KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY
B.Sc. Engineering 4th Year 2nd Term Examination, 2024
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) What is meant by Natural Language Processing (NLP)? How does it differ from Machine Learning (ML)? (07) [CO1]
   b) Why are NLP tasks sometimes become difficult for a machine to do? Explain with examples. (12) [CO2]
   c) Differentiate between sentence-level classification and token-level classification with suitable examples. (06) [CO5]
   d) Differentiate between linear regression and logistic regression. Which one is appropriate for binary text classification problem? Give a proper justification. (10) [CO2]

2. a) How does pre-training is related to fine-tuning? Explain. (07) [CO1]
   b) "TF-IDF has better feature representation than BOW."– Is it true? If so, justify. (08) [CO3]
   c) "FFNN is not suitable for processing sequential data." – Explain. (08) [CO2]
   d) Suppose, you need to solve a 3-class text classification problem. How do you perform the task deploying Bi-LSTM? Explain with proper figure. (12) [CO5]

3. a) Draw vanilla transformer architecture and explain its main parts. (06) [CO1]
   b) What are the main functions of query, key, and value vectors in transformer architecture? (06) [CO1]
   c) Discuss pre-training and fine-tuning procedure of BERT model. (15) [CO5]
   d) "BERT can be used as a generalized understanding model." – Explain. (08) [CO1]

4. a) How does BERT is utilized to extract features from the text? Explain with detailed diagram. (12) [CO5]
   b) What is sentiment analysis? How can BERT be used to deal with the sentiment analysis task? (08) [CO1]
   c) Discuss machine translation by vanilla transformer model. (08) [CO2]
   d) What is co-reference resolution? Explain with example(s). (07) [CO1]

SECTION B

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

5. a) Define corpus. What are the properties of a corpus? (05) [CO1]
   b) "Patterns are greedy." – Why? Design a regular expression to answer the following statement: "Find a computer with at least 6 GHz and 16 GB RAM within 800$." (10) [CO1]
   c) You have two strings "brief" and "drivers". Figure out whether "drive" is closer to "brief" or "drivers" using minimum edit distance algorithm. (14) [CO2]
   d) Define stop words. What are the applications of stop word detection in NLP tasks? (06) [CO1]

6. a) Define N-gram language model. How can you compute the probability of a sentence of n words P(w), $w = w_1 w_2 \ldots\ldots w_n$? How can you calculate the maximum likelihood by bigram probabilities? (10) [CO1]
   b) Consider the following corpus shown in the following figure. What is the most probable next word? (15) [CO3]

<s> I like Henry </s>
<s> I like college </s>
<s> Do Henry like college </s>
<s> Henry I am </s>
<s> Do I like Henry </s>
<s> Do I like college </s>

i) "<s> Do __" applying bigram probability model?
ii) "<s> Do I like __" applying bigram probability model?
iii) "<s> Do I like __" applying trigram probability model?

c) Define smoothing. How can you smooth the language model by interpolation? How to choose the value of 'λ' to maximize the probability? (10) [CO3]

7. a) What is the difference between tokens and embeddings? Explain the pipeline of creating embedding model. (08) [CO4]

b) Consider the following grammar G shown in the following figure. Is the string "The flight includes a meal." in L(G)? Answer it using CYK algorithm. (15) [CO3]

S → NP VP
NP → Det N
VP → V NP
V → includes
Det → the | a
N → meal | flight

c) Given a sequence of ice-cream observations (1 1 3) and an HMM λ = (A,B) in the following figure. Find the best hidden weather sequence Q (like H H H). (12) [CO2]

A		B	
start → HOT	0.25		
start → COLD	0.75		
HOT → HOT	0.60	P(1|HOT) =0.20 P(2|HOT) = 0.35 P(3|HOT) = 0.45	P(1|COLD) =0.50 P(2|COLD) = 0.40 P(3|COLD) = 0.10
HOT → COLD	0.40		
COLD → COLD	0.55		
COLD → HOT	0.45		

8. a) "POS tagging is a disambiguation task" – Justify the statement with example. (10) [CO1]

b) Consider the sentence: "It/PRP would/MD be/VB better/JJR to/TO book/? a/DT ticket/NN via/IN internet/NNP." The word "book" is often used as VB or NN. Given the probabilities below, find the right POS tag for the word "book". (12) [CO3]

P(VB|TO) = 0.67, P(NN|TO) = 0.02, P(book|VB) = 0.50,
P(book|NN) = 0.50, P(DT|VB) = 0.02, P(DT|NN) = 0.06.

c) For Hidden Markov Model (HMM) POS tagging, using the following formula, find the equation of calculating tag transition probabilities. (13) [CO2]

$$\hat{t}_1^n = \underset{t_1^n}{argmax}\, P(t_1^n|w_1^n)$$

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY
B.Sc. Engineering 4th Year 2nd Term Examination, 2024
Department of Computer Science and Engineering
CSE 4223
Digital System Design

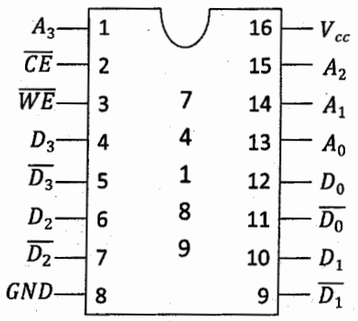
TIME: 3 hours FULL MARKS: 210

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 course outcomes.

SECTION A

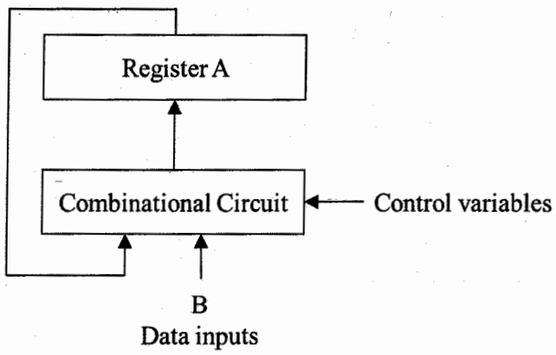
(Answer **ANY THREE** questions from this section in Script A)

1. a) The following figure shows a 64 bits RAM where total number of words is 16 and each word contains 4 bits. Now, using this chip make a 32×16 bits RAM where 32 is amount of word and 16 is size of each word. Also, show the connection of Address, Data and Control bus connections with RAM. (10) [CO1]



b) What is EEPROM? Design an 8×8 diod ROM and explain its operation. (10) [CO1]
c) What are the reasons that may not fit the design system with FPGA? Explain the architectural features of FPGAs. (10) [CO3]
d) "It is impossible to a system with FPGA that will consume less power than ASICs" – Justify the statement. (05) [CO3]

2. a) The block diagram of an accumulator that forms a sequential circuit is shown in the following figure. The A register and the associated combinational circuit constitute a sequential circuit. The combinational circuit replaces the ALU but cannot be separated from the register. (15) [CO1]



Now, using JK flip-flops, design one typical stage of a register that performs the following logic microoperations:

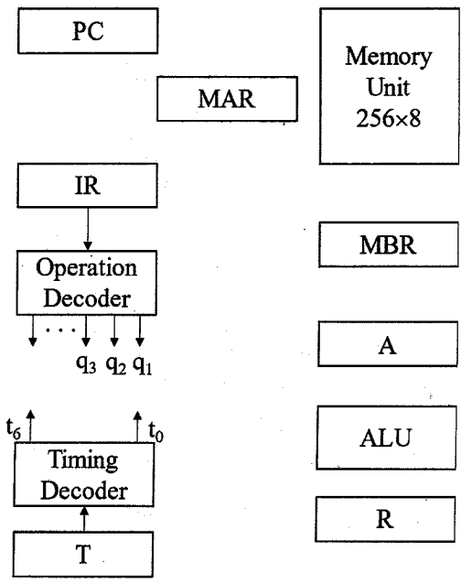
$P_1: A \leftarrow A - B$
$P_2: A \leftarrow \overline{A \vee B}$
$P_3: A \leftarrow \overline{A \wedge B}$
$P_4: A \leftarrow shrA$

b) The following figure shows a block diagram of a simple computer (right side). The register-transfer operations of different timing stages are shown in the following table (left side). (20) [CO2]

i) You have to design the control variables showing their Boolean functions.

ii) Also, complete the hardware connection of the block diagram with control variables showing the directions of the register transfer directions.

Instruction	Time	RTL Statement
FETCH	t_0:	$MAR \leftarrow PC$
	t_1:	$MBR \leftarrow M, PC \leftarrow PC + 1$
	t_2:	$IR \leftarrow MBR$
SUB	q_1t_3:	$MAR \leftarrow PC$
	q_1t_4:	$MBR \leftarrow M, PC \leftarrow PC + 1$
	q_1t_5:	$R \leftarrow MBR$
	q_1t_6:	$A \leftarrow A - R, T \leftarrow 0$
OR	q_2t_3:	$MAR \leftarrow PC$
	q_2t_4:	$MBR \leftarrow M, PC \leftarrow PC + 1$
	q_2t_5:	$R \leftarrow MBR$
	q_2t_6:	$A \leftarrow A \vee R, T \leftarrow 0$
BRZ	q_3t_3:	$MAR \leftarrow PC$
	q_3t_4:	$MBR \leftarrow M, PC \leftarrow PC + 1$
	q_3t_5:	$if(A = 0) then\ PC \leftarrow MBR, T \leftarrow 0$



3. a) A digital system has three register: AR, BR, and PR. Three flip-flops provide the control functions for the system: S is a flip-flop which is enabled by an external signal to start the system's operation; F and R are used for sequencing the microoperations. A fourth flip-flop, D, is set by the digital system when the operation is completed. The function of the system is described by the following register-transfer operations: (10) [CO1]

$S: PR \leftarrow 0, S \leftarrow 0, D \leftarrow 0, F \leftarrow 1$

$F: F \leftarrow 0, if\ (AR = 0)\ then\ (D \leftarrow 1)\ else\ (R \leftarrow 1)$

$R: PR \leftarrow PR + BR, AR \leftarrow AR - 1, R \leftarrow 0, F \leftarrow 1$

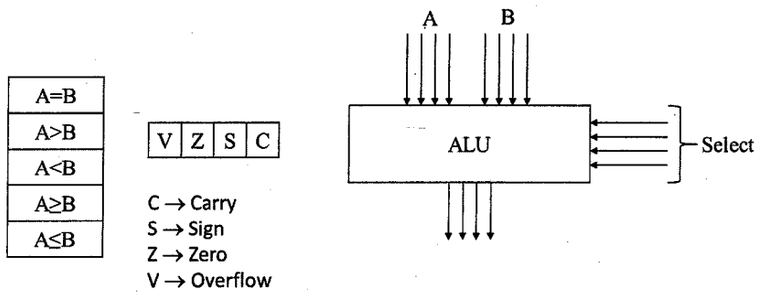
Now, analyze this digital system using an example so that you can explain what function that the system performs.

b) "One cannot tell from looking at an isolated register-transfer statement whether it represents a macro- or microoperation"– Justify the statement with proper example. (05) [CO1]

c) The following figure shows a status register. These status bits in status register also used in comparator circuit. Now, (10) [CO1]

i) Complete the combinational circuits to detect status bits form the ALU.

ii) How this status register used in compare operations? Explain with necessary circuit.



d) What are the differences between Circular Right Shift and Arithmetic Right Shift? Design a general shift register and write the corresponding Verilog module. (10) [CO3]

4. a) A CPU with a 2.0 GHz base clock speed utilizes a four-stage pipeline to process a program of 1000 instructions where 80% are standard `ADD` instructions requiring 10 total cycles and 20% are complex `MUL` instruction requiring 18 total cycles. However, CPU detects a high thermal load during the execution of `MUL` instructions and reduce the clock speed to 1.5 GHz for that specific instruction. (10) [CO2]

i) Calculate the clock period for both instructions.

ii) If thermal threshold is 150 instructions then calculate total execution time for the program.

b) Write an SAP-1 program for the expression: $30 - 12$ in machine code to show the output result. Also, show the necessary control word for each instruction's Fetch and Execution cycle generated by control unit. (20) [CO2]

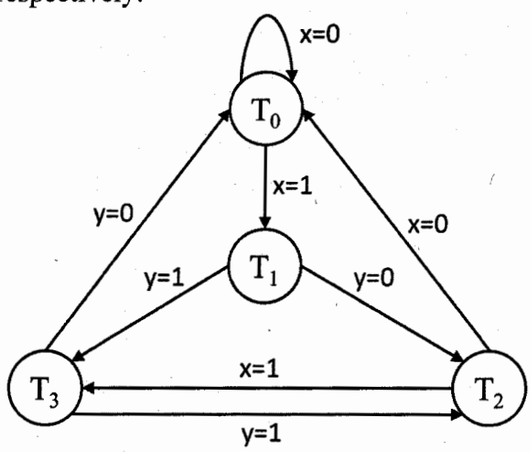
c) How can you make the Machine cycle variable for SAP-1? Explain with example. (05) [CO2]

SECTION B

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

5. a) Explain the relationship between control logic and data processor in a digital system using a block diagram. (07) [CO2]

b) The state diagram shown in the following figure has four states and two inputs x and y. Design the control unit using sequence register and decoder method using T and JK flip-flops for G_1 and G_2 respectively. (14) [CO2]



c) The following register-transfer operations specify an eight-states control of the sequence register and decoder type. G is a 3-bit sequence register and the outputs of the decoder are as follows for two inputs x and y. (14) [CO2]

$x'T_0: G \leftarrow T_0$

$y'T_2 + xT_0: G \leftarrow G + 1$

$T_4 + yT_2: G \leftarrow G + 3$

$T_1 + T_3 + T_5 + T_6: G \leftarrow G + 1$

$T_7: G \leftarrow T_0$

Draw the state diagram of the control, develop excitation table and draw flip-flop input function to design the control unit with T or D flip-flops.

6. a) What do you mean by reengineering and reverse engineering in the contexts of digital system design? Briefly mention different aspects of both. (07) [CO2]

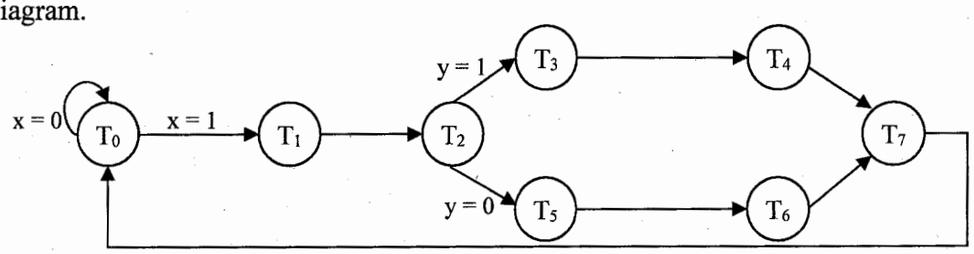
b) Consider a control unit with two JK flip-flops and 2×4 decoder. The inputs of the JK flip-flops are (14) [CO2]

$JG_2 = T_1$ $\quad KG_2 = T_3y$

$JG_1 = T_0x + T_2$ $\quad KG_1 = 1$

Here x and y are inputs of the system; T_0, T_1, T_2 and T_3 are the states of the system. Draw the logic diagram and develop excitation table for inputs $x, y \Rightarrow 00, 01, 10,$ and 11. Simplify the excitation table considering appropriate don't care conditions for x and y. Finally, redesign the control unit for T flip-flops based on the simplified excitation table.

c) Design the control unit using a PLA and other required elements for the following state diagram. (14) [CO2]



7. a) Draw block diagram of a typical micro-programmed control unit and briefly explain purpose of its individual components. (10) [CO2]

b) Design a micro-programmed control unit for a system that is able to display 0 to 9 through seven segment-display. The control unit has two inputs q_{odd} and q_{even}, and functions of the system for the inputs are below. (18) [CO2]

q_{odd}	q_{even}	Action
0	0	Stay in the initial state
0	1	Display even numbers only and return initial state
1	0	Display odd numbers only and return initial state
1	1	Display all the numbers and return initial state

c) Suppose, you need to design a small digital system which needs 15 individual instructions and its accumulator size is 16. Calculate the sizes of other registers and memory. (07) [CO2]

8. a) What are the different types of instructions in the studied digital computer design? Briefly describe significances of each type of instruction. (07) [CO2]

b) Draw block diagram of a small-scale digital computer with different registers and counters. Indicate data flow and control signal flow directions among the components with solid lines and dotted lines respectively. (14) [CO2]

c) "The digital computer operates in discrete steps controlled by the timing signals" – Justify the statement with an example of AND operation of studied small-scale computer. (07) [CO2]

d) The following sequence of micro-operations are performed in an accumulator considering their usual meaning. (07) [CO2]

```
CLA
AND
INC
ADD
INC
INC
CMA
```

Determine the content of A after each micro-operation if initially A=1001 and M=0110.

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY
B.Sc. Engineering 4th Year 2nd Term Examination, 2024
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) A cloud service provider monitors server response time (in ms) during peak hours. A sample of the recorded data is as follows: (15) [CO2]

85, 90, 95, 100, 110, 115, 130, 160, 210

i) Explain how a Q-Q plot would be used to assess whether the data follows a normal distribution.
ii) Identify whether the dataset indicates skewness or heavy tails.
iii) Discuss how such behavior affects system reliability.

b) An online advertising company tracks daily advertisement spend and revenue as follows: (10) [CO3]

Advertisement Spend (in $1000)	2	4	6	8	10	12
Revenue (in $1000)	5	9	12	15	18	21

i) Analyze the type and strength of correlation.
ii) Explain whether increasing advertisement spend is always beneficial.

c) Suppose that a data warehouse contains 20 dimensions, each with about five levels of granularity. (10) [CO1]

i) Users are mainly interested in four particular dimensions, each having three frequently accessed levels for rolling up and drilling down. How would you design a data cube structure to efficiently support this preference?
ii) At times, a user may want to drill through the cube, down to the raw data for one or two particular dimensions. How would you support this feature?

2. a) A smart home automation system records the states of various devices at different times. The states are represented as symmetric binary attributes, where 1 indicates "ON" and 0 indicates "OFF". Consider the following recorded states: (12) [CO1]

Device State	Light	AC	Heater	Fan
S_1	1	0	1	0
S_2	1	1	0	0
S_3	0	1	1	1

i) Compute dissimilarity between each pair.
ii) Identify which states are most similar.
iii) Explain how this can be used for anomaly detection.

b) "Strong associative rules are not necessarily interesting" – why? What is the cure? (10) [CO1]

c) A ride-sharing company collects trip data including time, location, driver, customer, and ride type information. Design a star schema for this system using the following: (13) [CO3]

Dimensions: Time, Location, Driver, Customer, Ride Type
Fact: Trip Details

Clearly identify the fact table and all dimension tables with appropriate attributes.

3. a) Define Frequent Pattern (FP). Explain support and confidence of an association rule. (05) [CO1]

b) Discuss the bottlenecks of Apriori algorithm. Differentiate between parallel projection and partition projection techniques. (10) [CO2]

c) Consider the following transaction database example. Construct a FP-Tree. Also, build conditional pattern base for D and E and recursively apply FP-growth on D and E. (20) [CO3]

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

4. a) A superstore maintains sales data with the following dimensions: Product (Category), Location (City), Time (Month). Generate the complete lattice of cuboids for this system. (12) [CO2]
   i) List all cuboids.
   ii) Arrange them in hierarchical order from base cuboid to apex cuboid.

   b) A fitness analytics company records the daily Calorie intake (in Kcal) of users as follows: (16) [CO3]

   1800, 2200, 1500, 1300, 2500, 2100, 1900, 3000, 1700, 2000

   i) Partition this data into three equal-frequency bins after sorting the values.
   ii) Apply smoothing by bin means and show the resulting values for each bin.
   iii) Explain how smoothing by bin boundaries differs from bin means. Which method would be more suitable for this dataset? Justify your answer.
   iv) Apply min-max normalization to transform the value 3000 into the range [0 – 1].

   c) What is Data Mart? Explain the categories of a Data Mart. (07) [CO1]

SECTION B

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

5. a) Discuss the initialization problem of k-means. Why does k-means++ select centroids probabilistically instead of directly selecting the farthest point? (12) [CO2]

   b) Explain how Genetic Algorithm can be used for k-means centroid initialization. Why is k-medoids more robust to outliers compared to k-means? (13) [CO2]

   c) Explain the strengths and weaknesses of k-means and k-medoid methods. (10) [CO1]

6. a) Consider the following 2D data points: (15) [CO3]

   (1, 1), (1, 2), (2, 1), (2, 2), (8, 8), (8, 9), (9, 8), (25, 25), (26, 25), (50, 50)

   For Density-Based Spatial Clustering of Applications with Noise (DBSCAN) assume MinPts = 4.
   i) Determine an appropriate ε using the elbow method.
   ii) Perform DBSCAN and clearly identify: core points, border points and noise.
   iii) State the final clusters.

   b) Explain the intuition behind mutual reachability distance in Hierarchical DBSCAN (HDBSCAN). (05) [CO1]

   c) Consider the 1D dataset: 0.0, 0.1, 0.2, 5.0, 5.1, 5.2, and 9.0. Using HDBSCAN with min-samples = 2 and min-cluster-size = 3, construct the Minimum Spanning Tree (MST). From the MST determine the final clusters and any noise points. (15) [CO3]

7. a) Explain Automatic Differentiation using a suitable example. (06) [CO2]

   b) Why CNN is better than MLP for image data? Explain the benefits of using pooling operation. Also, explain the difference between standard and depthwise separable convolution. (12) [CO1]

   c) Explain Squeeze-and-Excitation channel attention. (07) [CO1]

   d) Define parametric method of outlier detection. How can you detect univariate outlier for normal distributed data? Find outlier from the following data (if any): (10) [CO2]

   40.0, 48.9, 48.9, 49.0, 49.1, 49.2, 49.4, 49.3, 49.6, 49.1

8. a) Define outlier as a data mining problem. Give a classification of outliers. (07) [CO1]

   b) How do you differentiate between outlier and noise? Explain. (05) [CO1]

   c) Construct a KD-tree for the points: (12) [CO3]

   (2, 3), (3, 5), (5, 4), (6, 8), (8, 7), (9, 2), (4, 7)

   Find 3 nearest neighbors of the query point (5, 5). Use max-heap.

   d) For a 2D dataset, suppose the mean vector and covariance matrix are: (11) [CO3]

$$\mu = (5,5),\ \Sigma = \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix}$$

   For point $X = (11,10)$, calculate the Mahalanobis distance. Use chi-square critical value 5.991 for 2 degrees of freedom at 5% significance level and decide whether X is an outlier.