

# CS-5115 - Programming Prep for Grad Student

## PA1 - Data Structures

Arun Totad

WIN: 353146427

Date: 09/24/2024

- I give permission to the instructor to share my solution(s) with the class.
- The main goals of the homework / programming assignment: Create python program to demonstrate how real difference between 1-d Array and Array-List along with obtaining understanding to analyse complexity of the program.

PA1\_cs5115\_Totad\_092424.zip contains the following files:

- PA1\_Solution.py – Python code solution for the problem
- Array\_Output1.txt & Array\_Output2.txt – Gives the output of the program once executed.
- List\_Output1.txt & List\_Output2.txt – Gives the output of the program once executed.
- Plagiarism declaration.pdf - Plagiarism declaration
- Solution description.pdf – Solution description for the theoretical part of the assignment
- words\_alpha.txt – Local dataset for testing purposes

8) Perform a theoretical complexity analysis of your design (i.e., count number of operations/instructions and space usage) and then express that using asymptotic notation as a function of the input size (Pause: what is the input size of your problem?)

Answer:

The input size of the problem will be equal to size of the dataset elements i.e. in this case,  $n = 370104$

### **As per the code implementation**

Time Complexity:

- Loading the dataset:  $O(n)$
- Incremental Growth (Increase Strategy A):  $O(1)$
- Doubling Growth (Increase Strategy B):  $O(10 * n)$
- Fibonacci Growth (Increase Strategy C):  $O(n)$
- Binary Search: Binary search on a sorted list is performed in  $O(\log n)$ . But since insertion dominates:  $O(n)$

Space Complexity

- Loading the dataset:  $O(n)$
- Incremental Growth (Increase Strategy A):  $O(n)$

- Doubling Growth (Increase Strategy B):  $O(n)$
- Fibonacci Growth (Increase Strategy C):  $O(n)$
- Binary Search: Binary search on a sorted list is performed in  $O(\log n)$ . But since insertion dominates:  $O(n)$

### Theoretical Complexity Analysis

- **Time Complexity:** The overall time complexity of the program can be summarized as  $O(n)$  due to the dominant cost of growing and inserting elements into the array.
- **Space Complexity:** The space complexity is  $O(n)$ , as the entire dataset is stored and manipulated in memory.

### As per the code implementation

Time Complexity:

- Loading the dataset:  $O(n)$
- Incremental Growth (Increase Strategy A):  $O(n)$
- Doubling Growth (Increase Strategy B):  $O(n)$
- Fibonacci Growth (Increase Strategy C):  $O(n)$
- Binary Search: Binary search on a sorted list is performed in  $O(\log n)$ . But since insertion dominates:  $O(n)$

Space Complexity

- Loading the dataset:  $O(n)$
- Incremental Growth (Increase Strategy A):  $O(n)$
- Doubling Growth (Increase Strategy B):  $O(n)$
- Fibonacci Growth (Increase Strategy C):  $O(n)$
- Binary Search: Binary search on a sorted list is performed in  $O(\log n)$ . But since insertion dominates:  $O(n)$

### Theoretical Complexity Analysis

- **Time Complexity:** The overall time complexity of the program can be summarized as  $O(n)$  due to the dominant cost of growing and inserting elements into the array.
- **Space Complexity:** The space complexity is  $O(n)$ , as the entire dataset is stored and manipulated in memory.

10. Empirically measure the time and space complexity of your code

Answer:

## 1. List Implementation:

- Final size of the list: 370104 elements.
- Time for insertion (incrementing by 10 each time): ~0.09377 seconds.

## 2. Dynamic Array Implementation:

- Final size of the dynamic array: 360102 elements
- Time for insertion (doubling the array size): ~0.12489 seconds.

The image shows a Python IDE with two files open: `PA1_List_Implementation.py` and `PA1_Solution.py`. The left pane displays the source code, and the right pane shows the Jupyter Notebook console output.

**PA1\_List\_Implementation.py** (lines 7-36):

```
7 import urllib.request
8 import time
9
10 'Reading dataset for pre-processing'
11 def read_dataset(dataset = "", url = ""):
12     if(dataset != ""):
13         with open(dataset, "r") as file: #reading the file from locally
14             contentlist = file.readlines()
15             contentlist = [x.strip() for x in contentlist]
16     elif(url != ""): #reading the file from global
17         response = urllib.request.urlopen(url)
18         contentlist = response.read().decode().splitlines()
19     else:
20         print("Dataset not defined for processing")
21
22     print("-----Dataset fetched successfully-----")
23     print("Dataset --> ", dataset, url)
24     print()
25     print("Total size of Dataset: ", len(contentlist))
26     print("Elements of list = ", contentlist[0], " -> ", contentlist[len(c
27     #print("Any None values in List: ", any(x is None for x in List))
28     print("*****")
29     print()
30     return contentlist
31
32 'Display results after execution of strategies'
33 def print_results(List, timeElapsed):
34     print("Current size of list: ", len(List)) #Current size of the
35     print("Time elapsed: ", timeElapsed, " seconds")
36
```

**PA1\_Solution.py** (lines 224-253):

```
224 else:
225     print("Item already exist in the array")
226     end_time = time.time()
227     elapsed_time4 = end_time - start_time
228
229 'Main function'
230 if __name__ == "__main__":
231     print("*****")
232     print("CS5115: Programming for Graduate Student : PA 1 Solution")
233     print("*****")
234
235     contentlist = read_dataset("C:/Users/Arun Totad/Desktop/CS5115/Assignm
236     arrayType = input("Select type of dataset for processing: \n A) Array
237     if(arrayType.upper() == "A"):
238         arrayType = "Array"
239     elif arrayType.upper() == "B":
240         arrayType = "List"
241     else:
242         print("Mentioned data structure doesn't work for this program, Tr
243
244     ewol = incremental_strategies("IncrementBy2", contentlist, arrayType)
245     ewol = incremental_strategies("StrategyA", contentlist, arrayType)
246     ewol = incremental_strategies("StrategyB", contentlist, arrayType)
247     ewol = incremental_strategies("StrategyC", contentlist, arrayType)
248     print("Lets insert an element into the List")
249     ewol.insert(ewol)
250
251 #Graphical interface for executing each functionality individually.
252 ...
253
```

**Console Output:**

**In [818]:** contentlist = read\_dataset("C:/Users/Arun Totad/Desktop/CS5115/Assignment assets/PA1/PA1\_cs5115\_Totad\_092424/EOWL\_200 shuffled.txt")  
-----Dataset fetched successfully-----  
Dataset --> C:/Users/Arun Totad/Desktop/CS5115/Assignment assets/PA1/PA1\_cs5115\_Totad\_092424/EOWL\_200 shuffled.txt  
Total size of Dataset: 200  
Elements of list = together -> girl -> hour -> town -> multiply  
\*\*\*\*\*

**In [819]:** contentlist = read\_dataset("C:/Users/Arun Totad/Desktop/CS5115/Assignment assets/PA1/PA1\_cs5115\_Totad\_092424/EOWL\_200 sorted.txt")  
-----Dataset fetched successfully-----  
Dataset --> C:/Users/Arun Totad/Desktop/CS5115/Assignment assets/PA1/PA1\_cs5115\_Totad\_092424/EOWL\_200 sorted.txt  
Total size of Dataset: 200  
Elements of list = able -> fall -> mind -> simple -> young  
\*\*\*\*\*

**In [820]:** arrayType = input("Select type of dataset for processing: \n A) Array Implementation \n B) List Implementation \n")  
Array Implementation  
...: if(arrayType.upper() == "A"): ...: arrayType = "Array"  
...: elif arrayType.upper() == "B": ...: arrayType = "List"  
...: else: ...: print("Mentioned data structure doesn't work for this program, Try again with A/B")  
Select type of dataset for processing:  
A) Array Implementation  
B) List Implementation  
b  
**In [821]:**

```
PA1_Array_Implementation.py X PA1_List_Implementation.py X PA1_Solution.py X
231 print("Item already exist in the array")
232 end_time = time.time()
233 elapsed_time = end_time - start_time
234
235
236 'Main function'
237 if __name__ == "__main__":
238     print("*****")
239     print("CSS115: Programming for Graduate Student : PA 1 Solution")
240     print("*****")
241     print()
242
243 contentList = read_dataset("C:/Users/Arun Totad/Desktop/CSS115/Assignment assets/PA1/")
244 arrayType = input("Select type of dataset for processing: \n A) Array Implementation
245 if(arrayType.upper() == "A"):
246     arrayType = "Array"
247     print("-----Demonstrating Array Data Structure-----")
248     print()
249 elif arrayType.upper() == "B":
250     arrayType = "List"
251     print("-----Demonstrating List Data Structure-----")
252     print()
253 else:
254     print("Mentioned data structure doesn't work for this program, Try again with A/B")
255 ewol = incremental_strategies("IncrementBy2", contentList, arrayType)
256 ewol = incremental_strategies("StrategyA", contentList, arrayType)
257 ewol = incremental_strategies("StrategyB", contentList, arrayType)
258 ewol = incremental_strategies("StrategyC", contentList, arrayType)
259 print("Lets insert an element into the list")
260 ewol.insert(ewol)
261
262 #Graphical interface for executing each functionality individually.
263 ...
264 while(True):
265     executeall = input("Do you want to execute all strategies: Y/N ")
266     if (executeall.lower() == "Y"):
267
268 PA1_Array_Implementation.py X PA1_List_Implementation.py X PA1_Solution.py X
16 contentList = [x.strip() for x in contentList]
17 elif(url != ""): #reading the file from global
18     response = urllib.request.urlopen(url)
19     contentList = response.read().decode().splitlines()
20 else:
21     print("Dataset not defined for processing")
22
23 print("-----Dataset fetched successfully-----")
24 print("Dataset --> ",dataset, url)
25 print()
26 print("Total size of Dataset: ", len(contentList))
27 print("Elements of list = ", contentList[0], " -> ", contentList[len(contentList)//4]
28 #print("Any None values in list: ",any(x is None for x in List))
29 print("*****")
30 print()
31 return contentList
32
33 'Display results after execution of strategies'
34 def print_results(list, timeElapsed):
35     print("Current size of list: ", len(list)) #Current size of the
36     print("Time elapsed: ", timeElapsed, " seconds")
37     List.sort()
38     print("Elements of list = ", List[0], " -> ", List[len(List)//4], " -> ", List[len(List)-1])
39     #print("Any None values in list: ",any(x is None for x in List))
40     print("*****")
41
42
43 'Incremental Strategies'
44 def incremental_strategies(incrementor, dataset, arrayType):
45     print()
46     if(arrayType == "List"):
47         ewol = list()
48     elif(arrayType == "Array"):
49         ewol = [None] * len(dataset)
50
51
52 PA1_Array_Implementation.py X PA1_List_Implementation.py X PA1_Solution.py X
16 contentList = [x.strip() for x in contentList]
17 elif(url != ""): #reading the file from global
18     response = urllib.request.urlopen(url)
19     contentList = response.read().decode().splitlines()
20 else:
21     print("Dataset not defined for processing")
22
23 print("-----Dataset fetched successfully-----")
24 print("Dataset --> ",dataset, url)
25 print()
26 print("Total size of Dataset: ", len(contentList))
27 print("Elements of list = ", contentList[0], " -> ", contentList[len(contentList)//4]
28 #print("Any None values in list: ",any(x is None for x in List))
29 print("*****")
30 print()
31 return contentList
32
33 'Display results after execution of strategies'
34 def print_results(list, timeElapsed):
35     print("Current size of list: ", len(list)) #Current size of the
36     print("Time elapsed: ", timeElapsed, " seconds")
37     List.sort()
38     print("Elements of list = ", List[0], " -> ", List[len(List)//4], " -> ", List[len(List)-1])
39     #print("Any None values in list: ",any(x is None for x in List))
40     print("*****")
41
42
43 'Incremental Strategies'
44 def incremental_strategies(incrementor, dataset, arrayType):
45     print()
46     if(arrayType == "List"):
47         ewol = list()
48     elif(arrayType == "Array"):
49         ewol = [None] * len(dataset)
50
51
52 PA1_Array_Implementation.py X PA1_List_Implementation.py X PA1_Solution.py X
Name Type Size Value
arrayType str 5 Array
Help Variable Explorer
Console Z/A X
In [29]: runfile('C:/Users/Arun Totad/Desktop/PA1_Solution.py', wdir='C:/Users/Arun Totad/Desktop')
*****
CSS115: Programming for Graduate Student : PA 1 Solution
*****
-----Dataset fetched successfully-----
Dataset --> C:/Users/Arun Totad/Desktop/CSS115/Assignment assets/PA1/PA1_css115_Totad_092424/EDWL_200 shuffled.txt
Total size of Dataset: 200
Elements of list = together -> girl -> hour -> town -> multiply
*****
Select type of dataset for processing:
A) Array Implementation
B) List Implementation
A
-----Demonstrating Array Data Structure-----
Increment by 2
*****
Current size of list: 200
Time elapsed: 0.0 seconds
Elements of list = able -> fall -> mind -> simple -> young
*****
Strategy A: Increment by 10
*****
Current size of list: 200
Time elapsed: 0.0 seconds
Elements of list = able -> fall -> mind -> simple -> young
*****
Strategy B: Increment by Doubling
*****
Current size of list: 200
Time elapsed: 0.0 seconds
Elements of list = able -> fall -> mind -> simple -> young
*****
Current size of list: 200
Time elapsed: 0.0 seconds
Elements of list = able -> fall -> mind -> simple -> young
*****
Strategy C: Increment by fibonacci series
*****
Current size of list: 200
Time elapsed: 0.0 seconds
Elements of list = able -> fall -> mind -> simple -> young
*****
Lets insert an element into the list
Enter the new word that needs to be inserted into the array: Arun
Index number found for the new element insertion: 6
ewol[] before insertion: area -> base -> beauty
Current size of ewol[] before insertion: 200
new_ewol[] after insertion: area -> arun -> base
Current size of new_ewol[] after insertion: 201
Time elapsed: 0.0 seconds
*****
In [28]:
conda: base (Python 3.12.4) LSP: Python Line 41, Col 1 UTF-8 CRLF RW Mem 65%
```



```
PA1_Array_Implementation.py X PA1_List_Implementation.py X PA1_Solution.py X
16 contentlist = [x.strip() for x in contentlist]
17 elif(url != ""): #reading the file from global
18 response = urllib.request.urlopen(url)
19 contentlist = response.read().decode().splitlines()
20 else:
21 print("Dataset not defined for processing")
22
23 print("-----Dataset fetched successfully-----")
24 print("Dataset --> ",dataset, url)
25 print()
26 print("Total size of Dataset: ", len(contentlist))
27 print("Elements of list = ", contentlist[0], "-> ", contentlist[len(contentlist)//4])
28 #print("Any None values in list: ",any(x is None for x in list))
29 print("Time elapsed: ", timeElapsed, " seconds")
30 print()
31 return contentlist
32
33 'Display results after execution of strategies'
34 def print_results(list, timeElapsed):
35 print("Current size of list: ", len(list)) #Current size of the
36 print("Time elapsed: ", timeElapsed, " seconds")
37 list.sort()
38 print("Elements of list = ", list[0], "-> ", list[len(list)//4], "-> ", list[len(list)-1])
39 #print("Any None values in list: ",any(x is None for x in list))
40 print("*****")
41
42
43
44 'Incremental Strategies'
45 def incremental_strategies(incrementor, dataset, arrayType):
46 print()
47 if(arrayType == "list"):
48 ewol = list()
49
50 elif(arrayType == "Array"):
51 ewol = [None] * len(dataset)
```

Name	Type	Size	Value
arrayType	str	4	List

```
Console 2/A X
In [31]: runfile('C:/Users/Arun Totad/Desktop/PA1_Solution.py', wdir='C:/Users/Arun Totad/Desktop')
*****
CS5115: Programming for Graduate Student : PA 1 Solution
*****
-----Dataset fetched successfully-----
Dataset --> C:/Users/Arun Totad/Desktop/CS5115/Assignment assets/PA1/PA1_cs5115_Totad_092424/EOWL_200 shuffled.txt

Total size of Dataset: 200
Elements of list = together -> girl -> hour -> town -> multiply

Select type of dataset for processing:
A) Array Implementation
B) List Implementation
B
-----Demonstrating List Data Structure-----

Increment by 2
*****
Current size of list: 200
Time elapsed: 0.0 seconds
Elements of list = able -> fall -> mind -> simple -> young
*****
```

```
PA1_Array_Implementation.py X PA1_List_Implementation.py X PA1_Solution.py X
16 contentlist = [x.strip() for x in contentlist]
17 elif(url != ""): #reading the file from global
18 response = urllib.request.urlopen(url)
19 contentlist = response.read().decode().splitlines()
20 else:
21 print("Dataset not defined for processing")
22
23 print("-----Dataset fetched successfully-----")
24 print("Dataset --> ",dataset, url)
25 print()
26 print("Total size of Dataset: ", len(contentlist))
27 print("Elements of list = ", contentlist[0], "-> ", contentlist[len(contentlist)//4])
28 #print("Any None values in list: ",any(x is None for x in list))
29 print("Time elapsed: ", timeElapsed, " seconds")
30 print()
31 return contentlist
32
33 'Display results after execution of strategies'
34 def print_results(list, timeElapsed):
35 print("Current size of list: ", len(list)) #Current size of the
36 print("Time elapsed: ", timeElapsed, " seconds")
37 list.sort()
38 print("Elements of list = ", list[0], "-> ", list[len(list)//4], "-> ", list[len(list)-1])
39 #print("Any None values in list: ",any(x is None for x in list))
40 print("*****")
41
42
43
44 'Incremental Strategies'
45 def incremental_strategies(incrementor, dataset, arrayType):
46 print()
47 if(arrayType == "list"):
48 ewol = list()
49
50 elif(arrayType == "Array"):
51 ewol = [None] * len(dataset)
```

Name	Type	Size	Value
arrayType	str	4	List

```
Console 2/A X
Current size of list: 200
Time elapsed: 0.0 seconds
Elements of list = able -> fall -> mind -> simple -> young
*****

Strategy A: Increment by 10
*****
Current size of list: 200
Time elapsed: 0.0 seconds
Elements of list = able -> fall -> mind -> simple -> young
*****

Strategy B: Increment by Doubling
*****
Current size of list: 200
Time elapsed: 0.0 seconds
Elements of list = able -> fall -> mind -> simple -> young
*****

Strategy C: Increment by fibonacci series
*****
Current size of list: 200
Time elapsed: 0.0 seconds
Elements of list = able -> fall -> mind -> simple -> young
*****
Lets insert an element into the list
Enter the new word that needs to be inserted into the array: axes1
Index number found for the new element insertion: 1
*****
```

```
PA1_Array_Implementation.py X PA1_List_Implementation.py X PA1_Solution.py X
16 contentlist = [x.strip() for x in contentlist]
17 elif(url != ""): #reading the file from global
18 response = urllib.request.urlopen(url)
19 contentlist = response.read().decode().splitlines()
20 else:
21 print("Dataset not defined for processing")
22
23 print("-----Dataset fetched successfully-----")
24 print("Dataset --> ",dataset, url)
25 print()
26 print("Total size of Dataset: ", len(contentlist))
27 print("Elements of list = ", contentlist[0], "-> ", contentlist[len(contentlist)//4])
28 #print("Any None values in list: ",any(x is None for x in list))
29 print("Time elapsed: ", timeElapsed, " seconds")
30 print()
31 return contentlist
32
33 'Display results after execution of strategies'
34 def print_results(list, timeElapsed):
35 print("Current size of list: ", len(list)) #Current size of the
36 print("Time elapsed: ", timeElapsed, " seconds")
37 list.sort()
38 print("Elements of list = ", list[0], "-> ", list[len(list)//4], "-> ", list[len(list)-1])
39 #print("Any None values in list: ",any(x is None for x in list))
40 print("*****")
41
42
43
44 'Incremental Strategies'
45 def incremental_strategies(incrementor, dataset, arrayType):
46 print()
47 if(arrayType == "list"):
48 ewol = list()
49
50 elif(arrayType == "Array"):
51 ewol = [None] * len(dataset)
```

Name	Type	Size	Value
arrayType	str	4	List

```
Console 2/A X
Current size of list: 200
Time elapsed: 0.0 seconds
Elements of list = able -> fall -> mind -> simple -> young
*****
Lets insert an element into the list
Enter the new word that needs to be inserted into the array: axes1
Index number found for the new element insertion: 1

ewol[] before insertion: able -> above -> against
Current size of ewol[] before insertion: 200

new_ewol[] after insertion: able -> axes1 -> above
Current size of new_ewol[] after insertion: 201

Time elapsed: 0.0 seconds
*****

In [32]: |
```