

Question 1

Jeannette Marie Wing is Director of the Data Science Institute at Columbia University, where she is also a professor of computer science.

In a 2010 publication, Wing stated that ‘algorithms are at the heart of computational thinking and computer science’

- (a) What is an algorithm?



An algorithm is a series of steps (3 marks) that is used to solve a problem (2 marks).

(b) In her 2006 essay, entitled *Computational Thinking*, Jeannette Wing stated that ‘computational thinking is using **abstraction** and **decomposition** when attacking a large complex task or designing a large complex system.’

Describe what is meant by the terms **(i)** abstraction **(ii)** decomposition

Abstraction relates to the removing of any unnecessary details in a problem OR identifying the relevant details in a problem. (Either description - 5 marks)

Decomposition is the breaking down of a problem into smaller more manageable problems. (5 marks)

Question 2

State two ways in which **identifying patterns** can help us solve problems in computing.

5 Marks for each relevant answer:

eg.
Save time
More efficient
Makes future problems easier
Easier to make predictions

Question 3

Use **decomposition** to demonstrate how to sort the following strings, first in increasing alphabetical order and then in increasing numerical order:

C5, H3, D8, W4, H6

In your answer explain what you are doing and why you are doing it.

Our first step is we can ignore the numbers (2 marks)

We then sort our data alphabetically (1 mark)

Once the letters are sorted alphabetically we can look for any letters that are listed twice, i.e H, and sort that data numerically (2 marks)

Question 4

Bubble sort, **selection sort** and **insertion sort** are all examples of sorting algorithms you have studied in the previous weeks.

(a) What is the purpose of a sorting algorithm?

To arrange/sort our data alphabetically/numerically or in a desired fashion. (5 marks)

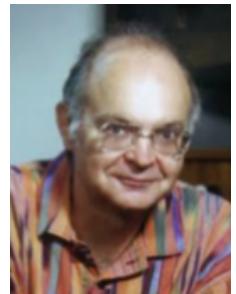
(b) A bubble sort algorithm is applied to the data set below. Which two numbers will swap position first?

87	43	2	35	1	100
----	----	---	----	---	-----

87 and 43 (5 marks)

Question 5

In his book *The Art of Computer Programming*, Donald Knuth states that “searching is the most time-consuming part of many programs, and the substitution of a good search method for a bad one often leads to a substantial increase in speed.”



- (a) The Python code below shows an implementation of a search algorithm.

Examine the code and answer the questions that follow:

```
1 myList=['John', 'Mary', 'Zoe', 'Alex', 'Seamus']
2 name=str(input("Enter name to lookup:"))
3
4 for index in range (len(myList)):
5
6     if myList[index]==name:
7         print("Name is at location: ",index)
8
9
10
```

- (i) State the name of the above search algorithm.

Linear search algorithm (5 marks)

- (ii) What data type is the variable ‘name’?

String (5 marks)

- (iii) **Explain in detail** how the algorithm would find the name ‘Alex’ in the list called myList.

In your answer, identify where the user enters the name to be searched for, how each name in the list is compared to the user input and what the output of this algorithm will be.

User will enter the name ‘Alex’ and it will be assigned to the variable ‘name’ (2 marks)
The for loop will step through the list (2 marks)
The for loop will compare the ‘name’ variable to each element in the list starting at index position[0] (3 marks)
If the two items don’t match then ‘name’ will be compared to the element in the next index position. (3 marks)
This will repeat until a match is found.
The program will output (‘Name is at location 3’) (5 marks)

Most of you lost marks in the last step of this problem, ye just said the output would be ‘3’, you got 2 marks for this answer but not the full 5 if you didn’t state the full sentence.

(iv) What notation is used to describe the worst case time complexity of this search algorithm?

Explain what this means in terms of the number of operations carried out by the algorithm if there is a change in the size of the list.

O(n) or Linear time complexity (3 marks)

If the size of the list increase then so does the number of comparisons that need to be made.
Similarly, if the list size decreases then the number of comparisons decrease (2 marks)

Question 6

Binary search is generally regarded as being a highly efficient search algorithm.

(a) State one **disadvantage** of a binary search algorithm.

Data must be sorted before being searched (5 marks)

(b) Step through how a binary search algorithm finds the integer 28 in the list L shown below. In your answer identify what index positions become the **Min, Max and Mid** during the search.

0	1	2	3	4	5	6	7
3	7	13	18	28	35	41	50
Min			Mid				Max

A lot of you lost marks in part b because you did not do as the question asked and identify values for the Min, Max and Mid values during the search. It didn't matter if you gave index position or numerical value.

The key value is 28 and the starting mid point value is 18/[3] are compared (2 marks)
28 is bigger than 18/[3] so the algorithm ignores all values from 18/[3] downwards. (2 marks)
28/[4] becomes the new min value and 35/[5] is the new midpoint value (3 marks)

28 is compared with 35/[5]

All values from 35/[5] upwards are ignored as they are bigger than the key [1 mark]
Only value left is 28/[4] which matches our key. [2 marks]

Question 7

The output below depicts a list after each pass through a list of integers taken by a certain sorting algorithm:

88, 46, 25, 11, 18, 11 and 22.

Study the data output to the user and answer the questions that follow.

```
Pass 1 : [88, 46, 25, 11, 18, 11, 22]
Pass 2 : [46, 25, 11, 18, 11, 22, 88]
Pass 3 : [25, 11, 18, 11, 22, 46, 88]
Pass 4 : [11, 18, 11, 22, 25, 46, 88]
Pass 5 : [11, 11, 18, 22, 25, 46, 88]
Pass 6 : [11, 11, 18, 22, 25, 46, 88]
Pass 7 : [11, 11, 18, 22, 25, 46, 88]
Sorted array :
[11, 11, 18, 22, 25, 46, 88]
```

- (i) State the name of the algorithm used to sort the list of integers.

Bubble sort, you can identify it because after each step the largest value is sorted. (5 marks)

- (ii) How does this algorithm work?

Index position [0] and [1] are compared and if they are out of order they are swapped [2 marks]
Next, index position [1] and [2] are compared and swapped if necessary [2 marks]
Comparison process is repeated with the hole list [2 marks]
Largest value in the list ends up in its final place at the end of the first run through [2 marks]
Whole list is compared again and second largest ends up in its final place after the second run through [2 marks]
Process repeats until list is sorted

- (iii) Refer to passes 5,6 and 7 and identify one limitation of the algorithm.

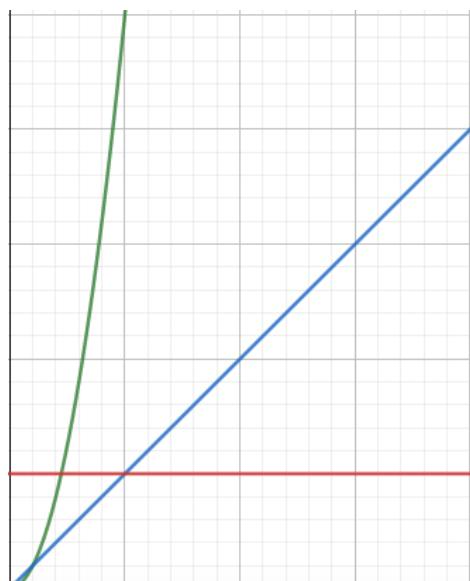
The list is sorted but it keeps running [3 marks]

(iv) How do you think you could overcome this limitation?

Any valid suggestion [2 marks]

Question 8

The graph below shows the plot of three different time complexities. Identify which colour line represents which time complexity and rank them in terms of **fastest, slow, slowest**.



2 Marks for each correct answer

Graph Colour	Time Complexity	Speed
Green	Quadratic $O(n^2)$	Slowest
Blue	Linear / $O(n)$	Slow
Red	Constant / $O(1)$	Fastest

Question 9

```
3 if condition_is_met == 'quadratic':
4     for outerloop in range len(list):
5         for innerloop in range len(list):
6             sum=sum+innerloop
7
8 elif condition_is_met == 'linear':
9     for counter in range len(list):
10        sum=sum+counter
11
12 elif condition_is_met== 'constant':
13     sum=(number*(number+1)/2)|
```

In the code above, depending on a certain condition being met a specific piece of code will be run. The code that is run will result in a different time complexity for the algorithm. Which time complexity would we use to describe this overall algorithm and why?

We would use the quadratic time complexity/ O(n^2) (5 marks)

Even though there are three different time complexities in the code, we use the worst case scenario to describe the complexity of the algorithm. (3 marks)

Question 10

1	Brand	Model	Accel	TopSpeed	Range
2	Tesla	Model 3 Lon	4.6 sec	233 km/h	450 km
3	Volkswagen	ID.3 Pure	10.0 sec	160 km/h	270 km
4	Polestar		2 4.7 sec	210 km/h	400 km
5	BMW	iX3	6.8 sec	180 km/h	360 km
6	Honda	e	9.5 sec	145 km/h	170 km
7	Lucid	Air	2.8 sec	250 km/h	610 km
8	Volkswagen	e-Golf	9.6 sec	150 km/h	190 km
9	Peugeot	e-208	8.1 sec	150 km/h	275 km
10	Tesla	Model 3 Star	5.6 sec	225 km/h	310 km
11	Audi	Q4 e-tron	6.3 sec	180 km/h	400 km
12	Mercedes	EQC 400 4M	5.1 sec	180 km/h	370 km

A student created a program to compare the information of some electric vehicles. The student wanted to find the **mean** and **mode** of the '**TopSpeed**' and '**Range**' of all the vehicles shown in the table above. However, when the student ran the code shown below they get a number of errors.

```
1 import csv
2
3 TopSpeed=[]
4
5 with open('03. Electric Car.csv','r') as file:
6     column_select=csv.DictReader(file)
7
8     for row in column_select:
9         TopSpeed.append(row['TopSpeed'])
10
11 # Calculating the mean
12 Total_TopSpeed=sum(TopSpeed)
13 Mean_TopSpeed=(Total_TopSpeed/len(TopSpeed))
14
15 # Calcualting the mode
16 Mode_TopSpeed=statistics.mode(TopSpeed)
17
18 print(Mean_TopSpeed)
19 print(Mode_TopSpeed)
20
```

Error 1

```
Total_TopSpeed=sum(TopSpeed)
TypeError: unsupported operand type(s) for +: 'int' and 'str'
```

Error 2

```
Mode_TopSpeed=statistics.mode(TopSpeed)
NameError: name 'statistics' is not defined
```

(a) What is the cause of Error 1? (*Hint: There are two things*).

Data is still in string format (3 marks)

Km/hr is still part of the string and needs to be removed. (3 marks)

Most of you identified that the data type was still a string but a good few of you forgot to identify the need to remove KM/hr characters

(b) How would you fix the code so this error no longer occurs?

Convert data to int or float type (3 marks)

Remove the Km/hr characters. (3 marks)

(c) Write the pseudocode or actual code for how you would fix this error.

See below

(d) How would you fix Error 2?

Import statistics module (3 marks)

(e) Write the pseudocode for how you would read in the data contained in the column ‘Range’ and output to the user the mean and mode of the data.

See below

Pseudo code part (c)
Any relevant pseudocode(5 marks)

Create new list for processed data

for item in Topspeed list
replace 'km/h' with ''
add data to new processed data list

for item in processed data list
convert item from string to float
add item to processed data list

Pseudo code part (e)
Any relevant pseudocode(10 marks)

import statistics
import csv

Create new list for range data
Create new list for range data processing

With file open
for row in columns
append row[Range] to range list (2 marks)

for item in Range list
replace 'km' with '' (2 marks)
add data to range data processing list

for item in range data processing lis
convert item from string to float (2 marks)
add item to range data processing list

print(statistics.mean(range data processing lis)) (2 marks)
print(statistics.mode(range data processing lis)) (2 marks)