CS 566 Parallel Processing, Spring 2006

Programming Assignment 2
Assigned: March 27
Due: May 1, 11:59pm

1 Goals

In this assignment, you will study various aspects of parallelization of *dynamic programming* using MPI, in the context of one of these problems. As for Assignment 1, you can work in pairs.

1. **Longest Common Subsequence:** for Amino-Acid sequences. Read Section 12.3.1.
2. **Optimal Matrix Parenthesization Problem:** Read Section 12.5.1.
3. **Any other problem of your choice involving dynamic programming.** Please get it approved by me before you begin. The problem must be at least as difficult as the above two problems.

Unlike Assignment 1, you now have to define the experiment to study parallel programming technqiues on the problem from the list above. As a general thumb rule, you can use the structure of steps for Assignment 1. The following are some of the main steps you will probably need.

1. Decide which algorithm formulations to use.
2. Decide which MPI primitives you will use.
3. Decide the parameters of interest, and what ranges to test them over.
4. Decide how to provide your input, and how to collect the output.
5. Decide how you will test your program.
6. **Report.** This is the most important component that will document your experiment, and all your readings (via graphs, tables etc.) and observations. As before, report on the computaiton time, efficiency, speedup.

2 On Completion

1. Put all your files including Makefile in directory `assignment2`. Turnin all the code, including the makefile, by executing:
   `turnin -c cs566 -p assignment2 assignment2`
2. Submit a hard copy of all the files including the code file.
3. Submit a hard-copy report describing your experiment. The report must have the following sections/information.

(a) **Formulations:** Describe each of the formulations. Include the list of MPI calls you used in the implementation of each formulation.

(b) **Parameter ranges:** For each of the formulations describe the range of parameters you experimented with.

(c) **Results:** Give the tables and plot the graphs showing the *timing*, *speedup*, and *efficiency* variations for each of the formulations, for the range of parameters you used. Remember to use appropriate scales for the graphs.

(d) **Analysis:** Analyse the results. Give your interpretation and reading of the results for each formulation and across formulations, i.e., also compare the results of the different formulations. You should address questions such as the following sample list:

- Which formulation is better? under what (or all) circumstances? Why?
- How and why does efficiency vary the way it does?
- How and why does speedup vary the way it does?
- Can you measure the computation time versus communication time? How do the formulations compare with each other with respect to this breakup of the time overhead?
- Can the communication overhead be overlapped with your computation overhead? Do your formulations/algorithms do so?

Explain why you observe what you observe. In particular, any anomalous observations should be explained.

(e) **Lessons:** What insights you learned from this assignment.

3 Grading

The problem is reasonably well-formulated but the experimental approach is very open-ended. Your goal is to get the most parallelism in solving the problem you have chosen. The depth of your choice of parallel formulations you choose to experiment with will be judged. Your assignment will also be judged on how *comprehensively* and *methodically* you have designed and run the experiment, and reported on the results. Your insights into the analysis of the results will also be considered in judging the assignment.

4 Reference Material/Chapters

Chapter 12 of the text book.