University of Colombo School of Computing
SCS 1304 - Problem Solving Strategies and Computation Approaches

Lab Sheet 10

Answer the following questions and you are allowed to use non programmable calculators.

On an election day in Srilanka, data flows in from thousands of polling stations across the country. The system must process, organize, and analyze this data in real-time for quick updates to the public. Several algorithms and strategies are applied by the data maintenance team for different tasks in the workflow.

Part A

Incoming vote count records from 5,000 polling stations must be sorted by the time of arrival to ensure proper sequencing in the update dashboard. The team chooses Quick Sort, using a Divide and Conquer approach.

1. **Explain the role of the pivot in Quick Sort and how it partitions the data.**
2. **Demonstrate the first two recursive steps of Quick Sort for the dataset**

   [245, 120, 375, 180, 320, 150].
3. **Calculate the worst-case number of comparisons needed to sort all 5,000 records.**
4. **Suggest one improvement to avoid the worst-case performance in Quick Sort.**

Part B

Once all records are ordered by arrival time, each district's data must be sorted by candidate ID before aggregation. The team decides to use Merge Sort for this step.

1. **Explain how the Divide and Conquer strategy works in Merge Sort with the example**

   [34, 27, 40, 21, 30, 25].
2. **Estimate the number of comparisons needed to sort 1,000 records per district using**

   $T(n) = n \times \log_2(n)$.
3. **If data is processed on 10 servers (each handling a group of districts), describe how parallel Merge Sort could improve performance.**

Part C

To identify suspicious voting patterns, the system multiplies large vote correlation matrices. For a smaller test case, consider:

Matrix A:

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$$

Matrix B:

$$\begin{bmatrix} 5 & 1 \\ 2 & 3 \end{bmatrix}$$

1. **Use Strassen's algorithm to multiply these matrices, showing all seven intermediate products and the final result.**
2. **Compare the total multiplication operations used in Strassen's method with the standard brute-force method for this case.**

Part D

The fraud detection unit runs an exhaustive search to detect duplicate or suspicious records by checking every transaction against all others – a Brute Force method.

1. **Explain why Brute Force is easy to implement but not efficient for large datasets.**
2. **If there are 100,000 transactions, calculate the approximate number of comparisons the Brute Force method would require.**
3. **Propose a Divide and Conquer-based approach that could replace Brute Force for this problem, and explain how it reduces complexity**

Part E

During the election data processing, a recursive function is used to split and process vote records:

```
ProcessData(n)
  if n == 1
    return 1
  else
    return 2 * ProcessData(n / 2) + n
```

1. **Derive the recurrence relation for the time complexity T(n) of the above function.**
2. **Use mathematical induction to prove that $T(n) = n \log_2 n + n$ is the solution for the recurrence.**
3. **Explain why the time complexity result aligns with the complexities of algorithms such as Merge Sort.**