

# CS 840 Microassignment 1

Due September 17, 2020 (anytime before 11:59 pm)

This is a “get going” assignment. The point is to start thinking about efficiency issues in general. It shouldn't take you more than a few minutes.

Suppose you have a large static file of keys, you plan to have them in sorted order, on disk. From time to time, you are to do a search for a key, determining whether it is present and the number of elements less than it. You are concerned about the worst time this will take for a search, so binary search takes about  $\lg n$  comparisons (interpolation takes more in the worst case, and van Emde Boas is not appropriate because of the large key domain or some other reason). But the  $\lg n$  comparisons is not the main issue, what else will be happening (at least as you are executing the first few comparisons) that will dominate the cost?

What could you do, in the preprocessing especially, that would alleviate (most of) this problem? What would be the effect?

Your answer need only be a paragraph or two.