

SP HW4 Report

Bxxxxxxx XXX

1. Thread Creation and Job Distribution

Threads were created before the training and testing steps. After each step, all threads, except for the main thread, would terminate.

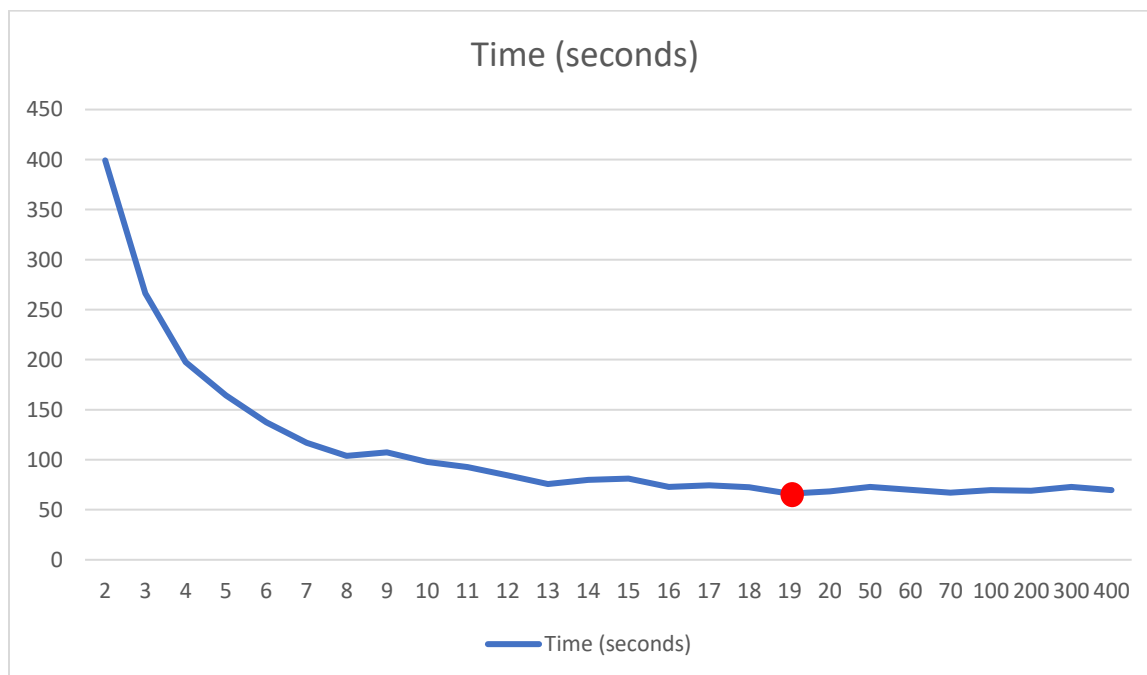
In the training process, each thread handled the creation of a tree. A thread would check if there were any trees left to create, and if so, would plant a tree and store its root into an array. If it discovered that there were no trees left to create, it would terminate itself.

In the testing process, each thread handled the prediction of an input. Similarly, it would first check if there were any more inputs left to process, and exit if there were none left.

2. Thread Count vs Time

Significant values that were held constant include: the number of trees planted (6000) and the amount of entries taken during the planting of each tree (128).

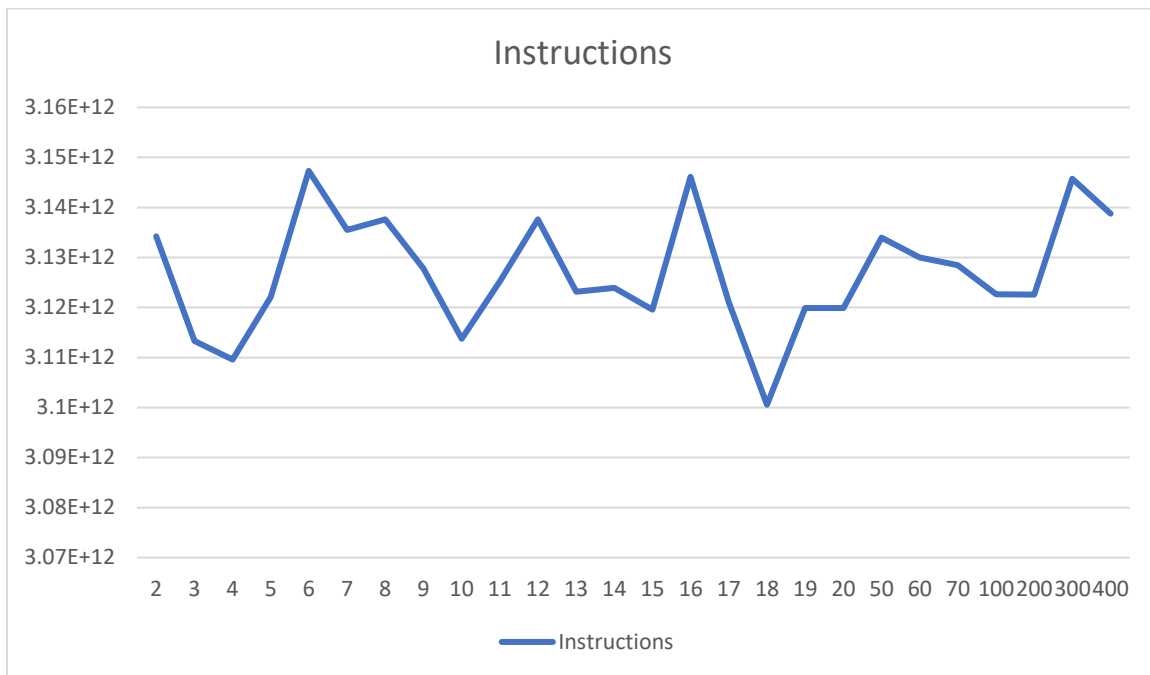
Note that this was tested on a workstation server with low loading capacity. To make the differences in time more noticeable, this report set the number of trees planted to 6000. The actual submission file lowers the number of trees planted to 2000, without any difference to the accuracy value.



At less than 5 threads, the computation time exceeded 3 minutes. The number then stabilizes after 15 threads, at which adding 10, 20, or even 300 makes little difference.

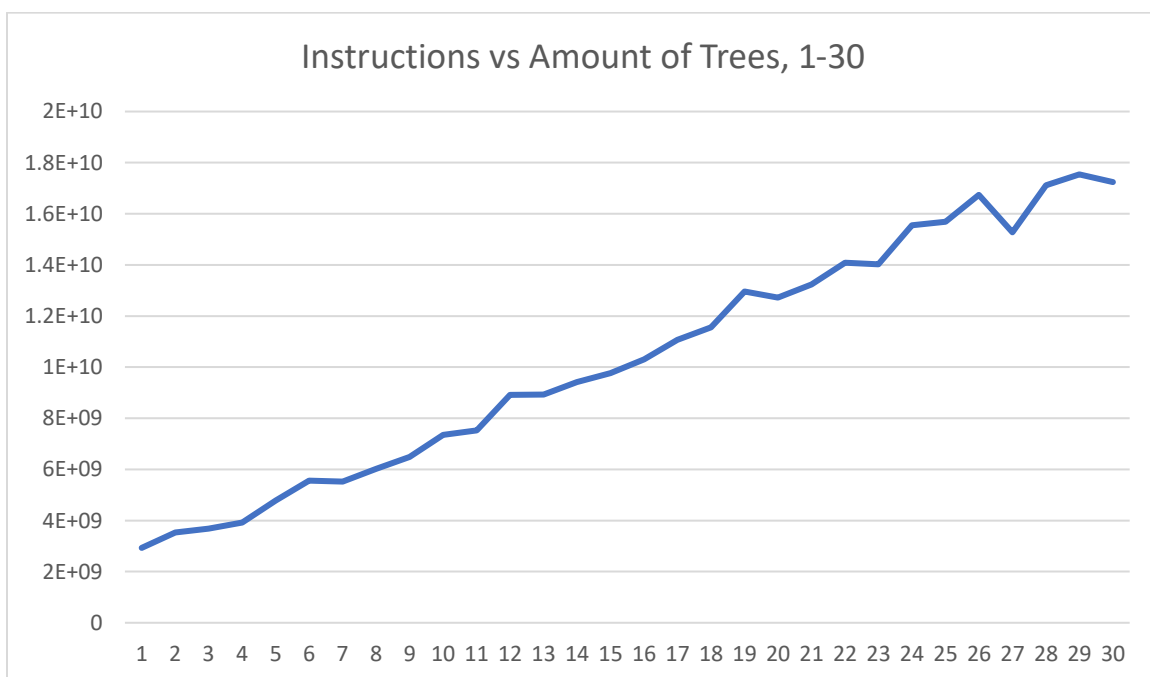
The minimum is at 19 threads, with a time of 66.1 seconds.

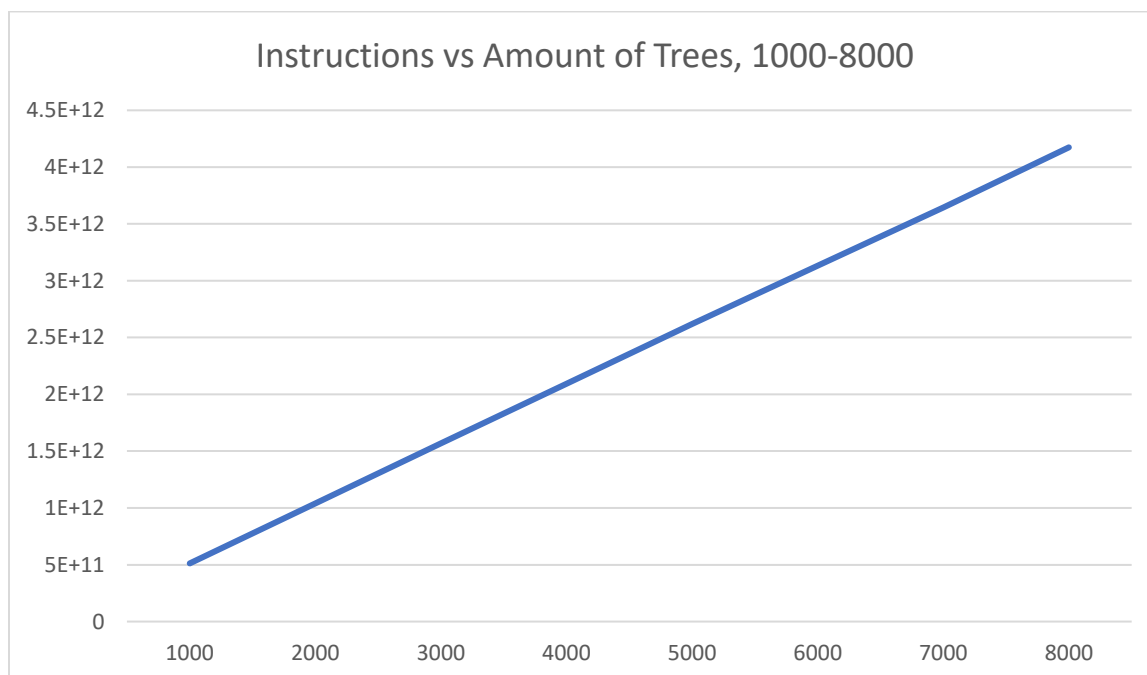
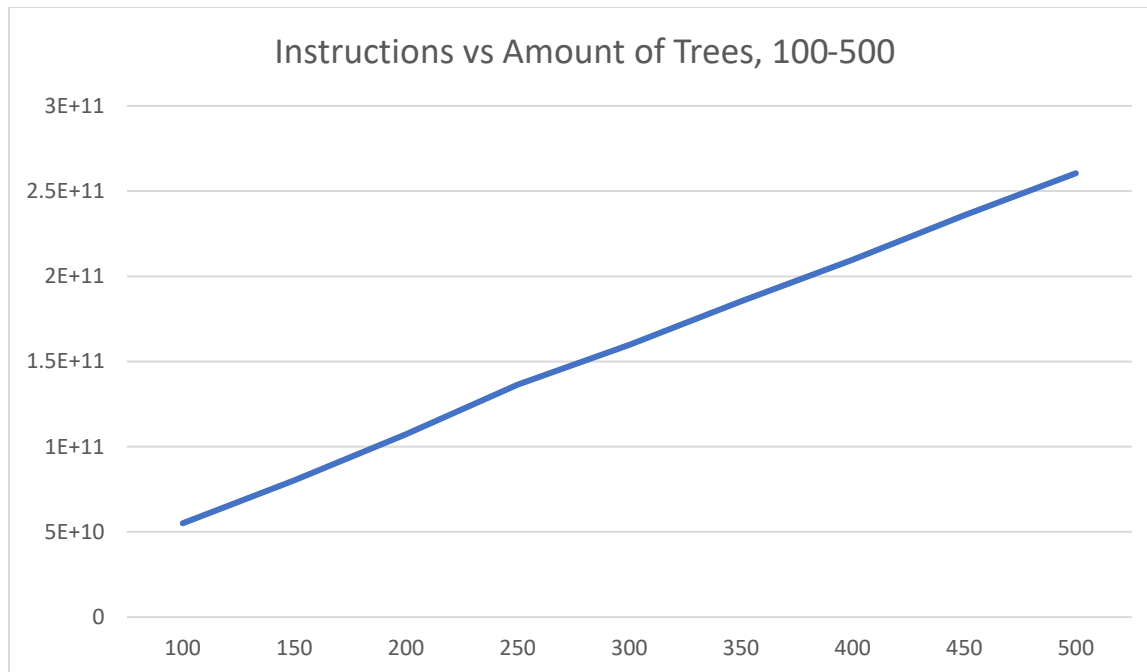
3. Thread Count vs Instructions Count



It seems that the amount of threads used made little difference to the number of instructions used. Like the previous section, the number of trees were fixed to 8000 and the amount of entries taken remained at 128.

4. Tree Count vs Instructions Count





The amount of entries taken during the planting of each tree and the number of threads remained constant at 128 and 20, respectively.

Based on the results, it seems that there is a direct correlation between the number of trees planted and the number of instructions called.

5. Other Discoveries

-Any amount of trees higher than 2000 did not improve the accuracy rate, which always stabilized at around 89%.

-There was also little difference in accuracy rate when the amount of entries taken to plant a tree rose from 50 to 128. In fact, it only increased the running time. It was only when the number dipped to below 20 did the accuracy rate take a significant hit.

-The efficiency of threads varied greatly between systems, especially those with different amounts of parallelization.

-On the workstation, creating more than 500 threads would cause the system to program to never finish. It would hang once all the trees were planted.