P5 Analysis

Henry Barklam (hgb13)

Marcus Deans (md374)

CS201: Data Structures and Algorithms

30 March 2020

We have adhered to the Duke Community Standard in completing this assignment.

Copy/Paste from running PercolationStats with these grid sizes:

grid sizes of 100, 200, 400, 800, 1600, and 3200

and using 20 trials

PercolationDFSFast (20 trials) Memory overload.

|  |  |  |  |
| --- | --- | --- | --- |
| Grid | Mean | Stddev | Total time |
| 100 | 0.593 | 0.014 | 0.083 |
| 200 | 0.591 | 0.010 | 0.106 |
| 400 | 0.590 | 0.006 | 0.805 |
| 800 | 0.594 | 0.004 | 4.462 |

PercolationBFS (20 trials)

|  |  |  |  |
| --- | --- | --- | --- |
| Grid | Mean | Stddev | Total time |
| 100 | 0.593 | 0.014 | 0.100 |
| 200 | 0.591 | 0.010 | 0.164 |
| 400 | 0.590 | 0.006 | 1.024 |
| 800 | 0.594 | 0.004 | 4.453 |
| 1600 | 0.592 | 0.002 | 23.720 |
| 3200 | 0.593 | 0.001 | 146.557 |

PercolationUF with QuickUWPC (20 trials)

|  |  |  |  |
| --- | --- | --- | --- |
| Grid | Mean | Stddev | Total time |
| 100 | 0.593 | 0.014 | 0.085 |
| 200 | 0.591 | 0.010 | 0.133 |
| 400 | 0.590 | 0.006 | 1.678 |
| 800 | 0.594 | 0.004 | 3.604 |
| 1600 | 0.592 | 0.002 | 14.436 |
| 3200 | 0.593 | 0.001 | 62.569 |

Answer these questions for PercolateUF with a QuickUWPC union-find object

1. How does doubling the grid size affect running time (keeping # trials fixed)

Doubling the grid size affects runtime at a quadratic rate, so the increase in runtime is greater than the increase in grid size. This is modeled by the graph and equation shown above.

1. How does doubling the number of trials affect running time.

PercolationUF with QuickUWPC (10 trials)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grid | Mean | Stddev | Total time | 20 trial times |
| 100 | 0.593 | 0.019 | 0.060 | 0.085 |
| 200 | 0.596 | 0.06 | 0.092 | 0.133 |
| 400 | 0.592 | 0.006 | 0.480 | 1.678 |
| 800 | 0.592 | 0.004 | 1.524 | 3.604 |
| 1600 | 0.594 | 0.002 | 6.479 | 14.436 |
| 3200 | 0.593 | 0.001 | 30.516 | 62.569 |

Doubling the number of trials increases runtime by approximately a factor of 2.5 for these grid sizes. With smaller grid sizes, doubling the number of trials has less effect and the runtime does not double. With larger grid sizes, doubling the umber of trials more than doubles the runtime. In any event, it can be shown that doubling the number of trials increases the runtime by a statistically significant amount.

1. Estimate the largest grid size you can run in 24 hours with 20 trials. Explain your reasoning.

This equation was obtained using the earlier data with 20 trials as can be seen on the graph:

Runtime = 7E-06x2 - 0.0015x + 0.4951 where x is the grid size, and runtime is in seconds.

There are 86400 seconds in a 24-hour period. This value can be substituted into the equation and the quadratic equation may be solved.

Therefore, the largest grid size that may be run in a 24 hour period is 111205 by 111205.