Ebrahim Qasem

COSC 311, Fall 2019

12/10/2019

External & Internal Sort tables and charts

Internal Sort x\*log(x). BASE 2

|  |  |  |
| --- | --- | --- |
| Internal Sort | | |
| Dataset Size (x) | x\*log2(x) | Time (y) |
| 10 | 33.2 | 34 |
| 20 | 86.4 | 88 |
| 30 | 147.3 | 148 |
| 50 | 282 | 286 |
| 100 | 664 | 672 |
| 200 | 1528 | 1,544 |
| 300 | 2469 | 2,488 |
| 500 | 4485 | 4,488 |
| 1,000 | 9970 | 9,976 |
| 2,000 | 21940 | 21,952 |
| 3,000 | 34650 | 34,904 |
| 5,000 | 61450 | 61,808 |
| 10,000 | 132900 | 133,616 |

Internal Sort x^2

|  |  |  |
| --- | --- | --- |
| Internal Sort | | |
| Dataset Size (x) | x^2 | Time (y) |
| 10 | 100 | 34 |
| 20 | 400 | 88 |
| 30 | 900 | 148 |
| 50 | 2500 | 286 |
| 100 | 10000 | 672 |
| 200 | 40000 | 1,544 |
| 300 | 90000 | 2,488 |
| 500 | 250000 | 4,488 |
| 1,000 | 1,000,000 | 9,976 |
| 2,000 | 4,000,000 | 21,952 |
| 3,000 | 9,000,000 | 34,904 |
| 5,000 | 25,000,000 | 61,808 |
| 10,000 | 100,000,000 | 133,616 |

External Sort x\*log(x). BASE 2

|  |  |  |
| --- | --- | --- |
| External Sort | | |
| Dataset Size (x) | x\*log2(x) | Time (y) |
| 10 | 33.2 | 55 |
| 20 | 86.4 | 145 |
| 30 | 147.3 | 246 |
| 50 | 282 | 485 |
| 100 | 664 | 1198 |
| 200 | 1528 | 2,767 |
| 300 | 2469 | 4,491 |
| 500 | 4485 | 8,082 |
| 1,000 | 9970 | 18,126 |
| 2,000 | 21940 | 39,958 |
| 3,000 | 34650 | 64,098 |
| 5,000 | 61450 | 113,999 |
| 10,000 | 132900 | 247,435 |

External Sort x^2

|  |  |  |
| --- | --- | --- |
| External Sort | | |
| Dataset Size (x) | x^2 | Time (y) |
| 10 | 100 | 55 |
| 20 | 400 | 145 |
| 30 | 900 | 246 |
| 50 | 2500 | 485 |
| 100 | 10000 | 1198 |
| 200 | 40000 | 2,767 |
| 300 | 90000 | 4,491 |
| 500 | 250000 | 8,082 |
| 1,000 | 1,000,000 | 18,126 |
| 2,000 | 4,000,000 | 39,958 |
| 3,000 | 9,000,000 | 64,098 |
| 5,000 | 25,000,000 | 113,999 |
| 10,000 | 100,000,000 | 247,435 |

Fitting the curve:

* For n \* log(n)

I took the size of data n and multiply by log n. Then, I selected the result column of that for x-axis and the time column for y-axis.

* For n^2

I did the same as for n\*log(n). I squared each size of data. Then, I selected the squared column for x-axis and the time column for y-axis.

Summary:

The algorithm confirmed O(n log n) for internal and external merge sort as we can see in the charts above. The error of the graph was calculated by multiplying log x by x and then taking the result of that with the time already calculated from the program.