

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY  
 B.Sc. Engineering 4<sup>th</sup> Year 2<sup>nd</sup> Term Examination, 2023  
 Department of Computer Science and Engineering  
 CSE 4221

Natural Language Processing

TIME: 3 hours

FULL MARKS: 210

- N.B. i) Answer **ANY THREE** questions from each section in separate scripts.  
 ii) Figures in the immediate right column of the questions indicate full marks.  
 iii) The rightmost column indicates the course outcomes.

**SECTION A**

(Answer **ANY THREE** questions from this section in Script A)

1. a) Define corpus. Write a regular expression to find an answer for the following statement. (08) [CO1]  
 "A computer with 6 GHz and 2 TB storage within 500\$".  
 b) What is lemmatization and stemming? How is lemmatization done? (06) [CO1]  
 c) How does Byte Pair Encoding (BPE) work for tokenization? Use the following training corpus to tokenize the test word "hugs". (14) [CO1]  
 "This is the hugging face course. It shows tokenizer algorithms".  
 d) Make a classification of the types of spelling errors and their solution approaches. (07) [CO1]
2. a) You have two strings  $St1$  and  $St2$ . Develop an algorithm to go from  $St1$  to  $St2$  with minimum cost of edit. Use insert, delete cost = 1 and update cost = 2. Explain your algorithm using example as  $St1$  = "INTENTION" and  $St2$  = "EXECUTION". (15) [CO3]  
 b) Define language model. "Predicting something is difficult". How does  $N$ -gram language model predict next word? Explain with suitable example. (14) [CO2]  
 c) How is language model evaluated? (06) [CO2]
3. a) Define text classification. "Naïve Bayes is a form of generative classifier" – justify the statement. (08) [CO3]  
 b) How does Naïve Bayes learn? Using information from the following table, find the class of the test sentence. (12) [CO3]

	Document	Class
Training	The movie is amazing	Positive
	I do not like the film	Negative
	Great acting and engaging	Positive
	Terrible movie and not recommended	Negative
Test	The acting is superb	?

- c) What is parsing in NLP? What are the differences between syntactic and semantic parsing? Explain. (05) [CO3]
- d) How does CYK algorithm work? For the following grammar  $G$ , is "baba" in  $L(G)$ ? (10) [CO3]

$$S \rightarrow AB|BC$$

$$A \rightarrow BA|a$$

$$B \rightarrow CC|b$$

$$C \rightarrow AB|a$$

4. a) "You have a true word  $w$  and a noisy channel that modifies  $w$  into  $w'$ ". How can you find the  $w$  from  $w'$ ? Explain. (10) [CO5]
- b) How can you define a word as a vector? How can you measure the similarity between two words  $u$  and  $v$ ? Show the similarity of the words from the following table. (15) [CO5]

	pie	data	computer
cherry	442	8	2
digital	5	1683	1670
information	5	3982	3325

- c) What is absolute discounting? Explain what are its advantages. (10) [CO5]

## SECTION B

(Answer **ANY THREE** questions from this section in Script B)

5. a) What is closed and open class of Parts of Speech (POS)? Explain with example. (05) [CO2]  
b) Explain different type of taggers and describe Hidden Markov Model (HMM) POS tagging using example. (10) [CO1]  
c) Explain Viterbi algorithm for POS tagging with example. (10) [CO1]  
d) How do conditional random fields enhance POS tagging through feature incorporation and why are features challenging to integrate into HMMs? (10) [CO1]
6. a) What is relation extraction? Explain this concept using pattern-based relation extraction and supervised learning algorithm. (10) [CO2]  
b) Explain semi-supervised relation extraction via bootstrapping. (08) [CO2]  
c) What are temporal lexical triggers? How do we detect events? (06) [CO2]  
d) How is the *tf-idf* score computed for terms in documents and how can it be used to calculate the cosine similarity between a query and documents using BM25? Demonstrate for the following dataset. (11) [CO5]  
Assume,  $|d| = 4$ ,  $|d_{avg}| = 3$ ,  $k = 1.3$ ,  $b = 0.75$ .  
Query: sweet love  
Doc 1: Sweet sweet nurse! Love?  
Doc 2: Sweet sorrow  
Doc 3: How sweet is love?  
Doc 4: Nurse!
7. a) Explain the evaluation of information retrieval systems and describe IR-based factoid question answering with a diagram. (08) [CO5]  
b) How can BERT-based models be effectively trained and evaluated for answer span extraction in information retrieval-based question answering systems? (09) [CO5]  
c) What is entity linking? Explain the process of entity mention detection using mathematical formulas and a diagram. (08) [CO5]  
d) What is knowledge-based question answering? Describe Watson's question answering system with the help of a diagram. (10) [CO5]
8. a) How does attention mechanism work in machine translation and why is it important for improving translation quality? (09) [CO4]  
b) Explain hourglass metaphor in text-to-speech with proper diagram. (08) [CO4]  
c) What is the difference between tokens and embeddings? Explain the pipeline of creating embedding model. (08) [CO4]  
d) Why is beam search necessary? Explain beam search with example. (10) [CO2]

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY  
B.Sc. Engineering 4<sup>th</sup> Year 2<sup>nd</sup> Term Examination, 2023  
Department of Computer Science and Engineering  
CSE 4223  
Digital System Design

TIME: 3 hours

FULL MARKS: 210

- N.B. i) Answer **ANY THREE** questions from each section in separate scripts.  
ii) Figures in the immediate right column of the questions indicate full marks.  
iii) The rightmost column indicates the course outcomes.

**SECTION A**

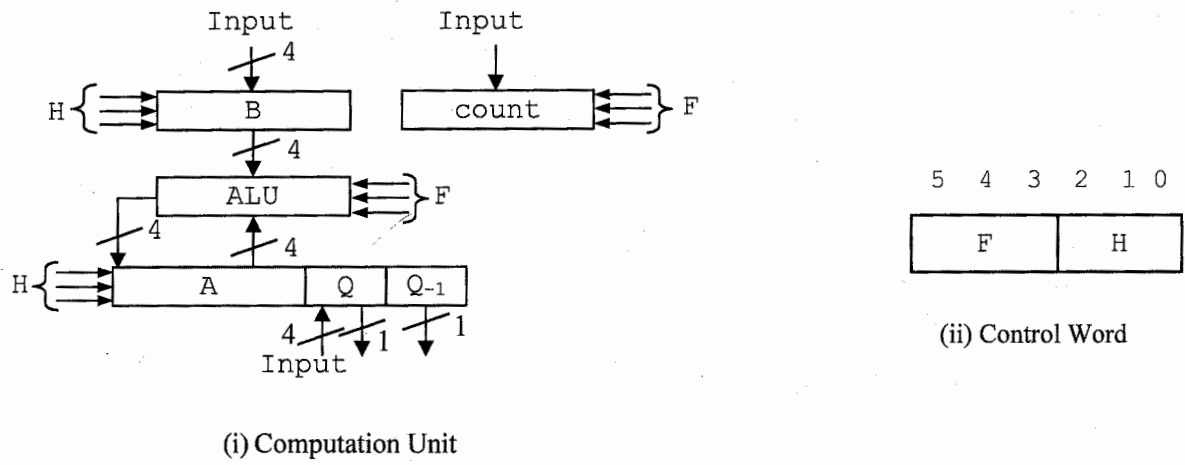
(Answer **ANY THREE** questions from this section in Script A)

1. a) "A counter is a special kind of register" – explain. Design a 5-bit controlled buffer register with a proper diagram. (15) [CO1]  
b) Why is ripple counter too slow for application? "A ring counter does not perform ordinary counting" – explain. Write some applications of ring counter. (10) [CO2]  
c) What is tri-state switch? Explain the operation of tri-state switch with an appropriate diagram. (10) [CO2]
2. a) Design a  $16 \times 4$  three-state diode ROM with on-chip decoding. Assume the data values for your design. (10) [CO1]  
b) Differentiate between static and dynamic RAM. Explain the operation of static and dynamic cells. (15) [CO2]  
c) "A 2-port memory organization is cheaper and faster than bus organization." Is it true? Write your opinion with proper explanation. (10) [CO2]
3. a) Draw SAP-1 architecture and explain each block in brief. (12) [CO2]  
b) Write an SAP-1 program for the expression:  $30 - 12 + 15 - 20$ . Then, convert it to machine code. (10) [CO1]  
c) Generate control word and draw timing diagram for SUB and OUT instructions of SAP-1. (13) [CO2]
4. a) What is register transfer logic? Write down the usual symbols used in register transfer logic. Also, provide example of each case. (10) [CO2]  
b) "A register can be represented in various ways in register transfer logic" – explain. What is instruction code? Give example(s). (10) [CO1]  
c) "In register transfer logic '+' symbol has two meanings" – justify it. (07) [CO1]  
d) What is FPGA? Write some applications of FPGA. (08) [CO3]

**SECTION B**

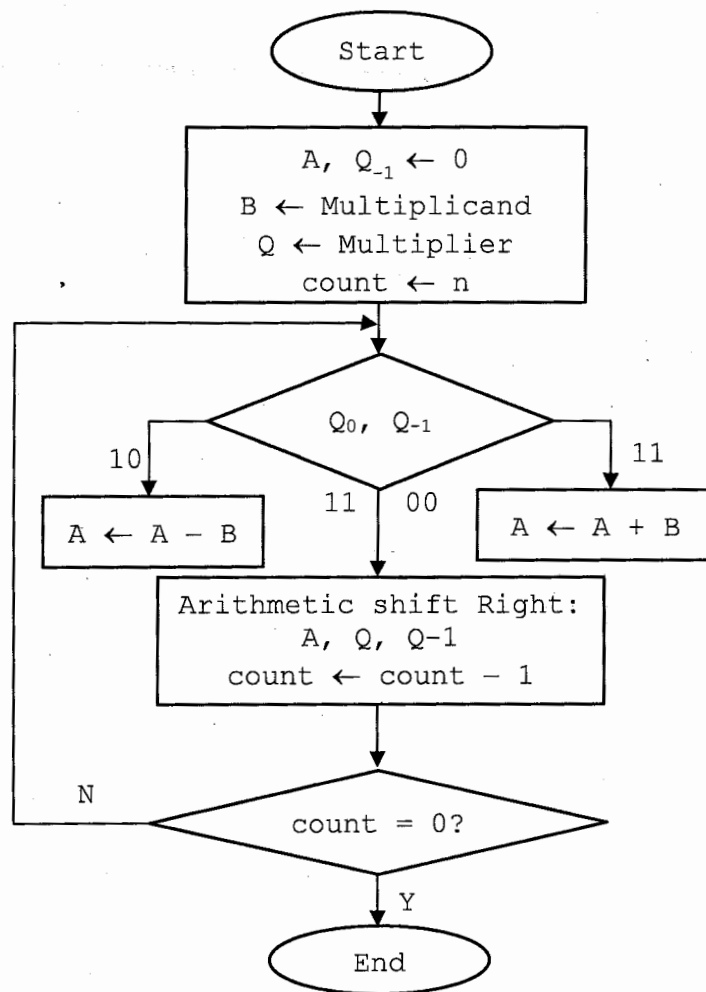
(Answer **ANY THREE** questions from this section in Script B)

5. a) Write the significance of control logic in any digital system. Briefly describe different methods of control logic design. (08) [CO2]  
b) You are given a system block diagram and flowchart with a function table of ALU shown in the Figure 5(b). Now, design a control unit using a sequence register and decoder method that multiplies two numbers using Booth's Algorithm, which is given in the flowchart (Figure 5(b)). (17) [CO2]  
c) Explain the relationship between control logic and data processor in a digital system using a block diagram. (10) [CO1]
6. a) Register-transfer operations with a 2-bit sequence register  $G$  are  $x'T_0 : G \leftarrow 00$ ;  $xT_0 : G \leftarrow 01$ ;  $y(T_1 + T_2) : G \leftarrow G + 1$ ;  $y'(T_2 + T_3) : G \leftarrow G - 1$ ;  $yT_3 : G \leftarrow 00$ . Draw the state diagram and design the control system with T-flipflop and decoder. Also, write a Verilog module for the control system. (20) [CO4]  
b) Let, you have implemented a PLA control unit. The total number of control bit outputs are 13. Among the 13 bits, 3 bits are connected to the sequence register. There are four external variables that determine or manipulate the workflow of the control unit. Now, you want to replace the PLA with a ROM. What will be the size of the ROM? Justify your choice. (10) [CO2]  
c) What are the similarities and differences between a PLA control and a control logic with sequence register and decoder? (05) [CO2]



Binary Code	Function of Selection Variable	
	F	H
000	count $\leftarrow$ input	A, Q, Q-1, B $\leftarrow$ 0
001	A+1	A, Q-1 $\leftarrow$ 0 B, Q $\leftarrow$ input
010	A+B	Shr: A, Q, Q-1
011	A+B+1	Shl: A, Q, Q-1
100	A-B-1	Arithmetic shift Right: A, Q, Q-1
101	A-B	Arithmetic shift Left: A, Q, Q-1
110	A-1	-
111	count-1	-

(iii) Function table



(iv) System Flowchart

Figure 5(b).

7. a) An instruction in address  $(0AC)_{16}$  in the computer has the operation-code of ADD (15) [CO1] instruction and an address part is  $(1CD)_{16}$ . The memory word of address  $(1CD)_{16}$  contains  $(1CBF)_{16}$ . Register A contains  $(1010)_{16}$ . Tabulate the contents of register PC, MAR, B, A and I after the instruction is executed. Repeat the problem two more times for operations AND and STO.
- b) A small computer only has memory reference instructions in the following table. (20) [CO1] Now, list the register-transfer statements for execution of memory reference instructions in case of fetch and execution cycle.

Table: Memory Reference Instructions

Symbol	Hexa-decimal code	Description	Function
AND	0 m	And to A	$A \leftarrow A \wedge M$
ADD	1 m	Add to A	$A \leftarrow A + M, E \leftarrow Cout$
STO	2 m	Store in A	$M \leftarrow A$
ISZ	3 m	Increment and skip if zero	$M \leftarrow M + 1, \text{ if } (M + 1 = 0) \text{ then } (pc \leftarrow pc + 1)$
BSB	4 m	Branch to Subroutine	$M \leftarrow pc + 5000, pc \leftarrow m + 1$
BUN	5 m	Branch unconditionally	$pc \leftarrow m$
LDA	6 m	Load memory Data to A	$A \leftarrow M$
HALT	7 m	Stop computer	$S \leftarrow 0$

8. a) Suppose, you need to design a digital system which needs 14 individual instructions (10) [CO2] and its accumulator size is 16. Calculate the sizes of other registers and memory, if maximum allocated memory size is 2048 words.
- b) The following program is a list of instructions in hexadecimal code. The computer (20) [CO1] executes the instructions starting from hexadecimal location 109.
- Write the program in symbolic form. Note that the last two values are operands.
  - Determine the contents of register A after each instruction being executed.

Location	Instruction
109	6800
110	1115
111	6200
112	6020
113	1116
114	6001
115	0A3F
116	00A5

- c) Explain the way for a small computer to synchronize its own action with an external (05) [CO2] I/O device (i.e., read-write operations).

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY  
B.Sc. Engineering 4<sup>th</sup> Year 2<sup>nd</sup> Term Examination, 2023  
Department of Computer Science and Engineering  
CSE 4225  
Embedded Systems

TIME: 3 hours

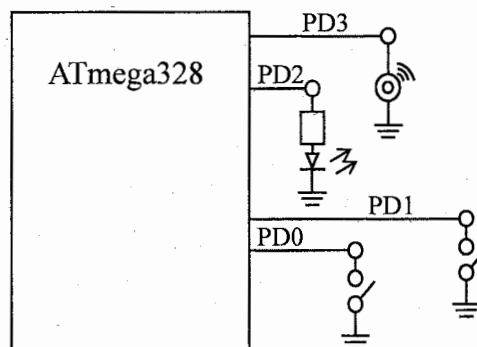
FULL MARKS: 210

- N.B. i) Answer **ANY THREE** questions from each section in separate scripts.  
ii) Figures in the immediate right column of the questions indicate full marks.

**SECTION A**

(Answer **ANY THREE** questions from this section in Script A)

1. a) What are the key features that make an embedded system different from general purpose computer? (07)  
b) What is meant by Evolvability of an embedded system? Explain in brief. (06)  
c) “Expertise with both software and hardware is needed to optimize design metrics in embedded systems” – justify the statement with proper examples. (09)  
d) For a particular product, the NRE cost and unit cost for the following three IC technologies are shown respectively as follows: (13)  
    i) FPGA: (\$10,000, \$50)  
    ii) ASIC: (\$50,000, \$10)  
    iii) VLSI: (\$200,000, \$5)  
    Determine precise volumes for which each technology yields the lowest total cost.
2. a) Illustrate the comparisons among three types of IC technologies: full-custom, semi-custom and PLD. (11)  
b) What is design-productivity gap? (06)  
c) “Adding designers can decrease individual productivity and at some point can actually lengthen the project completion time” – validate the statement mathematically with suitable example(s). (12)  
d) What are the challenges that are faced in designing embedded systems? (06)
3. a) What are the functions of sensors and actuators in embedded systems? Mention the desirable characteristics of sensors. (10)  
b) Show the hardware design of an automated washing machine and explain with a neat sketch. (11)  
c) An engineer designs a system consisting of two subsystems in series with the reliabilities  $R_1 = 0.99$  and  $R_2 = 0.85$  respectively. The cost of the two subsystems is approximately the same. The engineer decides to add two redundant components. Which of the followings is the best to do? (09)  
    i) Duplicate subsystems 1 and 2 in high-level redundancy.  
    ii) Duplicate subsystems 1 and 2 in low-level redundancy.  
    iii) Replace the second subsystem by a three component parallel system.  
d) Mention few examples of ASIP. (05)
4. a) Discuss an efficient waste management system for smart cities supported by the sensing-as-a-service. Provide appropriate figure(s). (09)  
b) Consider different input and output devices as shown in the following figure. Write the necessary instructions for an Arduino Uno to turn on a lamp and buzzer if the key is in ignition but seat belt is not latched. (09)



- c) What are the necessities of Hardware-Software (HW-SW) codesign? Show a model of the current HW-SW design process and then highlight the changes that occur during HW-SW codesign. (11)
- d) What is IoT Board? What are the different types of IoT Board? (06)

**SECTION B**

(Answer **ANY THREE** questions from this section in Script B)

5. a) What is Real Time Operating System (RTOS)? Differentiate between Hard Real-Time System and Soft Real-Time System with application. (09)
- b) How does RTOS offer its services? List the RTOS services under appropriate categories. (06)
- c) Discuss the RTOS kernel architecture with necessary figure(s). (10)
- d) Explain why Round Robin with time slicing is a hybrid scheduling algorithm (with appropriate example). (10)
6. a) What is Task? Differentiate between Task and Thread. Define Task Control Block (TCB) with its various components. (10)
- b) Which mechanism is used to coordinate between different processes? Discuss the following mechanisms briefly: (10)
- i) Message queue and Mailbox.
  - ii) Pipe.
- c) Suppose, you have to develop a smart card system. Now, discuss the followings for the mentioned application with necessary figure(s). (15)
- i) System functions and design metrics.
  - ii) Hardware architecture.
  - iii) Task and synchronization model.
7. a) Discuss the most common verification method for embedded system development. Draw the flow diagram of the detailed development process using the simulator. (10)
- b) Briefly describe different testing methods using software engineering approaches. How can you choose different test cases? (12)
- c) Suppose, you have to develop an Automatic Chocolate Vending Machine (ACVM). Now, (13)
- i) What are the specifications that you will suggest?
  - ii) Specify design metrics for ACVM.
8. a) Differentiate between Ad-hoc network and Infrastructure network with appropriate figure(s). (10)
- b) How does MODBUS protocol work in industrial automation and control system? Discuss how SCADA and HMI works in Operation Technology (OT) network. (12)
- c) How can you design a custom acceleration device using Hardware descriptor language? Discuss the followings with appropriate block diagram: (13)
- i) FPGA.
  - ii) DSP.

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY  
 B.Sc. Engineering 4<sup>th</sup> Year 2<sup>nd</sup> Term Examination, 2023  
 Department of Computer Science and Engineering  
 CSE 4233  
 Robotics

TIME: 3 hours

FULL MARKS: 210

- N.B. i) Answer **ANY THREE** questions from each section in separate scripts.  
 ii) Figures in the immediate right column of the questions indicate full marks.  
 iii) The rightmost column indicates course outcomes.

**SECTION A**

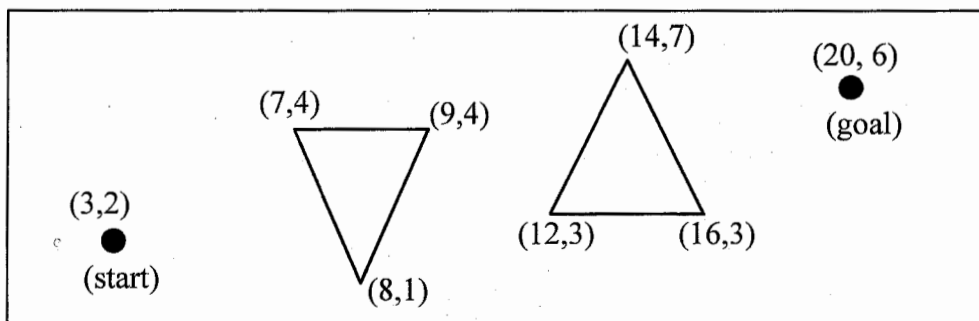
(Answer **ANY THREE** questions from this section in Script A)

1. a) What are the laws of robotics proposed by Isaac Asimov? How robots can be (11) [CO1] differentiated from the fixed automations?  
 b) Describe the historical journey of robots from the “Rossum’s Universal Robots” to (12) [CO1] modern day’s robot.  
 c) Explain how AI has added the dexterity and sophistication to the diversified field of (12) [CO2] robotics.
2. a) Derive the transformation matrix of mapping involving general frames which includes (07) [CO2] rotation along  $z$ -axis.  
 b) Suppose, frame {A} is the universal coordinate frame. Frame {B} is rotated relative to (08) [CO1] frame {A} about  $\hat{Z}$  by  $30^\circ$ , translated 35 units in  $\hat{X}_A$  and 17 units in  $\hat{Y}_A$ . Draw the frames and find  $A_P$ , where  $B_P = [13 \ 74 \ 0]^T$ .  
 c) Draw a cartesian manipulator robot, mention its different parts and describe the points (13) [CO3] to be considered for designing a manipulator.  
 d) What are the different types of joints used in robots? Mention their DOF. Define degree (07) [CO2] of redundancy.
3. a) Differentiate between algebraic and geometric approaches to solve the inverse kinematic (18) [CO2] problem. Using the algebraic approach, derive the joint angle for a 2R robotic arm given the position of the end-effector.  
 b) For an RPR robot, derive the D-H parameters for each joint. Then calculate the position (17) [CO2] and orientation of the end-effector relative to the base frame using the D-H parameters.
4. a) Explain the case of no solution and multiple solutions in case of inverse kinematics (05) [CO2] problems.  
 b) Find the Jacobian matrix and points of singularity of a 2 DOF manipulator. (17) [CO3]  
 c) Derive the equations for an outward and inward iterations of robot dynamics. (13) [CO1]

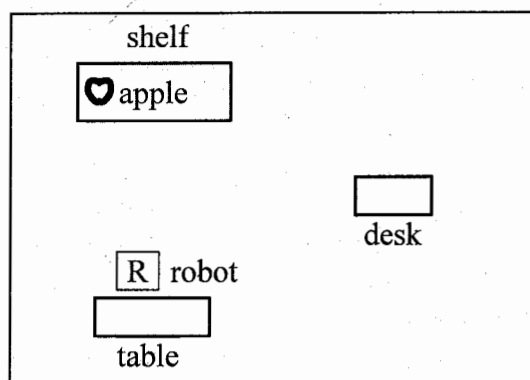
**SECTION B**

(Answer **ANY THREE** questions from this section in Script B)

5. a) Discuss the importance of sensors in robotics, using the Mars rover robot as an example (10) [CO1] to illustrate your explanation. The Mars rover is an autonomous robotic vehicle designed for exploration on the Martian surface.  
 b) Describe in brief the key design challenges you might face in developing legged robots. (09) [CO3]  
 c) Consider a point robot navigating among polygonal obstacles in a plane as shown in (16) [CO2] figure below. Construct a roadmap using a weighted, undirected visibility graph. Then apply Dijkstra’s algorithm to determine the shortest path.



6. a) Explain the significance of using four satellites in GPS for accurate positioning. Discuss whether GPS is suitable for a rescue robot operating in a dense forest environment and justify your answer. (09) [CO1]
- b) A robot R (shown in figure below) receives the command: 'Take the apple from the shelf and put it on the desk'. Define the planning domain of this task using PDDL. Clearly specify the necessary states and actions. Then apply a forward state-space search to achieve the goal state as specified in the task. (20) [CO3]



- c) Differentiate path generation in task space and joint space. (06) [CO4]
7. a) The second joint of a 6-axis robot is to go from an initial angle of  $20^\circ$  to an intermediate angle of  $80^\circ$  in 5 seconds and continue to its destination of  $25^\circ$  in another 5 seconds. Calculate the coefficients for third order polynomial in joint-space. Plot the joint angles, velocities and accelerations. Assume the joint stops at intermediate point. (15) [CO2]
- b) A mobile robot needs to navigate toward a target while avoiding obstacles. Design a fuzzy inference system to control the robot's steering angle and speed. The system should include two input variables – distance to obstacle (0 to 5 meters) and direction to target ( $-90^\circ$  to  $+90^\circ$ ); and two output variables – steering angle (degrees) and speed (m/s). (20) [CO4]
- Define the membership functions for all input and output variables. (Use triangular / trapezoidal shape).
  - Create 5 to 7 rules that maps the input variables to the output variable.
  - Simulate one step of the fuzzy logic process, detailing fuzzification, rule evaluation and defuzzification.
- Assume obstacle distance = 2 meters, target direction =  $30^\circ$ .
8. a) What are the popular goals to control in robotics? Briefly explain feed-forward and feed-back control with necessary block diagram and their application in robotics. (12) [CO1]
- b) Explain the key features of a Humanoid robot and their application. (13) [CO2]
- c) Describe zero moment point and mention its significance in designing robots. (10) [CO3]

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY  
 B.Sc. Engineering 4<sup>th</sup> Year 2<sup>nd</sup> Term Examination, 2023  
 Department of Computer Science and Engineering  
 CSE 4239  
 Data Mining

TIME: 3 hours

FULL MARKS: 210

- N.B. i) Answer **ANY THREE** questions from each section in separate scripts.  
 ii) Figures in the immediate right column of the questions indicate full marks.  
 iii) The rightmost column indicates course outcomes.

**SECTION A**

(Answer **ANY THREE** questions from this section in Script A)

1. a) Why do we need to mine data? Give some examples of what are not data mining. (10) [CO1]  
 b) Explain Quantile-Quantile (Q-Q) plot. What are the impacts of positive and negative correlation of data in daily life? Explain with examples. (10) [CO2]  
 c) What are the features of data matrix? Consider the following table containing attributes of binary type and show the dissimilarity among them. (15) [CO3]

	Feature of fruits	Sphere shape	Sweet	Sour	Crunchy
Object 1	Apple	Yes	Yes	Yes	Yes
Object 2	Banana	No	Yes	No	No
Object 3	Orange	Yes	No	Yes	No

2. a) What are the features of data warehouse? Explain them. (10) [CO2]  
 b) Consider the dataset below: (15) [CO3]  
     5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47  
     i) Obtain and interpret the quartiles.  
     ii) Determine and interpret the interquartile range.  
     iii) Find and interpret five-number summary.  
     iv) Identify potential outliers, if any.  
     v) Construct and interpret a boxplot.  
 c) Explain with flowchart – how do the data warehouse models help in the development of data warehouse. (10) [CO3]
3. a) Explain how to measure interestingness between dataset with example(s). (07) [CO1]  
 b) Consider the following example from a supermarket transactions database and construct a frequent pattern (FP)-Tree. Also, build conditional pattern base for D and E and recursively apply FP-growth on D and E. (18) [CO3]

TID	Items
1	{A, B}
2	{B, C, D}
3	{A, C, D, E}
4	{A, D, E}
5	{A, B, C}
6	{A, B, C, D}
7	{B, C}
8	{A, B, C}
9	{A, B, D}
10	{B, C, E}

- c) Discuss the bottlenecks of Apriori algorithm. Differentiate between parallel projection and partition projection techniques. (10) [CO2]
4. a) Explain the relationship among database processing, data mining processing and warehouse data using example(s). (10) [CO1]  
 b) What is bitmap indexing? Let, a shopping mall has the attributes named *Location, Items, Delivery systems, Discounts, Time, Customer* etc. Draw a star-net query model for this mall where both online and offline orders are processed and demonstrate it with example(s). (10) [CO3]  
 c) Define frequent pattern. Suppose, an association rule  $A \rightarrow B$  has support 60% and confidence 80%. What do you mean by this statement? Explain. (05) [CO2]  
 d) Find out the frequent itemset from the following table using Apriori algorithm where support threshold = 50% and confidence = 60%. (10) [CO3]

Transactions	T1	T2	T3	T4	T5	T6
List of Items	I1, I2, I3	I2, I3, I4	I4, I5	I1, I2, I4	I1, I2, I3, I5	I1, I2, I3, I4

**SECTION B**

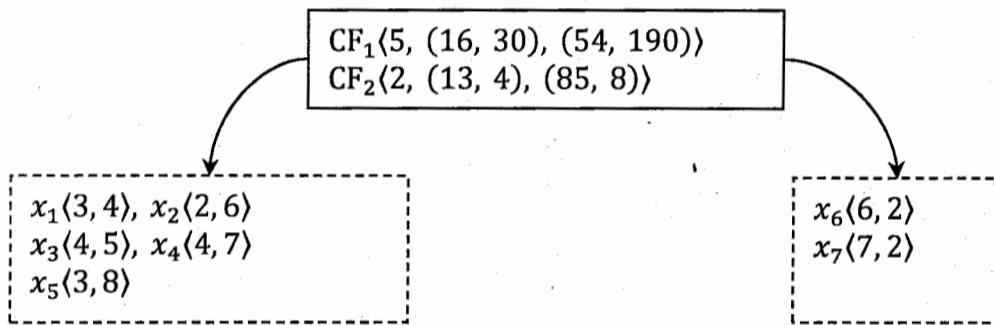
(Answer ANY THREE questions from this section in Script B)

5. a) What are the techniques of data reduction? Explain PCA as measure of dimensionality reduction with example. (12) [CO1]
- b) What are some common techniques for imputing missing values and in what scenarios should each be used? (07) [CO1]
- c) Suppose you have the following data of attribute age: (16) [CO3]  
 21, 24, 18, 16, 7, 54, 22, 28, 36, 22
- Partition this data into three equal-frequency bins.
  - What will be the resulting bin values after smoothing by bin means?
  - How does the smoothing by bin boundaries technique differ from bin means approach? Which method would you recommend for this dataset?
  - Apply min-max normalization to transform the value 54 into the range [0 – 1].

6. a) Explain why bitmap indices are highly efficient for queries involving multiple attributes but less suitable for queries involving a single attribute. (11) [CO2]
- b) Consider the following daily prices of two different cryptocurrencies A and B. If both cryptocurrencies are affected by the same market trends, will their prices rise or fall together? (11) [CO3]

Crypto A	2	3	5	4	6
Crypto B	5	8	10	11	14

- c) Define Clustering Feature (CF) and CF tree for BIRCH clustering. What will the resulting CF tree look like after adding the data points (7, 4), (8, 4), and (7, 9) to the following CF tree, given a distance threshold  $T < 1.5$  and branching factor of 2? (13) [CO3]



7. a) Consider the following patient data; which includes systolic blood pressure (SBP), diastolic blood pressure (DBP), and heart rate (HR). (20) [CO3]

Pateint ID	1	2	3	4	5	6	7	8	9	10
SBP (mmHg)	120	115	118	160	119	121	117	116	118	200
DBP (mmHg)	80	78	79	100	80	81	79	78	79	110
HR (bpm)	72	75	73	90	70	74	73	76	72	95

- Using a z-score of  $\pm 2$ , identify which patients are considered outliers for each feature.
  - Create histograms for SBP (bin-width of 10), DBP (bin-width of 5) and HR (bin-width of 5) and determine which patients are considered outliers by analyzing the histogram.
  - Which method do you think is more effective for this dataset? Provide a brief explanation of your choice.
- b) Explain the strengths and weaknesses of  $k$ -means and  $k$ -medoid methods. (08) [CO1]
- c) Discuss the trade-off between eager and lazy learners with proper example(s). (07) [CO1]
8. a) What is Cluster Based Local Outlier Factor (CBLOF)? Explain how you can use CBLOF to detect outliers. (11) [CO2]
- b) Perform agglomerative clustering on the following data using single linkage method with a distance threshold of 3, where the distance between two clusters is defined as:  
 $d(\text{Cluster}_1, \text{Cluster}_2) = \min\{d(p, q) \mid p \in \text{Cluster}_1, q \in \text{Cluster}_2\}$ .  
 Explain each step in brief. (12) [CO3]

Point	A	B	C	D	E
X	1	2	6	7	3
Y	2	3	6	8	3

- c) Define Mahalanobis distance. How can you use Mahalanobis distance as a measure of multivariate outlier detection? (12) [CO2]

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY  
 B.Sc. Engineering 4<sup>th</sup> Year 2<sup>nd</sup> Term Examination, 2023  
 Department of Computer Science and Engineering  
 CSE 4241  
 Biomedical Engineering

TIME: 3 hours

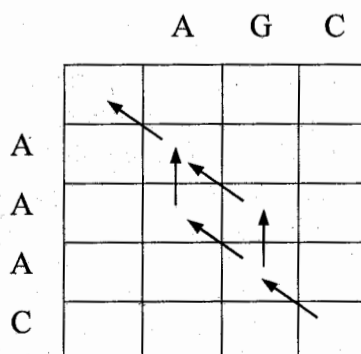
FULL MARKS: 210

- N.B. i) Answer **ANY THREE** questions from each section in separate scripts.  
 ii) Figures in the immediate right column of the questions indicate full marks.  
 iii) The rightmost column indicates the course outcomes.

**SECTION A**

(Answer **ANY THREE** questions from this section in Script A)

1. a) What are Bioinformatics and Medical Informatics? Write down the applications of Bioinformatics. (08) [CO1]
  - b) Apply the Local Dynamic programming algorithm with an affine gap penalty to align the following two sequences. Use the given scoring scheme where the gap opening penalty (h) is  $-3$  and the gap extension penalty (g) is  $-1$ . Show the alignment matrix and the optimal local alignment. (13) [CO2]  
 S1: A A G A  
 S2: T T A A G
  - c) Identify and describe the five categories of hazards present in a clinical environment. For each category, provide one real-life example and discuss its potential impact on patient and staff safety. (08) [CO4]
  - d) Why is it challenging to identify homologous relationships in the Midnight zone? Provide an example to illustrate your explanation. (06) [CO1]
2. a) What are sequence motifs and PROSITE? Explain the process of finding motifs using PROSITE. (10) [CO1]
  - b) Construct a BLOSUM scoring matrix using the following amino acid sequences: (15) [CO2]  
 seq 1 : K P T  
 seq 2 : K P V  
 seq 3 : V P A  
 seq 4 : T P V  
 seq 5 : T P K
  - c) Compute the sequence profile for the following multiple aligned sequences: (10) [CO2]  
 A T A A T A C  
 A T A A T A G  
 A T A A T T C  
 A T A T T A C  
 A T A A T A A
3. a) What are Genomes? What is the ultimate goal of gene prediction? (07) [CO1]
  - b) Why would a full dynamic programming approach not work for most genome alignment tasks? Would using BLAST be useful instead? Justify your answer. (12) [CO2]
  - c) Suppose, you have calculated a score matrix for a pair of DNA sequences. You have performed the trace-back calculation and found a result like following figure. Determine the high road alignment and low road alignment for these DNA sequences. (08) [CO2]



- d) Mr. X is an athlete experiencing muscle fatigue and low energy. He needs protein for energy production and muscle construction. Mr. Y has iron-deficiency anemia and needs a protein to support oxygen transport. Identify the required proteins for each and explain their roles with examples. (08) [CO1]

4. a) What are the pitfalls of biological databases? (04) [CO1]  
 b) Construct a phylogenetic tree by clustering approach from the following nucleotide sequences. (11) [CO2]  
 seq 1 : A T C G A T C G  
 seq 2 : G T A G A C G A  
 seq 3 : A C C G T A C G  
 seq 4 : T C A G T C A G  
 seq 5 : G C C T A C A G  
 c) Consider the following two sequences. Apply FASTA to compute the alignments. (10) [CO2]  
 A C G T C C T T C A T T  
 G T C T C A T G  
 d) Discuss the limitations of the Chou-Fasman algorithm in predicting protein secondary structures. How do neural networks address these limitations, particularly in improving the accuracy of 2D structure prediction? Provide examples to support your explanation. (10) [CO2]

**SECTION B**

(Answer ANY THREE questions from this section in Script B)

5. a) Which factors are important to be considered while designing medical instruments? Draw and explain the block diagram of a generalized medical instrument system. (12) [CO2]  
 b) Write down the significance of Na-K pump in electrophysiology. Calculate the resting potential for the following data. Assume body temperature (37°C) and values of the permeability coefficients as  $P_{Na} = 2 \times 10^8$  cm/s,  $P_K = 2 \times 10^6$  cm/s, and  $P_{Cl} = 4 \times 10^6$  cm/s. (08) [CO1]

Ion	Intracellular(millimoles/litre)	Extracellular(millimoles/litre)
Na <sup>+</sup>	10	140
K <sup>+</sup>	150	4
Cl <sup>-</sup>	4	120

- c) Design an Op-Amp based ECG amplifier circuit and explain the features of your designed circuit. (10) [CO3]  
 d) Briefly discuss on the sources of various noises and artifacts commonly available in ECG signal. (05) [CO1]
6. a) Explain different components of EEG in terms of signal characteristics and mental states. Also, discuss on the role of EEG in Brain Computer Interfacing. (12) [CO1]  
 b) Classify medical equipments. Explain the working principle of Heart-Lung machine as a life support equipment. (10) [CO4]  
 c) What should be the general considerations in the development of biosensor? (06) [CO3]  
 d) Briefly explain the working principle of a demand type synchronous pacemaker using suitable diagram. (07) [CO2]
7. a) Describe the working principle of biosensor in brief. Also, mention the typical sensing techniques for biosensors. (11) [CO3]  
 b) Write down the significance of electromagnetic spectrum on medical imaging. (06) [CO1]  
 c) How does computed tomography differ from conventional radiography? Briefly explain the basic principle of X-ray computed tomography and 3D rendered image reconstruction process. (12) [CO4]  
 d) Write short note on the transition from conventional film based imaging to digital imaging environment. (06) [CO1]
8. a) Describe the importance of using computers in patient monitoring system of ICU. Draw the block diagram of a computer based patient monitoring system. (10) [CO4]  
 b) What do you mean by eHealth? Write down the scopes and challenges of eHealth systems. (10) [CO1]  
 c) List the common risk factors associated with IoMT and suggest possible remedy measures. (07) [CO3]  
 d) Write short notes on: (08) [CO4]  
 i) Genomic Medicine  
 ii) ISO/IEEE 11073 Standard