Part 1: Selection Algorithms

Deterministic Selection Algorithm (Median of Medians)

Algorithm Explanation

The Deterministic Selection Algorithm (Median of Medians) finds the kth smallest element in an array in O(n) worst-case time complexity. The process involves:

- 1. Dividing the array into groups of 5.
- 2. Finding the median of each group.
- 3. Using the "median of medians" as the pivot to partition the array.
- 4. Recursively narrowing the search to find the desired element.

```
def median of medians(arr, k):
      return group[len(group) // 2]
  groups = [arr[i:i + 5] for i in range(0, len(arr), 5)]
  medians = [find_median(group) for group in groups]
```

```
high = [x for x in arr if x > median_of_medians_pivot]
```

```
Duplicates"),
  for array, k, description in test cases:
      start time = time.time()
      result = median of medians(array, k - 1) # k - 1 for 0-based indexing
      end time = time.time()
      print(f"Test Case: {description}")
if name == " main ":
```

- 1. Small Array:
 - Description: A simple array to validate correctness.
 - o Input: [3, 1, 2, 4, 5], k=2.
 - Output: 2.
 - o Time Taken: 0.0000020 seconds.
- 2. Large Array with Duplicates:

- Description: A large array with repeated values to test performance and handling of duplicates.
- o **Input**: Randomly generated array of size 1000, k=500.
- o **Output**: 49.
- o Time Taken: 0.0002180 seconds.
- 3. Reverse-Sorted Array:
 - o **Description**: A reverse-ordered array to ensure efficiency in sorted data.
 - Input: list(range(1000, 0, -1)), k=200.
 - o **Output**: 200.
 - Time Taken: 0.0003820 seconds.

Observations

- 1. The algorithm is highly efficient, even for large arrays, due to its O(n) complexity.
- 2. It handles duplicates and sorted data without any issues.
- 3. The observed time taken for small arrays is negligible, and performance scales well with input size.

Randomized Selection Algorithm (Randomized Quickselect)

Algorithm Explanation

The Randomized Selection Algorithm (Randomized Quickselect) finds the kth smallest element in an array using randomization. The key steps are:

- 1. Randomly selecting a pivot element.
- Partitioning the array into three parts:
 - o Elements smaller than the pivot.
 - Elements equal to the pivot.
 - Elements greater than the pivot.
- 3. Recursively narrowing the search based on k's position relative to the pivot partitions.
- 4. The algorithm achieves O(n) expected time complexity due to efficient partitioning and random pivot selection.

```
import random
import time

def randomized_select(arr, k):
   if len(arr) == 1:
```

```
return arr[0]
   # Randomly select a pivot
  pivot = random.choice(arr)
   # Partition the array into three parts
   low = [x for x in arr if x < pivot]</pre>
   high = [x for x in arr if x > pivot]
  pivots = [x for x in arr if x == pivot]
   # Determine where the k^th element lies
   if k < len(low):
       return randomized_select(low, k)
   elif k < len(low) + len(pivots):</pre>
      return pivots[0]
   else:
       return randomized_select(high, k - len(low) - len(pivots))
# Function to test the algorithm with timing
def test_randomized_select():
  test_cases = [
       # Edge Case 1: Small array
       ([3, 1, 2, 4, 5], 2, "Small Array"),
```

```
# Edge Case 2: Large array with duplicates
       ([random.randint(1, 100) for _ in range(1000)], 500, "Large Array with
Duplicates"),
      # Edge Case 3: Reverse-sorted array
       (list(range(1000, 0, -1)), 200, "Reverse-Sorted Array"),
  for array, k, description in test cases:
      start time = time.time()
      result = randomized select(array, k - 1) # k - 1 for 0-based indexing
      end time = time.time()
      print(f"Test Case: {description}")
      print(f"The {k}th smallest element is: {result}")
      print(f"Time taken: {end time - start time:.6f} seconds\n")
if name == " main ":
  test randomized select()
```

- 1. Small Array:
 - o **Description**: Validates correctness on a small input.
 - o Input: [3, 1, 2, 4, 5], k=2.
 - o Output: 2.
 - o Time Taken: 0.0000070 seconds.

2. Large Array with Duplicates:

o **Description**: Tests performance with duplicates.

• **Input**: Randomly generated array of size 1000, k=500.

o **Output**: 54.

o Time Taken: 0.0002290 seconds.

3. Reverse-Sorted Array:

• **Description**: Ensures correctness with reverse-ordered data.

○ Input: list(range(1000, 0, -1)), k=200.

o **Output**: 200.

Time Taken: 0.0003040 seconds.

Observations

- 1. The algorithm performs efficiently on small and large inputs.
- 2. The execution time scales linearly with the input size, as expected for an O(n) algorithm.
- 3. The results confirm correctness even with edge cases like duplicates and sorted inputs.

Comparison of Deterministic and Randomized Algorithms

Both the **Deterministic Selection Algorithm (Median of Medians)** and the **Randomized Selection Algorithm (Quickselect)** are efficient and widely used methods for finding the kth smallest element in an array. While they share the same O(n) linear time complexity for most inputs, their internal mechanisms, performance characteristics, and suitability for specific applications differ significantly.

Observations

- 1. Deterministic Algorithm (Median of Medians):
 - **Guaranteed Performance**: The deterministic algorithm guarantees O(n) time complexity even in the worst case. This is achieved through the careful selection of a pivot using the "median of medians" method, ensuring balanced partitioning of the input array.
 - Robustness: It is highly reliable and suitable for scenarios where predictability and worst-case guarantees are crucial, such as real-time systems, critical financial applications, and performance-sensitive environments.
 - Overhead: The process of calculating the median of medians adds some computational overhead compared to the randomized approach. This makes it slightly slower in practice, especially for smaller inputs or non-adversarial data distributions.

2. Randomized Algorithm (Quickselect):

 Efficiency in Practice: The randomized algorithm is generally faster in real-world scenarios due to its simplicity. By selecting the pivot randomly, it avoids the additional overhead of computing the median of medians.

- Expected Performance: While it has an expected O(n) time complexity for most inputs, there is a small chance of performance degradation in specific edge cases where random pivot selection leads to highly unbalanced partitions.
- Suitability: This algorithm is ideal for general-purpose use cases where the input is expected to follow typical distributions, such as sorting user-generated data or handling large datasets in non-critical systems.

Practical Implications

• Deterministic Algorithm:

- Best suited for applications requiring performance guarantees and robustness, where unpredictable behavior cannot be tolerated.
- Examples include real-time control systems, high-stakes financial systems, or environments where inputs might be adversarial (e.g., security-sensitive applications).

Randomized Algorithm:

- Preferred for general-purpose, real-world applications due to its simplicity and faster performance on average.
- Examples include data analysis pipelines, machine learning preprocessing, and other tasks where occasional performance variability is acceptable.

Overall Summary

While both algorithms are efficient and achieve linear time complexity, their suitability depends on the application context:

- The deterministic approach is more reliable and robust, making it the choice for critical systems with stringent performance requirements. However, this comes at the cost of slightly higher overhead.
- The randomized approach is faster and simpler for most real-world scenarios, making it the preferred choice when average-case performance is sufficient and execution speed is prioritized.

Part 2: Elementary Data Structures Implementation

Arrays and Matrices

Algorithm Explanation

1. Array Operations:

- Insertion: Uses Python's list.insert(index, value) to insert an element at a specific index.
- Deletion: Removes an element using list.pop(index).
- Access: Retrieves an element at a given index.
- Display: Returns the array content as a list.

```
class Array:
   def init (self):
      self.data = []
   def insert(self, index, value):
       if 0 <= index <= len(self.data):</pre>
           self.data.insert(index, value)
      else:
           raise IndexError("Index out of bounds")
   def delete(self, index):
       if 0 <= index < len(self.data):</pre>
          return self.data.pop(index)
      else:
           raise IndexError("Index out of bounds")
   def access(self, index):
       if 0 <= index < len(self.data):</pre>
          return self.data[index]
      else:
           raise IndexError("Index out of bounds")
   def display(self):
      return self.data
class Matrix:
  def __init__(self, rows, cols):
       self.data = [[0] * cols for _ in range(rows)]
```

```
def insert(self, row, col, value):
       if 0 <= row < len(self.data) and 0 <= col < len(self.data[0]):</pre>
           self.data[row][col] = value
       else:
           raise IndexError("Index out of bounds")
   def access(self, row, col):
       if 0 \le row \le len(self.data) and 0 \le col \le len(self.data[0]):
          return self.data[row][col]
      else:
           raise IndexError("Index out of bounds")
  def display(self):
      return self.data
 Example usage
if __name__ == "__main__":
  # Array Operations
  array = Array()
  array.insert(0, 10) # Insert at index 0
  array.insert(1, 20) # Insert at index 1
  array.insert(1, 15) # Insert at index 1
  print("Array after insertions:", array.display())
  print("Access element at index 1:", array.access(1))
  print("Deleted element at index 0:", array.delete(0))
  print("Array after deletion:", array.display())
```

```
# Matrix Operations

matrix = Matrix(3, 3)

matrix.insert(1, 1, 5)  # Insert value 5 at position (1, 1)

matrix.insert(0, 2, 10)  # Insert value 10 at position (0, 2)

print("Matrix after insertions:")

for row in matrix.display():
    print(row)

print("Access element at position (1, 1):", matrix.access(1, 1))
```

2. Matrix Operations:

- Insertion: Updates the value at a specified row and column using simple indexing.
- Access: Retrieves the value at a given row and column.
- o **Display**: Returns the entire matrix as a 2D list.

Edge Cases and Results

1. Array:

- o **Insertion**: Tested for beginning, middle, and end positions.
- Access: Retrieved values from valid indices.
- **Deletion**: Removed elements correctly, even at boundaries.

Output:

```
Array after insertions: [10, 15, 20]
Access element at index 1: 15
Deleted element at index 0: 10
Array after deletion: [15, 20]
```

2. Matrix:

- Insertion: Values added at different row-column positions.
- o Access: Retrieved values at specified positions.

Output:

```
Matrix after insertions:
[0, 0, 10]
[0, 5, 0]
[0, 0, 0]
Access element at position (1, 1): 5
```

Observations

1. Efficiency:

- o Operations on arrays and matrices are efficient for small and moderate sizes.
- Python's list structure simplifies implementation.

2. Error Handling:

o Index out-of-bounds access is correctly handled with IndexError.

3. Practical Applications:

- Arrays are suitable for scenarios requiring constant-time access and sequential data storage.
- Matrices are effective for representing 2D grids, such as in image processing or game boards.

Stacks and Queues Using Arrays

```
class Stack:
    def __init__(self):
        self.data = []

    def push(self, value):
        self.data.append(value)

    def pop(self):
        if not self.is_empty():
            return self.data.pop()
        else:
            raise IndexError("Pop from empty stack")

    def peek(self):
```

```
if not self.is_empty():
          return self.data[-1]
      else:
          raise IndexError("Peek from empty stack")
  def is_empty(self):
      return len(self.data) == 0
  def display(self):
      return self.data
class Queue:
  def __init__(self):
      self.data = []
  def enqueue(self, value):
      self.data.append(value)
  def dequeue(self):
      if not self.is_empty():
          return self.data.pop(0)
      else:
          raise IndexError("Dequeue from empty queue")
  def peek(self):
      if not self.is_empty():
          return self.data[0]
      else:
          raise IndexError("Peek from empty queue")
```

```
def is_empty(self):
      return len(self.data) == 0
  def display(self):
      return self.data
# Example usage
if __name__ == "__main__":
  # Stack Operations
  stack = Stack()
  stack.push(10)
  stack.push(20)
  stack.push(30)
  print("Stack after pushes:", stack.display())
  print("Peek top of stack:", stack.peek())
  print("Pop from stack:", stack.pop())
  print("Stack after pop:", stack.display())
   # Queue Operations
  queue = Queue()
  queue.enqueue(10)
  queue.enqueue(20)
  queue.enqueue(30)
  print("Queue after enqueues:", queue.display())
  print("Peek front of queue:", queue.peek())
  print("Dequeue from queue:", queue.dequeue())
```

Algorithm Explanation

- 1. Stack Operations:
 - **Push**: Adds an element to the top of the stack using list.append(value).
 - Pop: Removes and returns the top element using list.pop().
 - Peek: Returns the top element without removing it.
 - o **Is Empty**: Checks if the stack is empty by comparing the length to zero.
 - Display: Returns the current contents of the stack.
- 2. Queue Operations:
 - Enqueue: Adds an element to the rear of the queue using list.append(value).
 - \circ **Dequeue**: Removes and returns the front element using list.pop(\emptyset).
 - o Peek: Returns the front element without removing it.
 - o **Is Empty**: Checks if the queue is empty by comparing the length to zero.
 - o **Display**: Returns the current contents of the queue.

Edge Cases and Results

Stack:

- **Operations Tested**: Push, peek, pop, and operations on an empty stack.
- Output:

```
Stack after pushes: [10, 20, 30]
Peek top of stack: 30
Pop from stack: 30
Stack after pop: [10, 20]
```

Queue:

- Operations Tested: Enqueue, peek, dequeue, and operations on an empty queue.
- Output

```
Queue after enqueues: [10, 20, 30]
Peek front of queue: 10
Dequeue from queue: 10
Queue after dequeue: [20, 30]
```

Observations

1. Correctness:

- All operations behave as expected, including boundary cases (e.g., empty stack or queue).
- Error handling for empty operations is robust.

2. Efficiency:

- Stacks are efficient due to constant-time operations at the end of the list.
- Queues are less efficient because list.pop(0) requires shifting elements, making it O(n).

3. Practical Applications:

- Stacks: Ideal for LIFO tasks like function call management or undo operations.
- Queues: Useful for FIFO tasks like task scheduling or processing requests in order.

Singly Linked Lists

Algorithm Explanation

1. Insertion:

- At the Beginning: Creates a new node and points it to the current head.
 Updates the head to the new node.
- At the End: Traverses the list to the last node and appends the new node.

2. Deletion:

- o Removes a node by value.
- Special handling for:
 - The head node (updating the head pointer).
 - Non-existent values (raises a ValueError).

3. Traversal:

o Iterates through the list and collects all node values in a list.

4. Search:

Iterates through the list to check if a specific value exists.

```
class Node:
    def __init__(self, data):
        self.data = data
        self.next = None
```

```
class SinglyLinkedList:
  def insert_at_beginning(self, value):
          self.head = new_node
```

```
if self.head.data == value:
  elements.append(current.data)
```

```
return False
if __name__ == "__main__":
  linked_list.insert_at_beginning(10)
  linked_list.insert_at_beginning(5)
```

```
linked_list.delete_by_value(20)

print("List after deleting last element (20):", linked_list.traverse())
```

- 1. Insertion:
 - Inserted nodes at the beginning and end.
 - Output: [5, 10, 20]
- 2. Search:
 - Found an existing value (202020).
 - o Could not find a non-existent value (151515).

Output:

```
Search for 20: True
Search for 15: False
```

0

3. Deletion:

- Deleted a node from the middle (101010).
- Deleted the head node (555).
- Deleted the last remaining node (202020), leaving the list empty.

Output:

```
List after deleting 10: [5, 20]
List after deleting 5: [20]
List after deleting last element (20): []
```

Observations

1. Correctness:

- All operations (insertion, deletion, traversal, and search) work as expected.
- The algorithm handles edge cases like deleting the only element or searching for a non-existent value.

2. Efficiency:

- Operations like insertion and traversal are O(n) for a linked list.
- o Deletion and search require traversal, making them O(n) as well.

3. Practical Applications:

- Linked lists are ideal for scenarios where frequent insertions or deletions occur, especially at the beginning or middle of a list.
- They are not optimal for random access tasks due to sequential traversal.

Rooted Trees Using Linked Lists

Algorithm Explanation

- 1. Representation:
 - Each node (TreeNode) contains:
 - Value: The data stored in the node.
 - **Children**: A list of child nodes representing the relationships.
- 2. Operations:
 - Add Child: Appends a new node to the children list of a parent node.
 - **Remove Child**: Searches for a child by its value and removes it from the children list. If not found, raises a ValueError.
 - Traverse: Recursively traverses the tree in depth-first order, collecting all node values.

```
raise ValueError(f"Child with value {child_value} not found")
  def traverse(self):
if name == " main ":
  child2 = TreeNode("Child2")
```

```
childl.add_child(childl_1)
childl.add_child(childl_2)

# Add children to Child2
child2_1 = TreeNode("Child2_1")
child2.add_child(child2_1)

# Traversal
print("Tree traversal:", root.traverse())

# Remove a child
root.remove_child("Child2")
print("Tree traversal after removing Child2:", root.traverse())
```

- 1. Initial Tree Traversal:
 - Input: Root with children (Child1, Child2) and grandchildren (Child1_1, Child1_2, Child2_1).

Output:

```
Tree traversal: ['Root', 'Child1', 'Child1_1', 'Child1_2', 'Child2_1']
```

- 2. Remove a Child Node (Child2):
 - o **Input**: Remove a child node (Child2) with its own child (Child2_1).

Output:

```
Tree traversal after removing Child2: ['Root', 'Child1', 'Child1_1',
'Child1_2']
```

Observations

1. Correctness:

- The tree correctly represents parent-child relationships.
- o Removal operations correctly handle nodes with and without children.

2. Efficiency:

- Operations scale with the number of children at each node, making the traversal O(n), where n is the total number of nodes.
- Addition and removal are efficient due to direct list operations.

3. Practical Applications:

 Rooted trees are ideal for hierarchical data structures like file systems, organizational charts, and DOM trees in web development.

Performance Analysis of Task Scheduler App

Performance Results

The app uses arrays, matrices, stacks, queues, and rooted trees for task management. Below are the time measurements for each operation:

Time Complexity Analysis

1. **Array**:

- o **Insertion**: O(n), due to potential shifts.
- Deletion: O(n), as it may require shifting elements.

2. Matrix:

- Add Dependency: O(1), accessing a specific cell.
- Remove Dependency: O(1), accessing a specific cell.

3. **Stack**:

- **Push**: O(1), appends to the end.
- Pop: O(1), removes the last element.

4. Queue:

- **Enqueue**: O(1), appends to the end.
- **Dequeue**: O(n), removes the first element (shifting required).

5. **Tree**:

• Add Subtask: O(1), adds to the children list.

```
from Arrays_and_Matrics import Array, Matrix
from Stack_and_queue import Stack, Queue
from Rooted_Tree import TreeNode
def measure_time(func, *args):
  result = func(*args)
  end = time.time()
if __name__ == "__main__":
```

```
_, time_stack_push = measure_time(undo_stack.push, "Undo-Task")
_, time_stack_pop = measure_time(undo_stack.pop)
_, time_queue_enqueue = measure_time(task_queue.enqueue, tasks[1])
_, time_queue_dequeue = measure_time(task_queue.dequeue)
child task = TreeNode("Subtask")
```

```
print(f"Matrix Remove Dependency Time: {time_matrix_remove:.6f} seconds")

print(f"Stack Push Time: {time_stack_push:.6f} seconds")

print(f"Stack Pop Time: {time_stack_pop:.6f} seconds")

print(f"Queue Enqueue Time: {time_queue_enqueue:.6f} seconds")

print(f"Queue Dequeue Time: {time_queue_dequeue:.6f} seconds")

print(f"Tree Add Subtask Time: {time_tree_add:.6f} seconds")
```

```
python3 App.py
Array Add Task Time: 0.000002 seconds
Array Delete Task Time: 0.000001 seconds
Matrix Add Dependency Time: 0.000001 seconds
Matrix Remove Dependency Time: 0.000001 seconds
Stack Push Time: 0.000000 seconds
Stack Pop Time: 0.000001 seconds
Queue Enqueue Time: 0.000001 seconds
Queue Dequeue Time: 0.000001 seconds
Tree Add Subtask Time: 0.0000000 seconds
```

Trade-offs

1. Array vs. Linked List for Stacks and Queues:

- Arrays: Faster random access and less overhead, but insertion/deletion at arbitrary points can be slower due to shifting.
- Linked Lists: Efficient insertion and deletion anywhere, but additional memory overhead for pointers.

2. Matrix:

- o Ideal for representing fixed-size relationships, such as dependencies or graphs.
- Memory usage increases quadratically with size (n^2).

3. **Tree**:

- Suited for hierarchical data.
- Efficient traversal and addition of nodes but requires recursion for deep structures.

Practical Applications

1. Arrays:

Manage simple, ordered lists of tasks.

2. Matrices:

Represent task dependencies or adjacency in graphs.

3. **Stacks**:

Track undo actions in task modifications.

4. Queues:

Process tasks in order of arrival (FIFO).

5. Trees:

o Organize hierarchical tasks with subtasks.

References

- 1. Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms* (3rd ed.). MIT Press.
- 2. Goodrich, M. T., Tamassia, R., & Mount, D. M. (2011). *Data Structures and Algorithms in Python*. Wiley.
- Blum, M., Floyd, R. W., Pratt, V., Rivest, R. L., & Tarjan, R. E. (1973). Time bounds for selection. *Journal of Computer and System Sciences*, 7(4), 448–461. https://doi.org/10.1016/S0022-0000(73)80033-9
- 4. Hoare, C. A. R. (1961). Algorithm 65: Find. *Communications of the ACM*, 4(7), 321–322. https://doi.org/10.1145/366622.366647
- 5. Knuth, D. E. (1997). *The Art of Computer Programming, Volume 1: Fundamental Algorithms* (3rd ed.). Addison-Wesley.
- 6. Tarjan, R. E. (1979). Applications of path compression on balanced trees. *Journal of the ACM*, 26(4), 690–715. https://doi.org/10.1145/322154.322161
- 7. McConnell, J. J. (2008). *Analysis of Algorithms: An Active Learning Approach* (2nd ed.). Jones & Bartlett Learning.
- 8. Sedgewick, R., & Wayne, K. (2011). Algorithms (4th ed.). Addison-Wesley.