Embedded System Design and Modeling
ECE382N.23, Fall 2024

Homework #4
System Mapping & Synthesis

|	
Assigned:	October 24, 2024
Due:	November 6, 2024

Instructions:

- Please submit your assignment via Canvas. Submissions should include a single PDF with the writeup and a single Zip or Tar archive for any supplementary files (e.g., source files, which has to be compilable by simply running 'make' and should include a README with instructions for running each model).
- You may discuss the problems with your classmates but make sure to submit your own independent and individual solutions.
- Some questions might not have a clearly correct or wrong answer. In general, grading is based on your arguments and reasoning for arriving at a solution.

Problem 4.1: Task Mapping

Given the following task graph from Problem 1.2 in Homework 1 of a 2-layer neural network in which the first and second layers are partitioned into 4 and 2 tiles, respectively. Explore mappings of this graph onto a heterogeneous System-on-Chip (SoC) platform consisting of a dual-core CPU and a GPU with the following task execution times ('-' means that the task can not be mapped onto the respective processor type):
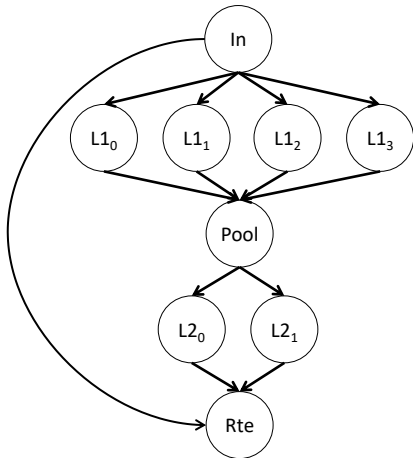
	CPU	GPU
In	1	-
L1	5	1
Pool	1	-
L2	10	2
Rte	1	-

(a) Apply a list scheduling algorithm to map and schedule one iteration of the graph onto the SoC platform. Use the maximum distance to the sink as priority function for the list scheduler. Show the step-by-step operation of the algorithm (time, state of the ready list with priorities and mapping decisions made in each step), as well as the final schedule. What is the latency of your schedule?

(b) Does a list scheduler always give a schedule with minimal latency? Is there a schedule for this graph with smaller latency than what is achieved by the list scheduler? If so, show the schedule and minimal latency.

(c) Now schedule the graph in a pipelined fashion. What is the highest throughput that can be achieved? Show a schedule that achieves this throughput while minimizing latency. What is the latency of one iteration of the graph in your schedule?

(d) **(extra credit)** Sketch the ILP formulation for the non-pipelined mapping (combined partitioning and scheduling) of the problem graph on the SoC platform. Limit the optimization problem to a basic (non-pipelined) schedule of a single iteration of the graph in the time window $0 \le t < T_{max}$. List all the inputs to your ILP and all constraints for unique mapping of tasks to processors, sequential execution on each processor and sequencing/dependency relations between actors. Formulate an objective to minimize overall latency (time to execute the single iteration of the graph).

Problem 4.2: Reading Assignment

Read the following paper and submit a written, including brief (executive) summary, 5 strengths, 5 weaknesses, and detailed comments/justification:

A. Deshwal, N. K. Jayakodi, B. K. Joardar, J. R. Doppa, P. P. Pande, "MOOS: A Multi-Objective Design Space Exploration and Optimization Framework for NoC Enabled Manycore Systems, *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 18, no. 5s, pp. 77:1-77:23, October 2019.