PUNE INSTITUTE OF COMPUTER TECHNOLOGY

DHANKAWADI, PUNE –43

LIST OF LAB ASSIGNMENTS

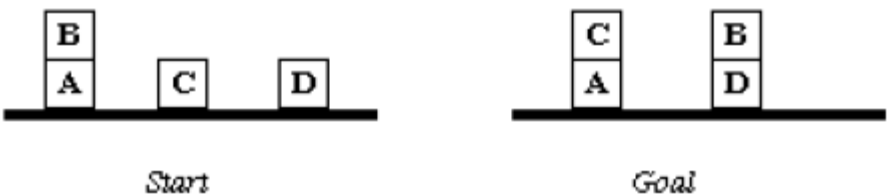# ACADEMIC YEAR: 2020- 2021

Department: Computer Engineering.   Date: 09/07/2020

Class: B.E Semester: I

Subject: Laboratory Practice I (410246)   Examination scheme:
                                                          TW-50, PR-50

Group A: 410241:High Performance Computing		
Four assignments individually and any one mini-project with group of 2-3 students.		
Sr.No.	Assign. No.	**Problem Statement**
1	A1	a) Implement Parallel Reduction using Min, Max, Sum and Average operations. b) Write a CUDA program that, given an N-element vector, find- • The maximum element in the vector • The minimum element in the vector • The arithmetic mean of the vector • The standard deviation of the values in the vector Test for input N and generate a randomized vector V of length N (N should be large). The program should generate output as the two computed maximum values as well as the time taken to find each value.
2	A2	**Vector and Matrix Operations-** Design parallel algorithm to 1. Add two large vectors 2. Multiply Vector and Matrix 3. Multiply two N × N arrays using n^2 processors
3	A3	**Parallel Sorting Algorithms-** For Bubble Sort and Merger Sort, based on existing sequential algorithms,

		design and implement parallel algorithm utilizing all resources available.
4	A4	**Parallel Search Algorithm-** Design and implement parallel algorithm utilizing all resources available. for 1. **Binary Search for Sorted Array** 2. **Best-First Search that (** traversal of graph to reach a target in the shortest possible path)
	Group A: Sample Mini Projects	
1		**Compression Module (Image /Video)** Large amount of bandwidth is required for transmission or storage of images. This has driven the research area of image compression to develop parallel algorithms that compress images. **OR** For video: RGB To YUV Transform concurrently on many core GPU
2		**Generic Compression:**Run length encoding concurrently on many core GPU
3		**Encoding:**Huffman encoding concurrently on many core GPU
	Group B: 410242: Artificial Intelligence & Robotics	
	Four assignments individually and any one mini-project with group of 2-3 students.	
5	B1	Solve 8-puzzle problem using A* algorithm. Assume any initial configuration and define goal configuration clearly. OR Solve following 6-tiles problem stepwise using A* algorithm, **Initial Configuration** B W B W B W **Final configuration:** B B B W W W Constraint: Tiles can be shifted left or right 1 or 2 positions with cost 1 and 2

		respectively.
6	B2	Implement any one of the following Expert System: 1. Medical Diagnosis of 10 diseases based on adequate symptoms 2. Identifying birds of India based on characteristics OR Develop elementary chatbot for suggesting investment as per the customers need.
7	B3	Implement goal stack planning for the following configurations from the blocks world, OR Implement syntax analysis for the assertive English statements. The stages to be executed are, Sentence segmentation Word tokenization Part-of-speech/morpho syntactic tagging Syntactic parsing (Use any of the parser like Stanford)
8	B4	Constraint Satisfaction Problem: Implement crypt-arithmetic problem or n-queens or graph coloring problem (Branch and Bound and Backtracking) OR Use Heuristic Search Techniques to Implement Hill-Climbing Algorithm.
		Group C: 410243-Data Analytics
Four assignments individually and any one mini-project with group of 2-3 students.		
9	C1	Download the Iris flower dataset or any other dataset into a DataFrame. (eg https://archive.ics.uci.edu/ml/datasets/Iris) Use Python/R and Perform following

		– - How many features are there and what are their types (e.g., numeric, nominal)? - Compute and display summary statistics for each feature available in the dataset. (eg. minimum value, maximum value, mean, range, standard deviation, variance and percentiles - Data Visualization-Create a histogram for each feature in the dataset to illustrate the feature distributions. Plot each histogram. - Create a boxplot for each feature in the dataset. All of the boxplots should be combined into a single plot. Compare distributions and identify outliers.
10	C2	Download Pima Indians Diabetes dataset. Use Naive Bayes" Algorithm for classification Load the data from CSV file and split it into training and test datasets. Summarize the properties in the training dataset so that we can calculate probabilities and make predictions. Classify samples from a test dataset and a summarized training dataset.
11	C3	Bigmart Sales Analysis: For data comprising of transaction records of a sales store. The data has 8523 rows of 12 variables. **Predict the sales of a store.** Sample Test data set available here https://datahack.analyticsvidhya.com/contest/practice-problem-big-mart-sales-iii/
12	C4	Twitter Data Analysis: **Use** Twitter data for sentiment analysis. The dataset is 3MB in size and has 31,962 tweets. **Identify the tweets which are hate tweets and which are not. S**ample Test data set available here https://datahack.analyticsvidhya.com/contest/practice-problem-twitter-sentiment-analysis/
		Group C: Sample Miniprojects
	1	Use Movies Dataset. Write the map and reduce methods to determine the average ratings of movies. The input consists of a series of lines, each containing a movie number, user number, rating, and a timestamp: The map should emit movie number and list of rating, and reduce should return for each movie

		number a list of average rating.
	2	Time Series Analysis: Use time series and forecast traffic on a mode of transportation. Use at least two methods. Sample Test data set available here https://datahack.analyticsvidhya.com/contest/practice-problem-time-series-2/
	3	Trip History Analysis: Use trip history dataset that is from a bike sharing service in the United States. The data is provided quarter-wise from 2010 (Q4) onwards. Each file has 7 columns. Predict the class of user. Make use of at least two classification algorithms and provide comparative analysis. Sample Test data set available here https://www.capitalbikeshare.com/trip-history-data

Subject Coordinator   Head, Dept. of Computer Engg.
(Hemlata P. Channe)   (Prof.M.S.Takalikar)