Nasir Hussain

43913

BS-Cy (5-1)

Assignment: 3

Instructor: Muhammad Usman Sharif

Analysis of Algorithms

Implement four sorting algorithms (Bubble Sort, Selection Sort, Merge Sort, and Quick Sort) on three different arrays of equal size.

- Array 1: at which best case scenario applies.
- Array 2: at which average case scenario applies.
- Array 3: at which worst case scenario applies.

For each sorting algorithm and array combination, measure the time taken to sort The array. You can use a built-in function or implement a custom function to calculate The execution time.

Analyze the results to compare the performance of the different sorting algorithms under various conditions.

Solution:

Bubble sort:

Code:

```
import copy
import time

def bubble_sort(arr):

    n = len(arr)
    comparisons = 0
    swaps = 0
    for i in range(n):
        swapped = False
        print(f"\nPass {i+1}:")
        for j in range(0, n - i - 1):
            comparisons += 1
            print(f" Comparing {arr[j]} and {arr[j + 1]}")
            if arr[j] > arr[j + 1]:
                print(f" Swapping {arr[j]} and {arr[j + 1]}")
                arr[j], arr[j + 1] = arr[j + 1], arr[j]
                swaps += 1
                swapped = True
            else:
                print(f" No swap needed for {arr[j]} and {arr[j + 1]}")
    if not swapped:
               print(" No swaps occurred in this pass. The array is already
```

```
return {"comparisons": comparisons, "swaps": swaps}
   best case sorted = copy.deepcopy(best_case)
    average case sorted = copy.deepcopy(average case)
    worst case sorted = copy.deepcopy(worst case)
   start time = time.time()
   display results ("Best", best case sorted, best time, best stats)
   start time = time.time()
    average stats = bubble sort(average case sorted)
   average_time = end_time - start_time
   display results ("Average", average case sorted, average time,
average stats)
   end time = time.time()
    display results ("Worst", worst case sorted, worst time, worst stats)
```

Running:

```
=== Best Case ===
Sorted Array: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
Time Taken: 0.0000000 seconds
Total Comparisons: 9
Total Swaps: 0
```

```
=== Average Case ===
Sorted Array: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
Time Taken: 0.001000 seconds
Total Comparisons: 30
Total Swaps: 10
```

```
=== Worst Case ===

Sorted Array: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

Time Taken: 0.003775 seconds

Total Comparisons: 45

Total Swaps: 45
```

TIME COMPEXITY:

Visual Representation of Time Complexity

- 1. Best Case (Already Sorted):
 - o **Passes Needed:** 1
 - o **Comparisons:** n-1
 - \circ **Swaps:** 0
 - o **Time Complexity:** O(n)
- 2. Average Case (Random Order):
 - o **Passes Needed:** Approximately n/2
 - \circ Comparisons: $(n^2)/2$
 - o **Swaps:** Depends on initial order
 - o **Time Complexity:** O(n²)
- 3. Worst Case (Reverse Sorted):
 - o **Passes Needed:** n-1
 - \circ Comparisons: $(n^2 n)/2$
 - o **Swaps:** $(n^2 n)/2$

o **Time Complexity:** O(n²)

Practical Implications

• Small Datasets:

o Bubble Sort can be acceptable for small arrays due to its simplicity.

• Large Datasets:

- o Inefficient due to O (n²) time complexity.
- o Better sorting algorithms (like Merge Sort, Quick Sort, or Heap Sort) should be used for scalability.

• Nearly Sorted Data:

o Optimized Bubble Sort performs well (O (n)) because it can terminate early if no swaps are needed.

Conclusion:

- **Best Case:** O (n) Efficient for already sorted arrays due to early termination.
- **Average Case:** O (n²) Quadratic time complexity makes it inefficient for larger, randomly ordered arrays.
- Worst Case: O (n²) highly inefficient for reverse-sorted arrays with maximum swaps and comparisons.

Summary of Time Complexities

Component	Operation	Time Complexity
Imports	Importing modules	O(1)
Function Definitions	Defining bubble_sort, display_results, main	O(1) each
bubble_sort Function	Entire Bubble Sort Algorithm	O(n²)
display_results	Displaying results	O(n)
main Function	Sorting Best, Average, Worst Cases	O(n ²)
Execution Check	Running main	O(1)

SELECTION SORT:

Code:

```
import time
def selection sort(arr):
   n = len(arr)
        for j in range(i + 1, n):
            comparisons += 1
    return {"comparisons": comparisons, "swaps": swaps}
def main():
   average_case = [3, 5, 1, 4, 2, 6, 9, 8, 10, 7]
```

```
# Make deep copies to preserve original arrays
best_case_sorted = copy.deepcopy(best_case)
average_case_sorted = copy.deepcopy(worst_case)
worst_case_sorted = copy.deepcopy(worst_case)

# Selection Sort on Best Case
print("Starting Selection Sort on Best Case:")
start_time = time.time()
best_stats = selection_sort(best_case_sorted)
end_time = time.time()
best_time = end_time - start_time
display_results("Best", best_case_sorted, best_time, best_stats)

print("\nStarting Selection Sort on Average Case:")
start_time = time.time()
average_stats = selection_sort(average_case_sorted)
end_time = time.time()
average_time = end_time - start_time
display_results("Average", average_case_sorted, average_time,
average_stats) # Step XIII: Display results (O(n))

# Selection Sort on Worst Case
print("\nStarting Selection_sort(worst_case_sorted)
end_time = time.time()
worst_stats = selection_sort(worst_case_sorted)
end_time = time.time()
worst_stats = selection_sort(worst_case_sorted)
end_time = time.time()
worst_time = end_time - start_time
display_results("Worst", worst_case_sorted, worst_time, worst_stats)

if __name__ == "__main__":
    main()
```

RUNNING:

```
=== Best Case ===

Sorted Array: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

Time Taken: 0.000000 seconds

Total Comparisons: 45

Total Swaps: 0
```

```
=== Average Case ===
Sorted Array: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
Time Taken: 0.003679 seconds
Total Comparisons: 45
Total Swaps: 4
```

```
=== Worst Case ===

Sorted Array: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

Time Taken: 0.000926 seconds

Total Comparisons: 45

Total Swaps: 5
```

TIME COMPLEXITY:

Visual Representation of Time Complexity

- 1. Best Case (Already Sorted):
 - Passes Needed: n
 - \circ Comparisons: (n*(n-1))/2
 - Swaps: 0 (if no new minimum is found, but in this implementation, a swap is only performed if a new minimum is found. Since the array is already sorted, no swaps occur.)
 - o **Time Complexity:** $O(n^2)$
- 2. Average Case (Random Order):
 - o Passes Needed: n
 - \circ Comparisons: (n*(n-1))/2
 - o **Swaps:** Approximately n (one swap per pass)
 - o **Time Complexity:** O(n²)
- 3. Worst Case (Reverse Sorted):
 - o Passes Needed: n
 - \circ Comparisons: (n*(n-1))/2
 - o **Swaps:** n (one swap per pass)
 - o **Time Complexity:** O(n²)

Practical Implications:

- Small Datasets:
 - Selection Sort can be acceptable for small arrays due to its simplicity and predictability.
- Large Datasets:
 - o Inefficient due to O (n²) time complexity.
 - Better sorting algorithms (like Merge Sort, Quick Sort, or Heap Sort) should be used for scalability.
- Memory Usage:

 Selection Sort is an in-place sorting algorithm with O (1) additional memory, making it memory-efficient.

• Stable Sorting:

 Standard Selection Sort is not stable. However, with modifications, it can be made stable.

CONCLUSION:

- **Best Case:** O (n²) No improvement even if the array is already sorted.
- Average Case: O (n²) consistently quadratic time complexity across all scenarios.
- Worst Case: $O(n^2)$ Maintains quadratic time complexity regardless of input.

Summary of Time Complexities

Component	Operation	Time Complexity
Imports	Importing modules	O(1)
Function Definitions	Defining selection_sort, display_results, main	O(1) each
selection_sort	Entire Selection Sort Algorithm	O(n²)
display_results	Displaying results	O(n)
main Function	Sorting Best, Average, Worst Cases	O(n²)
Execution Check	Running main	O(1)

MERGE SORT:

CODE:

```
def merge(arr, left, mid, right, comparisons, merges):
    L = arr[left:mid + 1]
    k = left
        comparisons += 1
        if L[i] <= R[j]:</pre>
            arr[k] = L[i]
            arr[k] = R[j]
        merges += 1
    while i < len(L):</pre>
        arr[k] = L[i]
        merges += 1
        arr[k] = R[j]
        merges += 1
    return comparisons, merges
return {"comparisons": comparisons, "merges": merges}
best case sorted = copy.deepcopy(best case)
average case sorted = copy.deepcopy(average case)
worst case sorted = copy.deepcopy(worst case)
```

```
print("Starting Merge Sort on Best Case:")
    start_time = time.time()
    best_stats = merge_sort(best_case_sorted)
    end_time = time.time()
    best_time = end_time - start_time
    display_results("Best", best_case_sorted, best_time, best_stats)

print("\nStarting Merge Sort on Average Case:")
    start_time = time.time()
    average_stats = merge_sort(average_case_sorted)
    end_time = time.time()
    average_time = end_time - start_time
    display_results("Average", average_case_sorted, average_time,
average_stats)

print("\nStarting Merge Sort on Worst Case:")
    start_time = time.time()
    worst_stats = merge_sort(worst_case_sorted)
    end_time = time.time()
    worst_time = end_time - start_time
    display_results("Worst", worst_case_sorted, worst_time, worst_stats)

if __name__ == "__main__":
    main()
```

Running:

```
=== Best Case ===
Sorted Array: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
Time Taken: 0.0000000 seconds
Total Comparisons: 19
Total Merges: 34

=== Average Case ===
Sorted Array: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
Time Taken: 0.0000000 seconds
Total Comparisons: 20
Total Merges: 34

=== Worst Case ===
Sorted Array: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
Time Taken: 0.0000000 seconds
Total Comparisons: 15
Total Merges: 34
```

TIME COMPLEXITY:

Detailed Explanation of Merge Sort Time Complexity:

Merge Sort is a divide-and-conquer algorithm that divides the input array into two halves, recursively sorts them, and then merges the sorted halves. Here's a more granular breakdown of its time complexity:

1. Dividing the Array:

- **Operation:** The array is repeatedly divided into halves until each subarray contains a single element.
- **Number of Divisions:** log₂ (n) divisions are needed to break the array down to single elements.
- **Time Complexity:** O(log n)

2. Merging the Arrays:

- **Operation:** Two sorted subarrays are merged into a single sorted array.
- **Number of Merge Operations:** Each level of division requires merging n elements in total.
- Time Complexity per Merge Level: O(n)
- Total Time for Merging Across All Levels: O(n log n)

3. Comparisons During Merge:

- **Operation:** Each element is compared once during the merge process.
- Time Complexity: O(n) per merge level
- **Total Comparisons:** O(n log n)

4. Overall Time Complexity:

- Combining Dividing and Merging: $O(n \log n) + O(n \log n) = O(n \log n)$
- **Explanation:** The algorithm consistently performs in O (n log n) time across all cases (best, average, worst) because the number of divisions and the merging process are unaffected by the initial order of elements.

5. Space Complexity:

- **Operation:** Merge Sort requires additional space for the temporary arrays during the merge process.
- Space Complexity: O(n)
- **Explanation:** Although the time complexity is efficient, Merge Sort uses extra memory proportional to the input size.

6. Stability:

- **Operation:** Merge Sort is a stable sorting algorithm.
- **Explanation:** It preserves the relative order of equal elements, which is beneficial in scenarios where the order carries meaning.

Visual Representation of Time Complexity

1. Best Case (Already Sorted):

Passes Needed: log₂(n)

• **Comparisons:** O(n log n)

• Merges: O(n log n)

• Time Complexity: O(n log n)

2. Average Case (Random Order):

• Passes Needed: log₂(n)

• Comparisons: O(n log n)

• Merges: O(n log n)

• Time Complexity: O(n log n)

3. Worst Case (Reverse Sorted):

• Passes Needed: log₂(n)

• Comparisons: O(n log n)

Merges: O(n log n)

• Time Complexity: O(n log n)

Practical Implications

Large Datasets:

- o **Advantage:** Efficient sorting with O (n log n) time complexity.
- o **Disadvantage:** Requires additional memory for temporary arrays.

• Small Datasets:

 Advantage: Still efficient, but simpler algorithms like Insertion Sort might perform better due to lower constant factors.

• Memory Constraints:

 Consideration: Since Merge Sort requires additional space, it may not be suitable for systems with limited memory.

• Stability Requirement:

 Advantage: Being a stable sort, Merge Sort is ideal when the stability of elements matters.

Conclusion:

- Merge Sort Characteristics:
- **Time Complexity:** O (n log n) across all cases.
- **Space Complexity:** O (n) due to additional temporary arrays.
- Stability: Merge Sort is stable by default.
- **Efficiency:** Highly efficient for large datasets compared to simple algorithms like Bubble Sort or Selection Sort.

• Implementation Highlights:

- **Recursive Division:** The array is recursively divided into halves until single-element subarrays are reached.
- **Merging Process:** Sorted subarrays are merged back together in a manner that maintains order.
- **Counting Operations:** The implementation counts the number of comparisons and merges to provide insight into the algorithm's performance.

• Performance Metrics:

- Comparisons: Reflects the number of element comparisons made during sorting.
- Merges: Indicates the number of times elements are merged into the main array.
- **Time Taken:** Measures the actual time consumed by the sorting process.

• Best Practices:

- **Deep Copies:** Making deep copies of the original arrays ensures that each sorting test starts with the unsorted data.
- **Time Measurement:** Using the time module to measure the duration of sorting provides empirical evidence of the algorithm's efficiency.
- **Detailed Logging:** Printing the state of the array after each merge helps in understanding how the algorithm progresses.

Summary of Time Complexities

Component	Operation	Time Complexity
Imports	Importing modules	O(1)
Function Definitions	Defining merge_sort, display_results, main	O(1) each
merge_sort	Entire Merge Sort Algorithm	O(n log n)
display_results	Displaying results	O(n)
main Function	Sorting Best, Average, Worst Cases	O(n log n)
Execution Check	Running main	O(1)

QUICK SORT:

Code:

```
import time
    comparisons = 0
    swaps = 0
           pi, comparisons, swaps = partition(arr, low, high, comparisons,
swaps)
           quick sort(arr, pi + 1, high)
    def partition(arr, low, high, comparisons, swaps):
           comparisons += 1
               arr[i], arr[j] = arr[j], arr[i]
               print(f" Swapping {arr[j]} and {arr[i]}")
       print(f" Array after partitioning: {arr}")
       return i + 1, comparisons, swaps
    return {"comparisons": comparisons, "swaps": swaps}
   best case = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] # Best Case: Already sorted
```

```
average_case = [3, 5, 1, 4, 2, 6, 9, 8, 10, 7] # Average Case: Random
order
  worst_case = [10, 9, 8, 7, 6, 5, 4, 3, 2, 1] # Worst Case: Reverse
sorted

best_case_sorted = copy.deepcopy(best_case)
  average_case_sorted = copy.deepcopy(average_case)
  worst_case_sorted = copy.deepcopy(worst_case)

print("Starting Quick Sort on Best Case:")
  start_time = time.time()
  best_stats = quick_sort(best_case_sorted)
  end_time = time.time()
  best_time = end_time - start_time
  display_results("Best", best_case_sorted, best_time, best_stats)

print("\nStarting Quick Sort on Average Case:")
  start_time = time.time()
  average_stats = quick_sort(average_case_sorted)
  end_time = time.time()
  average_time = end_time - start_time
  display_results("Average", average_case_sorted, average_time,
average_stats)

print("\nStarting Quick Sort on Worst Case:")
  start_time = time.time()
  worst_stats = quick_sort(worst_case_sorted)
  end_time = time.time()
  worst_stats = quick_sort(worst_case_sorted)
  end_time = time.time()
  worst_stats = quick_sort(worst_case_sorted, worst_time, worst_stats)

if __name__ == "__main__":
    main()
```

Running:

```
=== Best Case ===

Sorted Array: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

Time Taken: 0.0000000 seconds

Total Comparisons: 45

Total Swaps: 54
```

```
=== Average Case ===
Sorted Array: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
Time Taken: 0.016070 seconds
Total Comparisons: 23
Total Swaps: 22
```

=== Worst Case ===

Sorted Array: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

Time Taken: 0.001104 seconds

Total Comparisons: 45

Total Swaps: 29

TIME COMPLEXITY:

Detailed Explanation of Quick Sort Time Complexity

Quick Sort is a highly efficient sorting algorithm based on the divide-and-conquer paradigm. Here's a more granular breakdown of its time complexity:

1. Choosing a Pivot:

- o **Operation:** Select a pivot element from the array.
- **Time Complexity:** O (1) if choosing the last element as pivot, as in the provided implementation.

2. Partitioning the Array:

- Operation: Rearrange elements in the array such that elements less than or equal to the pivot are on the left, and elements greater than the pivot are on the right.
- **Time Complexity:** O (n), where n is the number of elements in the current subarray.
- o **Explanation:** Each element is compared to the pivot once.

3. Recursive Sorting of Subarrays:

- o **Operation:** Recursively apply Quick Sort to the subarrays formed by partitioning.
- o **Time Complexity:** O (log n) levels of recursion on average.
- **Explanation:** Each recursive call splits the array approximately in half, leading to a logarithmic number of levels.

4. Overall Time Complexity:

- o **Best Case:** O(n log n)
 - **Explanation:** Balanced partitions lead to logarithmic recursion depth with linear work at each level.
- o **Average Case:** O(n log n)
 - **Explanation:** Even with some unbalanced partitions, the overall time remains logarithmic multiplied by linear.
- Worst Case: O(n²)
 - **Explanation:** Highly unbalanced partitions (e.g., when the smallest or largest element is always chosen as pivot) lead to linear recursion depth with linear work at each level.

5. Space Complexity:

- o **Operation:** Requires O (log n) space for the recursion stack on average.
- **Explanation:** Each recursive call adds a layer to the stack, and with balanced partitions, the depth is logarithmic.

6. Stability:

- o **Operation:** Quick Sort is not a stable sort by default.
- o **Explanation:** Equal elements may not retain their original order after sorting.

7. Optimizations to Prevent Worst Case:

- o **Randomized Pivot Selection:** Choosing a random pivot reduces the probability of encountering the worst case.
- o **Median-of-Three Pivot Selection:** Choosing the median of the first, middle, and last elements as pivot can lead to more balanced partitions.

Practical Implications

- Large Datasets:
 - Advantage: Quick Sort is generally faster in practice compared to other O(n log
 n) algorithms like Merge Sort, especially due to better cache performance and inplace sorting.
 - o **Disadvantage:** Vulnerable to worst-case performance without optimizations.
- Small Datasets:
 - o **Advantage:** Efficient and quick to sort.
 - **Additional Note:** Sometimes, insertion sort is used for very small subarrays to optimize performance further.
- Memory Usage:
 - Advantage: In-place sorting with O (log n) additional space for the recursion stack.
 - o **Disadvantage:** Merge Sort requires O (n) additional space.
- Stability Requirement:
 - o **Consideration:** If stability is required, Quick Sort may not be suitable unless modified to be stable.

For Best Case:

- Time Complexity:
 - quick sort(best case sorted): O(n log n) on average
 - **display results:** O(n)
- Overall: $O(n \log n) + O(n) = O(n \log n)$

Best Case Average:

- Time Complexity:
 - quick_sort(average_case_sorted): O(n log n) on average
 - **display_results:** O(n)

• Overall: $O(n \log n) + O(n) = O(n \log n)$

Worst Case:

• Time Complexity:

• quick_sort(worst_case_sorted): O(n²) in the worst case

• **display_results:** O(n)

• **Overall:** $O(n^2) + O(n) = O(n^2)$

Summary of Time Complexities

Component	Operation	Time Complexity
Imports	Importing modules	O(1)
Function Definitions	Defining quick_sort, display_results, main	O(1) each
quick_sort	Entire Quick Sort Algorithm	O(n log n) on average, O(n²) worst case
display_results	Displaying results	O(n)
main Function	Sorting Best, Average, Worst Cases	O(n log n) on average, O(n²) worst case
Execution Check	Running main	O(1)

Conclusion

The provided Python script offers a comprehensive exploration of the Quick Sort algorithm across best, average, and worst-case scenarios. By incorporating detailed logging, performance metrics, and structured output, it serves as both an educational tool and a practical means of analyzing the algorithm's behavior under various conditions.

Key Points:

- Quick Sort Characteristics:
 - o **Time Complexity:** O (n log n) on average, O (n^2) in the worst case.
 - o **Space Complexity:** O (log n) due to recursion stack.
 - o **Stability:** Not stable by default.
 - o **Efficiency:** Highly efficient for large datasets with appropriate pivot selection.

• Implementation Highlights:

- **Recursive Division:** The array is recursively divided into partitions based on the pivot.
- **Partitioning Process:** Elements are rearranged around the pivot to ensure left <= pivot < right.
- **Counting Operations:** The implementation tracks the number of comparisons and swaps to provide insight into the algorithm's performance.

• Performance Metrics:

- Comparisons: Reflects the number of element comparisons made during sorting.
- **Swaps:** Indicates the number of times elements are swapped.
- **Time Taken:** Measures the actual time consumed by the sorting process.

• Best Practices:

- **Deep Copies:** Making deep copies of the original arrays ensures that each sorting test starts with the unsorted data.
- **Time Measurement:** Using the time module to measure the duration of sorting provides empirical evidence of the algorithm's efficiency.
- **Detailed Logging:** Printing the state of the array after each partition helps in understanding how the algorithm progresses.