CSP2348.3 - Data Structures

Assignment 2 - Algorithm implementation and analysis

Student ID: 10636908

Name: Poorav Sharma

Due Date: 05/05/2023

Table of Contents

[Introduction 3](#_Toc134354844)

[Question 1: Array sorting algorithm: design, implementation, and analysis 3](#_Toc134354845)

[Question 1.1: Bubble sort algorithm 3](#_Toc134354846)

[Question 1.2: Improvement to the bubble sort algorithm 4](#_Toc134354847)

[Question 1.2.1: Observation 1 4](#_Toc134354848)

[Question 1.2.2: Observation 2 5](#_Toc134354849)

[Question 1.2.3: Observation 3 7](#_Toc134354850)

[Question 1.3: Variation of the bubble algorithm 9](#_Toc134354851)

[Question 1.3.1: Sink-down sort algorithm 9](#_Toc134354852)

[Question 1.3.2: Bi-directional-bubble sort (BDBS) algorithm 11](#_Toc134354853)

[Question 2: Summary of theoretical analysis of array-based sorting algorithms 14](#_Toc134354854)

[Question 3: Algorithm analysis by experimental study 16](#_Toc134354855)

[Question 3.1: Developing a system to test various cases of sorting algorithms. 16](#_Toc134354856)

[Question 3.1.1: Testing individual sorting algorithm. 16](#_Toc134354857)

[Question 3.2.1: Testing multiple sorting algorithms. 19](#_Toc134354858)

[Question 3.2: Conducting an experimental study. 20](#_Toc134354859)

[Conclusion 22](#_Toc134354860)

[References 22](#_Toc134354861)

# Introduction

This report showcases the design, analysis, implementation, and performances of the array-based sorting algorithms. It answers three questions which requires the application of mathematical theory to algorithm analysis and analyse complexity and performance of the algorithms associated with it. The first question is answered by using algorithm design, algorithm analysis using O-notation, and implementing the algorithm in Python. The second question is based on theoretical algorithm analysis. It contains a table of array sorting algorithms and its complexities. The third question shows an experimental study on several array-based sorting algorithms.

# Question 1: Array sorting algorithm: design, implementation, and analysis

## Question 1.1: Bubble sort algorithm

a) Pseudocode

sorting array in ascending order

n = number of elements in array

FOR repeat until i equals n

FOR repeat until j equals n-1

IF array[j] is bigger than array [j+1]

swap (array[j], [j+1])

END IF

END FOR

END FOR

b) Analysis

The algorithm will compare the element beside each other until the largest number is at the end of the array. The comparison will happen n-1 time in inner loop because that is the maximum number for comparison that can be made in one iteration. The algorithm will repeat iteration n number of times. Therefore, the number of comparisons made in this algorithm will be n\*(n-1). The O-notation is O(n2) because n\*(n-1) = n2-n.

As for the number of copying operation it will depend on the how unsorted the array is. In the best case where the array is already sorted then the number of copying operation will be 0. In the worst-case scenario where the array is arranged from biggest to smallest (descending) order the number of copying operation will equal n\*(n-1)/2. The average at worst case will be half of the equation which is n2-n/4. The O-notation is O(n2) for worst case and average because (n\*(n-1))/2 = (n2-n)/2 and (n\*(n-1))/2 \* ½ = (n2-n)/4.

c) Testing algorithm

**Image 1.1** below shows the array with 20 elements when it is unsorted. Afterwards it shows the array when it has been sorted by the algorithm created. Below it also shows the number of comparisons it made, and the number of Chart, scatter chart

Description automatically generatedcopies (swaps) made.

**Image 1.1** Test result of bubble algorithms.

## Question 1.2: Improvement to the bubble sort algorithm

### Question 1.2.1: Observation 1

a) Pseudocode

sorting array in ascending order

n = number of elements in array

FOR repeat until i equals n

FOR repeat until j equal n-i-1

IF array[j] bigger than array [j+1]

swap (array[j], [j+1])

END IF

END FOR

END FOR

b) Analysis

The algorithm will compare the element beside each other until the largest number is at the end of the array. The comparison will happen n-1-i time in inner loop because the elements that have been sorted doesn’t have to be compared again. The algorithm will repeat iteration n number of times. Therefore, the maximum number of comparisons made in this algorithm will be halved making n\*(n-1) = (n\*(n-1))/2. The O-notation is O(n2) because (n\*(n-1))/2 = (n2-n)/2.

As for the number of copying operation it will depend on the how unsorted the array is. In the best case where the array is already sorted then the number of copying operation will be 0. In the worst-case scenario where the array is arranged from biggest to smallest (descending) order the number of copying operation will equal (n\*(n-1))/2. The average at worst case will be half of the equation which is (n2-n)/4. The O-notation is O(n2) for worst case and average because (n\*(n-1))/2 = (n2-n)/2 and (n\*(n-1))/2 \* ½ = (n2-n)/4.

c) Testing algorithm

Scatter chart

Description automatically generated with low confidence**Image 1.2.1** below shows the array with 20 elements when it is unsorted. Afterwards it shows the array when it has been sorted by the new observation 1 algorithm created. Below it also shows the number of comparisons it made have been halved compared to the original algorithm. It also shows the number of copies (swaps) made.

**Image 1.2.1** Test result of Observation 1-BubbleSort algorithm

### Question 1.2.2: Observation 2

a) Pseudocode

sorting array in ascending order

n = number of elements in array

FOR repeat until i equals n

Set swapped to False.

FOR repeat until j equal n-1

IF array[j] bigger than array [j+1]

swap (array[j], [j+1])

Set swapped to True.

END IF

END FOR

IF swapped is not True

End loop

END IF

END FOR

b) Analysis

The algorithm will compare the element beside each other until the largest number is at the end of the array. The comparison will happen n-1 time in inner loop. The algorithm will repeat iteration n number of times. If the loop was in the worst case (array in descending order), the maximum number of comparisons made in this algorithm will be n\*(n-1). The O-notation for worst case is O(n2) because n\*(n-1) = (n2-n). In the best case where the array is already sorted then the algorithm will only iterate n-1 times in the inner loop and terminate. This is because it no swapped is made the Boolean and if statement placed will active and break away from the loop and stop unwanted comparisons. The O-notation for best case is O(n) because n-1 = n.

As for the number of copying operation it will depend on the how unsorted the array is. In the best case where the array is already sorted then the number of copying operation will be 0. In the worst-case scenario where the array is arranged from biggest to smallest (descending) order the number of copying operation will equal (n\*(n-1))/2. The average at worst case will be half of the equation which is (n2-n)/4. The O-notation is O(n2) for worst case and average because (n\*(n-1))/2 = (n2-n)/2 and (n\*(n-1))/2 \* ½ = (n2-n)/4.

c) Testing algorithm

**Image 1.2.2** below shows the array with 20 elements when it is unsorted. Afterwards it shows the array when it has been sorted by the new observation 2 algorithm created. Below it also shows the number of comparisons and the number of copies (swaps) made. In the second test a sorted array is put into new sorting algorithm to test whether the algorithm terminates as required if there are no elements being swapped.

Graphical user interface, text, application

Description automatically generated

**Image 1.2.2** Test result of Observation 2-BubbleSort algorithm

### Question 1.2.3: Observation 3

a) Pseudocode

sorting array in ascending order

n = number of elements in array

FOR repeat until i equals n

Set swapped to False.

FOR repeat until j equal n-1-i

IF array[j] bigger than array [j+1]

swap (array[j], [j+1])

Set swapped to True.

END IF

END FOR

IF swapped is not True

End loop

END IF

END FOR

b) Analysis

The algorithm will compare the element beside each other until the largest number is at the end of the array. The comparison will happen n-1-i times in inner loop because the elements that have been sorted doesn’t have to be compared again. The algorithm will repeat iteration n number of times. If the loop was in the worst case (array in descending order), the maximum number of comparisons made in this algorithm will be n\*(n-1)/2. The O-notation for worst case is O(n2) because (n\*(n-1))/2 = (n2-n)/2. In the best case where the array is already sorted then the algorithm will only iterate n-1 times in the inner loop and terminate. This is because it no swapped is made the Boolean and if statement placed will active and break away from the loop and stop unwanted comparisons. The O-notation for best case is O(n) because n-1 = n.

As for the number of copying operation it will depend on the how unsorted the array is. In the best case where the array is already sorted, the number of copying operation will be 0. In the worst-case scenario where the array is arranged from biggest to smallest (descending) order the number of copying operation will equal (n\*(n-1))/2. The average worst case will be half of the equation which is (n2-n)/4. The O-notation is O(n2) for worst case and average because (n\*(n-1))/2 = (n2-n)/2 and (n\*(n-1))/2 \* ½ = (n2-n)/4.

c) Testing algorithm

**Image 1.2.3** below shows the array with 20 elements when it is unsorted. Afterwards it shows the array when it has been sorted by the new observation 3 algorithm created which utilises new components used in observation 1 and 2. The test also shows the number of comparisons and the number of copies (swaps) made. In the second test a sorted array is put into new sorting algorithm to test whether the algorithm terminates as required if there are no elements being swapped. The last test shows the worst case (array with element in descending order). It shows the maximum number of comparisons and copies(swaps) the algorithm needed to sort the array. ((n\*(n-1)/2) = (20\*(20-19))/2 = (20\*19)/2 = 380/2 = 190.

Text

Description automatically generated

**Image 1.2.3** Test result of Observation 3-BubbleSort algorithm

## Question 1.3: Variation of the bubble algorithm

### Question 1.3.1: Sink-down sort algorithm

a) Pseudocode

sorting array in ascending order

n = number of elements in array

FOR repeat until i equals n

Iterate i until it equals n

Set swapped to False.

FOR repeat until j equal i

Set j as n-1 loop until equals j equals i

IF array[j] bigger than array [j-1]

swap (array[j], [j-1])

Set swapped to True.

END IF

END FOR

IF swapped is not True

End loop

END IF

END FOR

b) Analysis

This new variation of bubble sort is called sink sort does the exact same thing as a bubble sort except for the order in which the array is sorted in. Unlike bubble sort which sorts the array in ascending order, sink sort algorithm sorts the arrays in descending order. The sink sort applies strategies used in observation 3. Therefore, when the array is already sorted the number of comparisons made is n-1. The O-notion for comparisons for best case is O(n) because n-1 = n. In the worst case (array is in ascending order) the maximum comparisons are (n\*(n-1))/2. The O-notation for worst case is O(n2) because (n\*(n-1))/2 = (n2-n)/2 .

The number of copies made depends on how unsorted the array is. In the best case where the array is already sorted, the number of copying operation will be 0. In the worst-case scenario where the array is arranged from smallest to biggest (ascending) order the number of copying operation will equal (n\*(n-1))/2. The average worst case will be half of the equation which is (n2-n)/4. The O-notation is O(n2) for worst case and average because (n\*(n-1))/2 = (n2-n)/2 and (n\*(n-1))/2 \* ½ = (n2-n)/4.

c) Testing algorithm

**Image 1.3.1** below shows the array with 20 elements when it is unsorted. Afterwards it shows the array when it has been sorted by using the sink sort algorithm which utilises strategies used in observation 1 and 2. The test also shows the number of comparisons and the number of copies (swaps) made. In the second test a sorted array is put into sink sort algorithm to test whether the algorithm terminates as required if there are no elements being swapped. The last test shows the worst case (array with element in ascending order). It shows the maximum number of comparisons and copies(swaps) the algorithm needed to sort the array. ((n\*(n-1)/2) = (20\*(20-19))/2 = (20\*19)/2 = 380/2 = 190.

Graphical user interface, text, application

Description automatically generated

**Image 1.3.1** Test result of Sink Sort algorithm

### Question 1.3.2: Bi-directional-bubble sort (BDBS) algorithm

a) Pseudocode

sorting array in ascending order

n = number of elements in array

FOR repeat until i equals n

Iterate i until it equals n

Set swapped to False.

Set j as i

FOR repeat until j equal n-1-i

IF array[j] bigger than array [j+1]

swap (array[j], [j+1])

Set swapped to True.

END IF

END FOR

IF swapped is not True

End loop

END IF

Set swapped to False.

FOR repeat until k equal i

Set k as n-1 loop until equals k equals i

IF array[k] bigger than array [k-1]

swap (array[k], [k-1])

Set swapped to True.

END IF

END FOR

IF swapped is not True

End loop

END IF

END FOR

b) Analysis

This new variation of bubble sort algorithm is called Bi-directional-bubble sort (BDBS) algorithm. This sort does the exact same thing as a bubble sort except but this time it also utilises sink sort algorithm in its inner loop. This means in the inner look of this algorithm there are two for loop. One for loop scans the array from left to right and places the largest element on right (end). The second loop scans the array from right to left and places the smallest element on the left (beginning) of the array. The number of comparisons made is the same as sink sort and observation 3 bubble sort. This is because every time an element is sorted in the beginning or end using the two loops they are removed from the scan in the future loop. It uses the Boolean swapped to end the loop if there were no swaps in either of the inner loops. When the array is already sorted (best case - array is in ascending order) the number of comparisons made is n-1. The O-notion for comparisons for best case is O(n) because n-1 = n. In the worst case (array is in ascending order) the maximum comparisons are (n\*(n-1))/2. The O-notation for worst case is O(n2) because (n\*(n-1))/2 = (n2-n)/2 .

The number of copies made depends on how unsorted the array is. In the best case where the array is already sorted, the number of copying operation will be 0. In the worst-case scenario where the array is arranged from biggest to smallest (descending) order the number of copying operation will equal (n\*(n-1))/2. The average worst case will be half of the equation which is (n2-n)/4. The O-notation is O(n2) for worst case and average because (n\*(n-1))/2 = (n2-n)/2 and (n\*(n-1))/2 \* ½ = (n2-n)/4.

c) Testing algorithm

**Image 1.3.2** below shows the array with 20 elements when it is unsorted. Afterwards it shows the array when it has been sorted by using the sink sort algorithm which utilises strategies used in observation 1 and 2. The test also shows the number of comparisons and the number of copies (swaps) made. In the second test a sorted array is put into sink sort algorithm to test whether the algorithm terminates as required if there are no elements being swapped. The last test shows the worst case (array with element in ascending order). It shows the maximum number of comparisons the algorithm needed to sort the array. ((n\*(n-1)/2) = (20\*(20-19))/2 = (20\*19)/2 = 380/2 = 190. The maximum number of copies is one less because the array has a element that repeats twice.

Graphical user interface, text, application

Description automatically generated

**Image 1.3.2** Test result of Bi-directional-bubble sort (BDBS) algorithm

# Question 2: Summary of theoretical analysis of array-based sorting algorithms

**Note** thatin **Table 1** only quick sort has the same average complexity as its best case. As for all the other sorting algorithm with worst-cases its worst- case time complexity is the same as their average O-notation.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sorting algorithm** | **Number of comparisons** | **Number of copies** | **Time complexity** | **Space complexity** |
| Selection | ~ n2/2 | ~ 2n | O(n2) | O(1) |
| Insertion | ~ n2/4 | ~ n2/4 | O(n2) | O(1) |
| Merge | ~ n log2n | ~ 2n log2n | O(n log2n) | O(n) |
| Quick | \*BC: ~ n log2n  \*WC: ~ n2/2 | \*BC: ~ 2n/3 log2n  \*WC: 0 | \*BC: O(n log2n)  \*WC: O(n2) | \*BC: O(log2n)  \*WC: O(n) |
| Heap | ~2\*(n log2n) | ~ n log2n | O(n log2n) | O(1) |
| Bubble | n(n-1) | BC: 0  WC: n(n-1)/2 | O(n2) | O(1) |
| Obs1-Bubble | n(n-1)/2 | BC:0  WC: n(n-1)/2 | O(n2) | O(1) |
| Obs2-Bubble | BC: n-1  WC:n(n-1) | BC:0  WC: n(n-1)/2 | BC: O(n)  WC: O(n2) | O(1) |
| Obs3-Bubble | BC: n-1  WC: n(n-1)/2 | BC:0  WC: n(n-1)/2 | BC: O(n)  WC: O(n2) | O(1) |
| Sink-down | BC: n-1  WC: n(n-1)/2 | BC:0  WC: n(n-1)/2 | BC: O(n)  WC: O(n2) | O(1) |
| Bi-directional | BC: n-1  WC: n(n-1)/2 | BC:0  WC: n(n-1)/2 | BC: O(n)  WC: O(n2) | O(1) |

Notes: \*: BC – Best Case; WC- Worst Case

**Table 1**: Array sorting algorithm complexity (by Theoretical analysis)

# Question 3: Algorithm analysis by experimental study

## Question 3.1: Developing a system to test various cases of sorting algorithms.

### Question 3.1.1: Testing individual sorting algorithm.

The images below are the test result for each sorting algorithm.

Selection Sort

Graphical user interface, text, application

Description automatically generated **Image 3.1.1** selection sort test

Iteration Sort

**Image 3.1.2** Iteration sort test Graphical user interface, text

Description automatically generated with medium confidence

Merge Sort

A picture containing text

Description automatically generated

**Image 3.1.3** Merge sort test

Quick Sort

Graphical user interface, text, application

Description automatically generated

**Image 3.1.4** Quick sort test

Heap Sort

Graphical user interface, text

Description automatically generated with medium confidence

**Image 3.1.5** Heap sort test

Bubble Sort

Graphical user interface

Description automatically generated with low confidence

**Image 3.1.6** Bubble sort test

Obs1bubble Sort

Graphical user interface, text

Description automatically generated with medium confidence

**Image 3.1.7** Obs1bubble sort test

Obs2bubble Sort

Graphical user interface, text, application, email

Description automatically generated

**Image 3.1.8** Obs2bubble sort test

Obs3bubble Sort

A picture containing graphical user interface

Description automatically generated

**Image 3.1.9** Obs3bubble sort test

Sink-down Sort

Text

Description automatically generated with low confidence

**Image 3.1.10** Sink-down sort test

Bi-directional Sort

A picture containing text

Description automatically generated

**Image 3.1.11** Bi-directional sort test

### Question 3.2.1: Testing multiple sorting algorithms.

Table

Description automatically generated

**Image 3.2.1** Tests all sorting algorithms using the same array and displays it in table format.

## Question 3.2: Conducting an experimental study.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ***Sorting Algorithm*** | *n=100* | *n=200* | *n=400* | *n=800* | *n=1000* | *n=2000* |
| Selection | 4950 | 19900 | 79800 | 319600 | 499500 | 1999000 |
| Insertion | 2556 | 10236 | 39773 | 160925 | 252104 | 1003571 |
| Merge | 2717 | 6403 | 14818 | 33606 | 43551 | 97058 |
| Quick | 639 | 1551 | 3718 | 8487 | 11285 | 24829 |
| Heap | 1532 | 3665.1 | 8494 | 19289 | 25003 | 56006 |
| Bubble | 9900 | 3651 | 159600 | 639200 | 999000 | 3998000 |
| Obs1-Bubble | 4950 | 19900 | 79800 | 319600 | 499500 | 1999000 |
| Obs2-Bubble | 9098 | 37053 | 149784 | 617147 | 960838 | 3892652 |
| Obs3-Bubble | 4906 | 19721 | 79438 | 319002 | 498580 | 1997444 |
| Sink-down | 4883 | 19823 | 79438 | 318739 | 498410 | 1997403 |
| Bi-directional | 3857 | 15625 | 60602 | 245614 | 381180 | 1515287 |

**Table 2**: Experimental study: Average number of comparisons for sorting arrays of n integers

(Over 10 runs).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sorting Algorithm** | n=100 | n=200 | n=400 | n=800 | n=1000 | n=2000 |
| Selection | 0.200 | 0.900 | 3.201 | 13.702 | 21.005 | 82.319 |
| Insertion | 0.546 | 0.600 | 2.300 | 10.202 | 16.904 | 67.615 |
| Merge | 0.100 | 0.200 | 0.800 | 1.200 | 1.100 | 3.001 |
| Quick | 0.000 | 0.200 | 0.300 | 0.700 | 1.000 | 2.701 |
| Heap | 0.200 | 0.200 | 0.801 | 1.400 | 2.000 | 4.201 |
| Bubble | 0.700 | 2.500 | 10.227 | 44.310 | 69.216 | 282.910 |
| Obs1-Bubble | 0.300 | 1.901 | 6.377 | 26.706 | 41.928 | 170.438 |
| Obs2-Bubble | 0.700 | 2.401 | 10.202 | 43.710 | 68.297 | 281.463 |
| Obs3-Bubble | 0.300 | 1.600 | 6.201 | 27.406 | 43.010 | 174.525 |
| Sink-down | 0.500 | 1.700 | 7.002 | 28.806 | 45.310 | 182.275 |
| Bi-directional | 0.200 | 1.500 | 5.601 | 24.206 | 37.508 | 150.484 |

**Table 3**: Experimental study: Average running time (in ms) for sorting arrays of n integers

(Over 10 runs).

# Conclusion

In this assignment I created the bubble sort algorithm in question. I further improved it according to the observation and created variant bubble sort such as sink down sort and bi-directional bubble sort. In the second question I analysed various sorting algorithms to fill in the table. I used my analysis from question one to fill in the bubble sort rows. For the third question I built a program that allows users to test all the sorting algorithms listed individually or as a group, so that I can compare them. By testing all the sorting algorithms and finding their average comparisons and runtime, I noticed that quick sort is the fastest. In the test with 100 elements in the array the quick sort algorithms was able to sort the array the fastest(program couldn’t measure time so its time was 0 ms). It also had the fast time and least amount of comparison through out the tests. The normal bubble sort and the second observation have the longest runtime and comparisons because they don’t have strategies that to reduce unnecessary comparisons. The tests done in question 3 further proves that the analysis done on the question 1 and 2 is correct because it is showing results that I was expecting when analysing the algorithms.

# References

Programiz. (n.d.). Merge sort algorithm. Retrieved May 7, 2023, from <https://www.programiz.com/dsa/merge-sort>

Hussein, B. (2016, June 20). Introduction to data structures: 9. Merge sort [Video]. YouTube. <https://www.youtube.com/watch?v=RvqmZLuCQrw>

Code Studio. (n.d.). Www.codingninjas.com. <https://www.codingninjas.com/codestudio/library/time-and-space-complexities-of-sorting-algorithms-explained>

Dr Jitian XIAO, CSP2348\_M3\_Array Searching Algorithms and Analysis, Edit Cowan University, Data Structures - CSP2348.3

Dr Jitian XIAO, CSP2348\_M7\_Heaps and heap sorting, Edit Cowan University, Data Structures - CSP2348.3