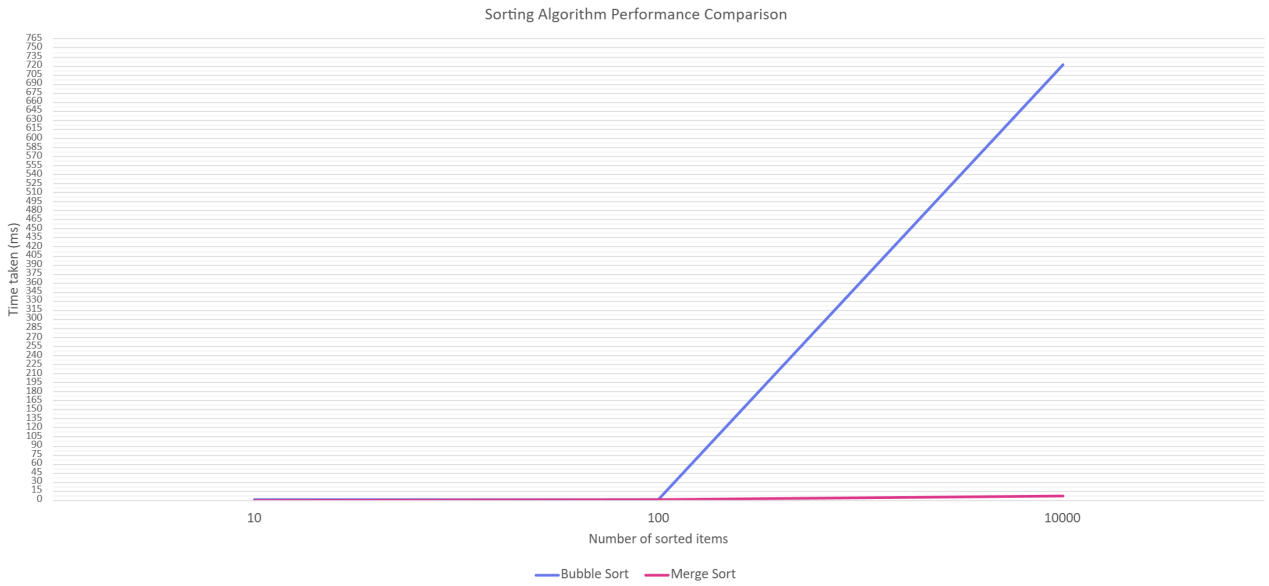
CS1IP Coursework 2

Sort Comparison

Freddie Taylor



The two sorts initially perform similarly - with the 10 and 100 item benchmarks lacking any easily perceivable speed differences, despite the different time complexities. This is because the $O(n \log_2 n)$ complexity of the merge sort and $O(n^2)$ of bubble result in an extremely similar number of worst-case comparison operations at lower item counts

	10 items	100 items	10000 items
Merge Sort	$\approx$ 33 comparisons	$\approx$ 664 comparisons	$\approx$ 132,877
Bubble Sort	$\approx$ 45 comparisons	$\approx$ 4950 comparisons	$\approx$ 49,995,000

The differences in comparison count where items ≤ 100 between the two sorts shown above are relatively inconsequential to modern CPUs unless performed frequently enough for the differences to compound, as shown by the results all rounding approximately to a millisecond. This suggests it's not always that beneficial to use a more complex, 'faster' sort when you are handling smaller arrays due to the marginal improvements found.

Bubble Sort also has the advantage that it can detect when an array is sorted after every iteration through it, which gives it a potential best case time complexity of $O(n)$

The major performance difference is observed when the number of items increases to 10,000. Due to merge sort's logarithmic time complexity, it performs, at most, 99,867,123 less comparison operations, showing that the performance difference of the two significantly increases as item count increases. This is reflected in the results found from the comparison benchmarks. The time taken inflates from similar times between the two sorts ($0 - 1ms$) in the smaller sorts to $\sim$ 750 ms (With memoisation, tested on a Ryzen 5 7600) for bubble sort, and $\sim$ 6 ms for merge sort 10,000 items. This is because logarithmic time complexity's rate of growth decreases as item count grows, while the quadratic complexity of bubble sort has a linear rate of growth.

The practical effect of this is that merge sort is much faster for larger data sets.
I also tested the algorithms without memoisation of card value parsing. Merge sort's result is effectively the same for 10,000, at $\sim$ 7 ms, while bubble sort's increases even more to $\sim$ 1700 ms, showing how the few extra CPU cycles added every iteration have a major impact on overall performance.