Henry Yu UID: 303 889 334

Lab 5 Report

1.
    a. The item collection include the alignment_score_matrix, which contains how the scores are calculated, string_1, and string_2 which are the character strings we are comparing, the matrix score which contains the answer, and file_name_1 , file_name_2 which are the inputs.
    b. The step collections are read_file1 and read_file2 which contains the way to read the file and outputs it as string_1 and string_2. There is also score_gen which generates the first column and first row of the score matrix. Finally there is the smith_waterman which makes and completes the final output of score.
    c. The dependencies between the 4 step collections are that read_file1 and read_file2 can be ran in parallel but both must be completed before score_gen is started. Finally, score_gen must be completed before we start on the smith_waterman algorithm. The communications between the step collections include sending the output of read_file1 and read_file2 to the smith_waterman step, and sending score_gen to the smith_waterman step. Also, each process that has to come in a certain order must receive an input that tells them that the operation before was successful. This solution is the most efficient because each process is running when it can.
2. To implement the main function, it would consist of initializing the item collections and calling the step collections. The kernel would consist of the implementation of each step, which would basically contain parts of the code that was divided by the step function.
3. The advantage of using CnC over using MPI is that it contains when something needs to be execute and in what order. The user doesn't needt o specify what needs to be ran in parallel. However, the disadvantage to CnC is that the user can't specify what should be ran in parallel.