**CS 4350-5350**

**High Performance Computing with Embedded Systems**

**Python and Slurm Process Timing Analysis**

**(500 pts.)**

Date: 12-3-2020

Due: 12-11-2020 (5 pm)

**Given:** Three independent processes: Random Forest, Mass-Spring-Damper System (MSD System) , and SIR Epidemic Model

**1. Find: (400 pts.)**

Utilizing your knowledge timing methods in Python, analyze the computational time of three separate processes. Each process should have a minimum amount of computational complexity to make it useful for analysis (Random Forest > 500 trees, SMD System > 100 nodes, and SIR = 2 Coupled Compartments) . Your method should include functions for each process. You should be able to call a process from within a main function in your Python code. The ability to specify the number of times to run the process internally to calculate a mean value should be an argument you pass through. You also need to be able to specify how many times you will run this process to capture mean values for a final vector of mean value times for that particular experiment.

***Example ( Process\_1([input] no\_of\_trials\_for\_process, [input]no\_mean\_values\_to\_ x\_mean\_i = [x\_mean\_0, x\_mean\_1, ... ,x\_mean\_9]\_i'  and this will create a matrix, X\_mean = [x\_mean\_0, x\_mean\_1, ..., x\_mean\_9]  for a size X\_mean = 10 x 10.***

You will do this for three separate processes: Process\_1(), Process\_2, and Process\_3() to generate three data arrays: X\_mean\_P1, X\_mean\_P2, and X\_mean\_P3 . These should all be size=100x10. Now

**a)** Create scatter plots for your 10 experiments with 10 mean values for processes 1 through 3. You should use the subplot function to create a row of three plots for the three processes. Label these appropriately (x-axis[ trial no] and y-axis [time, in seconds]). Remember, you will have 10 data points per experiment; so, you will have 10 curves with 10 points per curve on your scatter plot for each process.

**b)** Utilizing your three process arrays, create three heat maps of the data arrays (ie. a heat map is creating an image from the data in the array. You can easily normalize each array [0 to 1] and scale by 255. Convert those values into uint8() and save the image as a grayscale image. Do this for each data array (X\_mean\_P1, X\_mean\_P2 and X\_mean\_P3). Using your subplot command plot these in a 1x3 format.

**c)** Calculate standard deviation and mean for each column in your original data X\_mean\_mean\_i and std\_i for i=0 to 9. Do this for all three processes. Plot your mean values (X\_mean\_mean\_i) for all three processes together and do the same for the standard deviations (std\_i). Use subplot to display these in a 1x2 array with proper labeling. Turn in all source code, plots and data with a Word document displaying results (**a**) through (**c**).

**2. Find: (100 pts.)** Utilizing your knowledge of Slurm, analyze the computational time (in theory only) of three separate processes. Each process should have a minimum amount of computational complexity to make it useful for analysis (Random Forest > 500 trees, MSD System > 100 nodes, and SIR = 2 Coupled Compartments).

**a)** Write a Slurm batch file to run all three processes simultaneously to three separate CPU cores. Write a command line **sinfo** to glean important information on the run time for each process.

**b)** Write a slurm batch file that allows you to pass the number of trees through to your Random Forest program from 100-1000 in increments of 100.

**c)** Write a Slurm batch file to vary the number of nodes (N) and damping coefficient (C\_d) of the MSD system for N from 10 to 100 by increments of 10 and C\_d from 0 to 1 in increments of 0.1.

**d)** Write a Slurm batch file to vary B and g in your SIR model, B from 0.0 to 0.5 in increments of 0.05 and g from 0.0 to 0.33333 in increments of 0.03333. (**a** through **d** should utilize 4 CPU cores for run-time).

**Extra Credit (100 pts.)** Run one of the above batch files in Slurm (**a-d**) and report your results.

**References:**

*https://www.youtube.com/watch?v=K\_JIPrcPHCg*

*https://www.youtube.com/watch?v=8N8gb4BSu\_4*

*https://www.youtube.com/watch?v=gg2X6NCor8E*

*https://www.youtube.com/watch?v=RbUy3JZJw24&t=1s*

*https://www.youtube.com/watch?v=0Rj\_xNuyOyQ*

**Additional Resources:**

https://www.youtube.com/watch?v=7zJUceJiYxQ

https://www.youtube.com/watch?v=LRJMQO7Ercw&t=8s

https://www.youtube.com/watch?v=grZVs1UUxug&t=696s