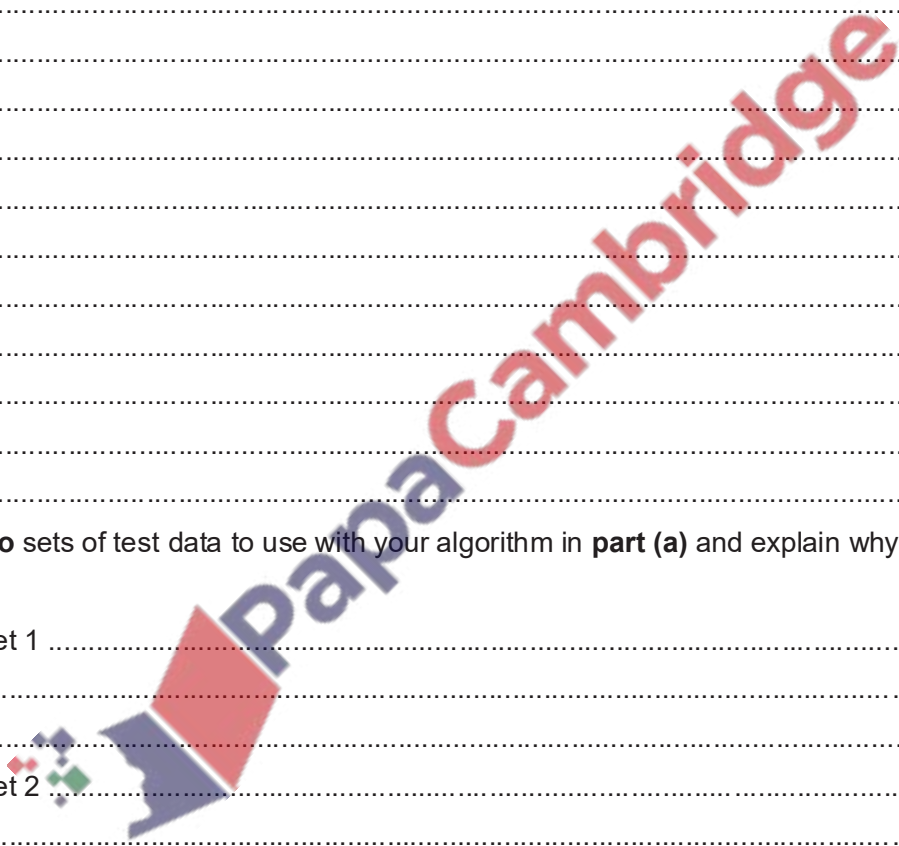
March 2019 P22

5 A programmer restricts input values to less than 90 and greater than 60.

(a) State whether this is called validation or verification.

....

Name the check that needs to be used.

....[2]

(b) State **three** different types of test data the programmer would need to use. Give an example of each type and the reason that the programmer chose that test data.

Type 1 ....

Example ....

Reason ....

....

Type 2 ....

Example ....

Reason ....

....

Type 3 ....

Example ....

Reason ....

....[9]

Q 12.83 Summer2019 P22

5 Explain what is meant by **validation** and **verification**.

Give an example for each one.

Validation ....

....

....

....

Example ....

....

....

Verification ....

....

....

....

Example ....

....

.... [6]

Summer2019 P21

4 This section of program code may be used as a validation check.

1 PRINT "Input a value between 0 and 100 inclusive"

2 INPUT Value

3 WHILE Value < 0 OR Value > 100

4 PRINT "Invalid value, try again"

5 INPUT Value

6 ENDWHILE

7 PRINT "Accepted: ", Value

(a) Give a name for this type of validation check.

.... [1]

(b) Describe what is happening in this validation check.

....

....

....

....

.... [2] (

c) Complete the trace table for this program code using the test data: 200, 300, −1, 50, 60[3]

Value	OUTPUT

(d) Draw a flowchart to represent this section of program code.

Summer 15 P22

3 (a) The flowchart below inputs six single digit numbers. The predefined function MOD gives the value of the remainder, for example, Y ← 10 MOD 3 gives the value Y = 1
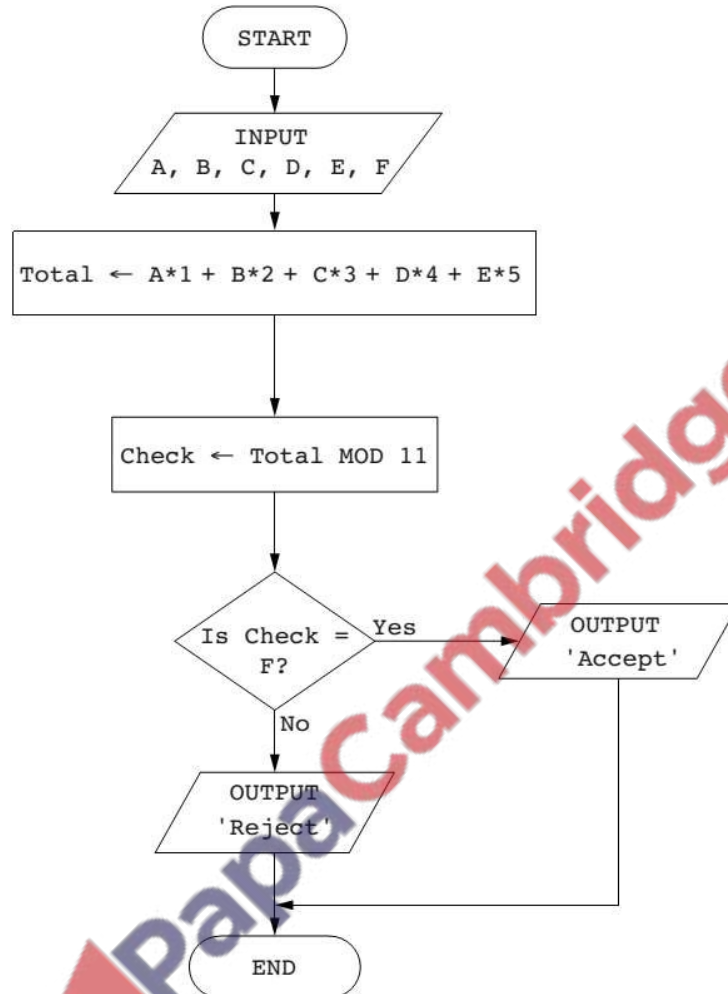
Complete a trace table for each of the two sets of input data.
Set 1 5, 2, 4, 3, 1, 5
Set 2 3, 2, 1, 0, 7, 3

Trace table set 1: 5, 2, 4, 3, 1, 5

A	B	C	D	E	F	Total	Check	Output

Trace table set 2: 3, 2, 1, 0, 7, 3

A	B	C	D	E	F	Total	Check	Output

Candidate Example response

Example candidate response – high

Complete a trace table for each of the two sets of input data.

Set 1   5, 2, 4, 3, 1, 5

Set 2   3, 2, 1, 0, 7, 3

Trace table set 1   5, 2, 4, 3, 1, 5

A	B	C	D	E	F	Total	Check	Output
5	2	4	3	1	5	38	5	Accept

Trace table set 2   3, 2, 1, 0, 7, 3

A	B	C	D	E	F	Total	Check	Output
3	2	1	0	7	3	45	1	Reject

[4]

Examiner comment – high

The candidate has completed both trace tables correctly.

Total mark awarded = 4 out of 4

Example candidate response – middle

Complete a trace table for each of the two sets of input data.

Set 1   5, 2, 4, 3, 1, 5

Set 2   3, 2, 1, 0, 7, 3

Trace table set 1 5, 2, 4, 3, 1, 5

5   1   2

A	B	C	D	E	F	Total	Check	Output
5	2	4	3	1	5	38	5	Accept
5	4	12	12	5	5			

Trace table set 2 3, 2, 1, 0, 7, 3

A	B	C	D	E	F	Total	Check	Output
3	2	1	0	7	3	45	01	Reject
3	4	3	0	35	3			

[4]

Examiner comment – middle

The candidate has completed A, B, C, D, E and F in both trace tables correctly, the working out to check the calculation against F has incorrectly been included so the marks cannot be awarded. Total, Check and Output are correct for one mark in each trace table. 01 has been accepted as 1.

Total mark awarded = 2 out of 4
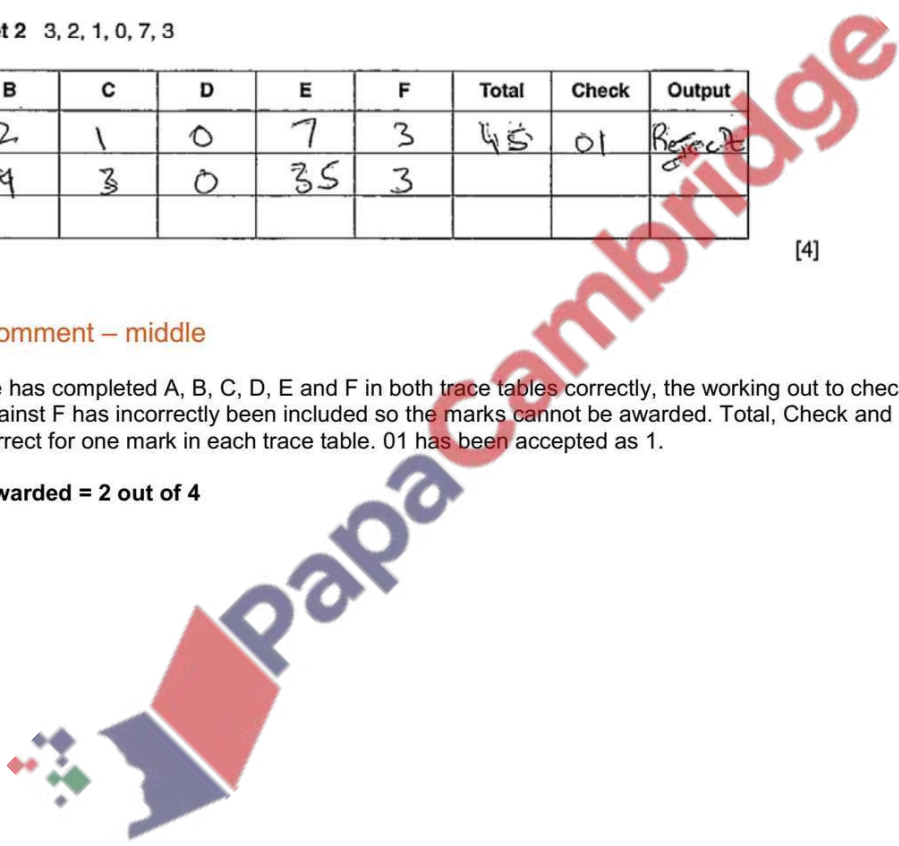
Example candidate response – low

Complete a trace table for each of the two sets of input data.

Set 1   5, 2, 4, 3, 1, 5

Set 2   3, 2, 1, 0, 7, 3

Trace table set 1 5, 2, 4, 3, 1, 5

A	B	C	D	E	F	Total	Check	Output
5	2	4	3	1	5			
5	4	12	12	5	5	38	5	Accept
//	"	//	//	//	//	//	//	Accept

Trace table set 2 3, 2, 1, 0, 7, 3

A	B	C	D	E	F	Total	Check	Output
3	2	1	0	7	3			
3	4	3	0	35	3	45	1	Reject
//	//	//	//	//	//	//	//	Reject

[4]

Examiner comment – low

The candidate has completed A, B, C, D, E and F in both trace tables correctly, the working out to check the calculation against F has incorrectly been included so the marks cannot be awarded. Total and Check are correct, Output is incorrect as both Accept and Reject have been repeated. No marks can be awarded.

Total mark awarded = 0 out of 4