Scientific Computation Project 1 class feedback

November 15, 2024

The class did well on the assignment overall. A few comments on the specific questions:

- Part 1 question 1: The class did very well overall. A few students didn't include the effect of m in their cost analysis.

- Part 1 question 2: Here, for many submissions, we were looking for greater depth in the analysis. Visual inspection of timing results was not considered sufficient, and the use of appropriate axis scales and least-squares fits provide much greater insight. Note that it is difficult to visually distinguish between linear and $N \log N$ behavior. We also were looking for careful variation of both m and n to show how the superiority of method 2 for large m depends on n.

- Part 2 question 1: Many submissions used a variant of Rabin-Karp, but it is crucial to store pattern hashes in a dictionary so that $\mathcal{O}(1)$ lookup can be used when the rolling hash calculations take place.

- Part 2 question 2: Many discussion of the cost focused on worst-case scenarios which are highly unlikely to arise. In such cases, it is important to consider other cases as well which are more useful for understanding the efficiency of a code/algorithm. When considering if a given code is efficient, it is essential to consider the efficiency of *both* the underlying algorithm and its implementation. Many focused on just the latter.