CSCI 6212 – Design and Analysis of Algorithms (Fall 2023)

Project Assignment 2

Abde Manaaf Ghadiali – G29583342

Problem Statement

(Staircase/Pareto-Optimal Points) Let P be a set of n points in a 2-dimensional plane. A point p in P is Pareto-Optimal if no other point is both above and to the right of p. The sorted sequence of Pareto-Optimal points describes a top-right staircase with the interior points of P below and to the left of the staircase. We want to compute this staircase. Describe a divide and conquer algorithm to compute the staircase of P in O(n * log(n)) time.

Theoretical Analysis

Pseudo Code (Using Divide and Conquer Approach)

```
function ParetoOptimal(P)
  if P has one point
    return P
  else
    // Divide P into two sets L and R such that L contains the left half points and R contains the right half points.
    // Call ParetoOptimal(L) and ParetoOptimal(R) recursively to get the Pareto-Optimal points in L and R respectively.
    //Merge the two sets of Pareto-Optimal points using a merge-like process to get the Pareto-Optimal points in P.
    return Pareto-Optimal points in P.
```

Experimental Analysis

- **GitHub Project Repository Link –** **https://github.com/default741/CSCI_6212_Course_Notes/blob/main/project-files/project-02/staircase_pareto_optimal.py**
- **Program Listing**

  I have executed the code for values of n ranging from 2 to 1E5 with increments of 25% of the nearest multiple of 10, which can be seen in the following section of Output Numerical Data.

- **Data Normalization Notes**

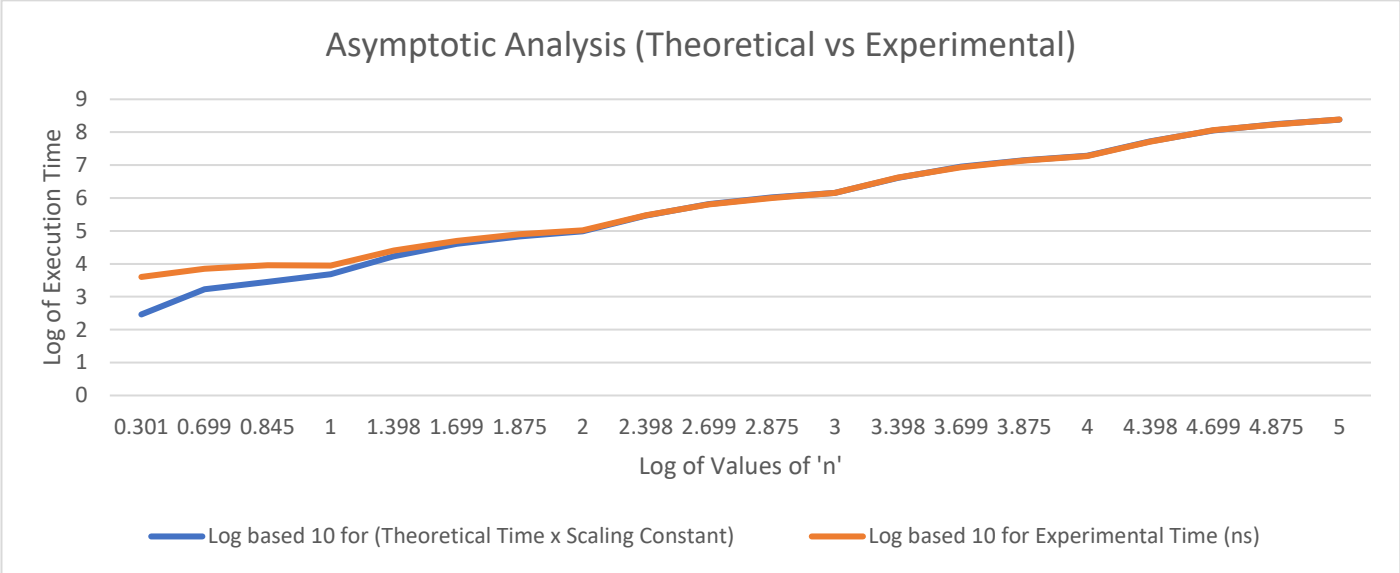
  To Normalize the theoretical time, we take the average of the experimental time and divide that by the average of the theoretical time, that give us the scaling constant which we multiply with the theoretical time. This gets the theoretical and experimental values in the same range. But the values of n are still too large to compare with the execution times. So, we log the values of n as well as we log the values of theoretical and experimental execution Times.

- **Output Numerical Data**

Value of 'n'	Theoretical Time Taken (O(n * log n))	Experimental Time Taken (ns)	Scaling Constant	Theoretical Time x Scaling Constant	Log based 10 for 'n'	Log based 10 for (Theoretical Time x Scaling Constant)	Log based 10 for Experimental Time (ns)
2.00E+00	2.00	4000.03	144.88	289.77	0.301	2.462048184	3.602063298

5.00E+00	11.61	7100.05	144.88	1682.04	0.699	3.225836959	3.851261656
7.00E+00	19.65	8999.95	144.88	2847.17	0.845	3.454413551	3.954240199
1.00E+01	33.22	8899.95	144.88	4812.92	1	3.682408416	3.949387633
2.50E+01	116.10	25199.96	144.88	16820.42	1.398	4.225836959	4.401399835
5.00E+01	282.19	49500.03	144.88	40885.01	1.699	4.611564132	4.694605442
7.50E+01	467.16	79699.91	144.88	67683.86	1.875	4.830485141	4.901457822
1.00E+02	664.39	103799.98	144.88	96258.35	2	4.983438412	5.016197251
2.50E+02	1991.45	297100.05	144.88	288527.19	2.398	5.460186739	5.472902725
5.00E+02	4482.89	642000.00	144.88	649496.00	2.699	5.812576478	5.807535027
7.50E+02	7163.06	1009500.00	144.88	1037807.45	2.875	6.016116783	6.004106325
1.00E+03	9965.78	1422800.01	144.88	1443875.25	3	6.159529671	6.153143859
2.50E+03	28219.28	4150800.06	144.88	4088501.22	3.398	6.611564132	6.618131814
5.00E+03	61438.56	8541499.96	144.88	8901418.71	3.699	6.94945923	6.931534143
7.50E+03	96545.06	13665499.98	144.88	13987762.59	3.875	7.145748253	7.135625526
1.00E+04	132877.12	18722599.95	144.88	19251669.95	4	7.284468408	7.272366158
2.50E+04	365241.01	51124199.99	144.88	52917305.94	4.398	7.723597726	7.708626525
5.00E+04	780482.02	115167099.98	144.88	113078774.54	4.699	8.053381093	8.061328431
7.50E+04	1214595.22	174509400.04	144.88	175974507.06	4.875	8.245449757	8.241818825
1.00E+05	1660964.05	242957099.92	144.88	240645874.38	5	8.381378421	8.385529595

- **Graph**



- **Graph Observations**

  We can visually observe from the graph that the Theoretical Execution Time and Experimental Execution Time follow a similar path for same values of n. Since we assume that the variable assignment and conditional checking and expression evaluation takes constant time for theoretical values, it does not hold true for experimental values. Hence, we see the Bumps in the graph for the experimental values.

Conclusion

Hence, we can conclude that the problem statement's (Staircase/Pareto-Optimal Points) Theoretical Execution Time of O(n * log n) using divide and conquer approach matches with the experimental result.