

CS 5050 Final 99 Points

This is an open book exam. This means you can use the internet, books, software, and personal notes to produce the answers. Think of it like a set of problems you must solve at work. There are limits though. The work you submit must be your own, so no collaboration with any other person (including but not limited to people in the class). No copy and pasting of any internet content. By submitting the completed exam, you are affirming that you have followed these rules.

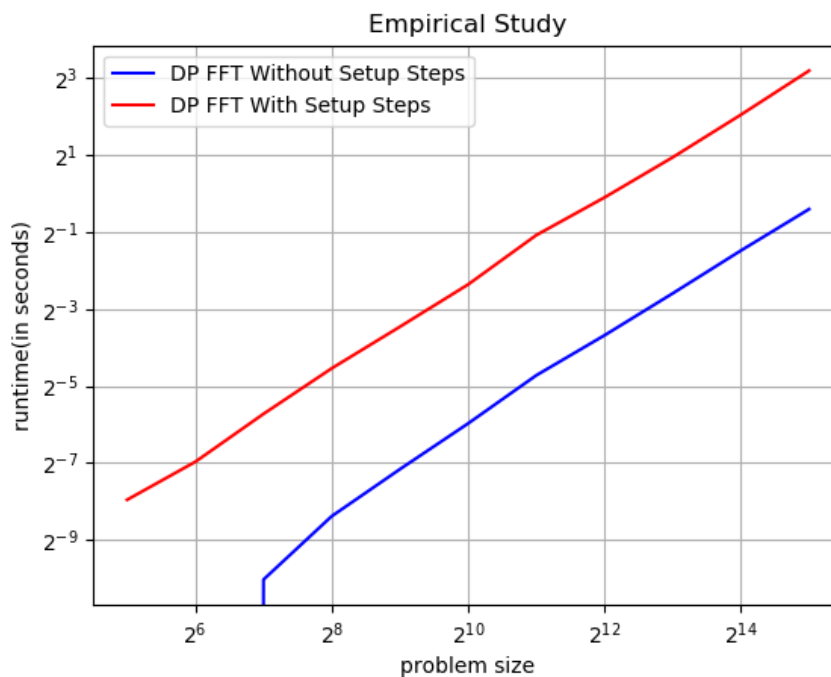
You must submit your completed exam as one pdf file. You may enter your answers by editing this .doc file or printing it, writing on it and then scanning in. Or some combination of the two. Some of the answers require writing some code. For these, the best way would be to do a screen snip that shows the new code inserted into any existing code and/or as new functions. Add a few comments so I can understand what you did. The advantage of the snip is that the formatting will be preserved so the code will be easier to read.

For this last “assignment” I will not accept late work.

Please keep your answers concise. There is no need to give me an essay as an answer. There are 9 questions in this exam and they are each worth 11 points.

Q8) Many students were surprised to find that the iterative version of the FFT was not clearly faster than the recursive FFT. One reason is that the following steps are repeated each time: a) allocating the solution array, computing the omega tables, and computing the reverse-bit-shuffle to fill in the base cases.

In general, these steps may be done once only and used every time from there on. This is called amortization, where the time taken is spread out over all future runs of the algorithm. Do the following: Rerun your iterative code, not counting these three setup steps and compare the results. Show a graph of the two versions of the code, one counting the setup steps, the other without. What is the speedup?



As shown in the graph, without the setup steps the dp algorithm runs about 11 times faster.

Q9) COVID-19 is a new virus that has mutated from previously observed viruses like SARS. COVID-19 continues to mutate as it spreads through the population. One of our assignments was to determine the mutation “distance” between two DNA sequences and identify the most likely mutations that occurred to change one sequence to the other. Do some research and identify and briefly describe three uses of the DNA alignment algorithm that is helping scientists understand COVID-19 and help find treatments to save lives.

- ChemRxiv uses DNA sequence alignment algorithm to compare the sequences and homology modeling of COVID-19 with other viruses Coronaviridae family. This study allows them to analyze the receptor binding sites of the samples to check for possible bonding with certain ligands with anti-viral properties. Knowing this will help with creating a vaccine.
- Because COVID-19 is an RNA virus, like the flu, it is more prone to mutation which makes it difficult to make a vaccine. Dr. John Rose from the department of pathology at Yale Medicine is using the DNA Alignment algorithm to track how different the current virus spreading in the US from the original samples being studied from China. He has found that COVID-19 is mutating very slowly and when it does mutate the changes are very small.
- Nextstrain uses the DNA alignment algorithm to track how quickly COVID-19 is mutating as well as where in the genome the mutations occur. This helps with drug development so that treatments for the illness will not have adverse effects.