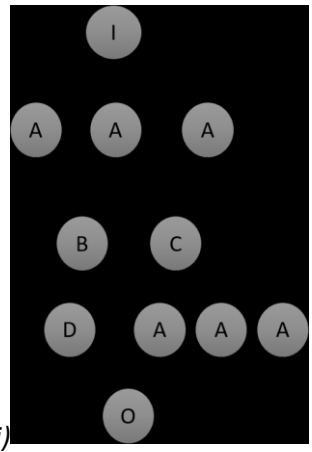
BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI (RAJ.)
II SEMESTER 2022-23
Mid Semester Test (OPEN BOOK)

Course No.: CS F422 **Course Title:** Parallel Computing
Date: 16th March (2-3:30) **Maximum Marks:** 25% (25M)

Note:
- Overwritten answers will not be accepted for rechecks
- Write all parts of a question together.

Q1. Consider a data-dependence graph given in (i). Identify all sources of parallelism. Compute average degree of concurrency assuming each node is of unit time. For the code given in (ii), draw a data dependency graph and identify sources of parallelism. [4M]

(i)



(ii)

```
double sqrt_newton (double n)
{
  int no_itr=10000,i;
  double x = 1;
  for (i=0;i<no_itr;i++)
    {
    double nx = (x+n/x)/2;
    x = nx;
    }
  return x;
}
```

Q2. Consider the following sequential algorithm for sorting a list of data elements. [6M]

```
void sort(
      int  a[] ,
      int  n   ) {
   int phase, i, temp;

   for (phase = 0; phase < n; phase++)
      if (phase % 2 == 0) {
         for (i = 1; i < n; i += 2)
            if (a[i-1] > a[i]) {
               temp = a[i];
               a[i] = a[i-1];
               a[i-1] = temp;
            }
      } else {
         for (i = 1; i < n-1; i += 2)
            if (a[i] > a[i+1]) {
               temp = a[i];
               a[i] = a[i+1];
               a[i+1] = temp;
            }
      }
}
```

(a) Draw a data dependency graph for the code given. Identify opportunities for parallelism.
(b) Implement a parallel version of this algorithm using Pthreads.
(c) Compute speedup, cost and efficiency of parallel algorithm. Are there any parallel overheads?
(d) Determine whether your parallel algorithm is cost-optimal. What are the minimum number of threads that ensure cost-optimality?
(e) Calculate the required increase in problem size to maintain the efficiency given that one more processor is added.

Q3. Briefly answer the following.

(a) Using a diagram, explain why log barrier complexity is better than linear barrier in Pthreads?

(b) In which case, guided scheduling class performs better than dynamic scheduling class in OpenMP loop parallelization. Give an example.

(c) Compute the extra time (asymptotically) required for a p-processor EREW PRAM to simulate a p-processor ARBITRARY PRAM?

(d) "False sharing is not possible if all processors in a shared memory system use different bytes in a single cache line". State true/false with justification.

(e) "The time complexity for parallel reduction on EREW PRAM is $\Theta(\log n)$ if $p = n/(\log n)^2$ where n is the number of data elements". State true/false with justification.

[5*1=5M]

Q4. The speedups of parallel program I to III are given below. [4M]

	Speedup		
Processors	Prog I	Prog II	Prog III
1	1	1	1
2	1.67	1.89	1.89
3	2.14	2.63	2.68
4	2.5	3.23	3.39
5	2.78	3.68	4.03
6	3	4	4.62
7	3.18	4.22	5.15
8	3.33	4.35	5.63

(a) Compute speed up and scaled speed up for 16 processors for all three programs.

(b) Explain the nature of parallel overhead in each of these programs.

Q5. Consider a Twitter application which generates data in the following format *<UserID, Follower ID>*. Sample data is shown in the following table. Assume that there are no duplicate records in the file.

UserID	FollowerID
A	B
A	K
B	X
B	Y
C	X
D	K
A	X

User *u1* is more popular than *u2* if *k1>k2* where *k1* and *k2* are number of followers of *u1* and *u2* respectively

We need to find the top-10 most popular users on twitter.

(a) Design and implement a parallel version using MPI.

(b) Calculate time complexity, speedup, cost, and efficiency for the parallel version.

(c) Design Hadoop map-reduce application with pseudo code for map() and reduce() tasks.

[6M]

--&--