KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY
B.Sc. Engineering 4th Year 2nd Term Examination, 2020
Department of Computer Science and Engineering
CSE 4223
Digital System Design

TIME: 1.5 hours FULL MARKS: 120

N.B. i) Answer **ANY TWO** questions from each section in separate scripts.
ii) Figures in the right margin indicate full marks.

SECTION A

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

1. a) What is a register? Design a 3-bit tri-state buffer register with necessary diagram. (20)
   b) Differentiate between serial loading and parallel loading with example(s). (10)

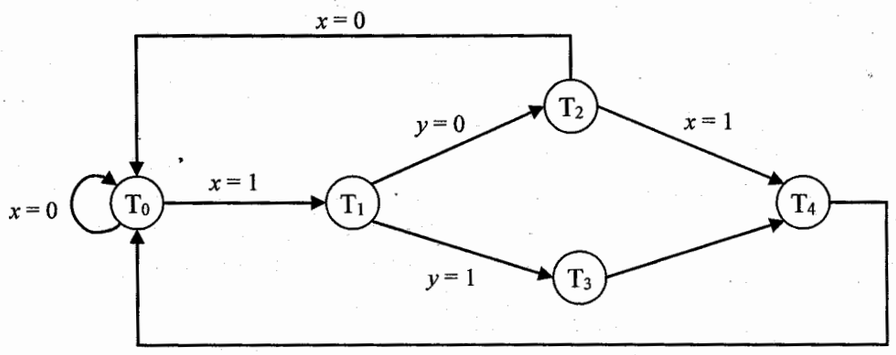
2. a) Draw timing diagram and generate Control word for LDA, ADD, SUB, and OUT instructions in case of SAP-1 computer. (20)
   b) Write SAP-1 program for the expression: 32 + 20 – 16 + 10. (10)

3. a) "In register transfer logic '+' symbol has two meanings" – Explain with example. (12)
   b) What is instruction code? Explain. (08)
   c) Distinguish between Scratchpad memory and 2-port memory Processor Organization. (10)

SECTION B

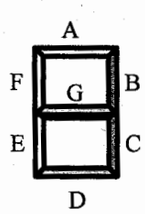
(Answer **ANY TWO** questions from this section in Script B)

4. a) Is it possible to distinguish data and control information in a Digital System only looking to the values in processor or memory resistors? Justify your answer. (06)
   b) Explain the relation between control logic and data processor in a digital system using a block diagram. (06)
   c) A control unit has two inputs x and y, and five states shown in the figure below. (18)



Design the control unit using sequence register and decoder method. Consider T, JK, and SR flip-flops for G3, G2, and G1, respectively for ODD Roll. For EVEN Roll, consider T, SR, and JK flip-flops for G3, G2 and G1, respectively.

5. a) Does PLA only may act as control unit? Justify your answer. (05)
   b) What are the provisions necessary in address sequencing in the micro-program control? (05)
   c) Consider a seven segment display as a data processor which may display 0 to 9 in ten different states. Design a digital system with micro-program control (i.e., state diagram, block diagram, binary micro-program) which may display digits of your OWN Roll number. System will consider an external input d; for $d = 0$, system will stay in initial state and for $d = 1$, system will display the digits and return to initial state. (20)



6. a) Why are the different timing signals required in addition to clock pulse in a digital system? Explain with a proper example. (07)

   b) What are the basic memory reference instructions considered in the studied small-scale digital computer? Briefly explain their functions. (09)

   c) An instruction in address $(A98)_{16}$ in the computer has the operation code of the ADD instruction and an address part $(B98)_{16}$. The memory word of address $(B98)_{16}$ contains the number $(XXXX)_{16}$, where XXXX means the last four digits of your Roll number. Register A contains $(A0B0)_{16}$. Tabulate the contents of register PC, MAR, B, A and I, after the instruction is executed. Repeat the problem for the memory reference code STO. (14)

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY
B.Sc. Engineering 4th Year 2nd Term Examination, 2020
Department of Computer Science and Engineering
CSE 4239
Data Mining

TIME: 1.5 hours FULL MARKS: 120

N.B. i) Answer **ANY TWO** questions from each section in separate scripts.
ii) Figures in the right margin indicate full marks.

SECTION A

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

1. a) What is data mining? Suppose that the data for analysis includes attribute 'age'. The age values for the data tuples are: (18)

13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 40

i) Show a boxplot of the data.
ii) Plot an equal-width histrogram of width 10.
iii) iii) Calculate z-score normalization and min-max normalization by setting min = 0 and max = 1.

b) Suppose a hospital tested the age and body fat data for 8 selected adults with the following results. Calculate the correlation coefficient (pearson's product moment coefficent). (12)

Age	54	54	56	57	58	58	60	61
% fat	42.5	28.8	33.4	30.2	34.1	32.9	41.2	35.7

2. a) Suppose a data warehouse consists of four dimensions; date, spectator, location and game; and the two measures, count and charge; where charge is the fare that a spectator pays when watching a game on a given date. Spectators may be students, adults or seniors, with each category having its own charge rate. (20)

i) Draw a 'Snowflake schema' and a 'Fact constellations' diagram for the data warehouse.
ii) Starting with the base cuboid [date, spectator, location, game], which specific OLAP operations should one perform in order to list the total charge paid by student spectators at GM_Place in 2010?
iii) Taking this cube as an example, discuss advantages and problems of using a bitman index structure.

b) Suppose that a data'warehouse contains 20 dimensions, each with about five levels of granularity. (10)

(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?

3. a) Assume a base cuboid of 10 dimensions contains only three base cells: (1) $(a_1, d_2, d_3, ..., d_9, d_{10})$, (2) $(d_1, b_2, d_3, ..., d_9, d_{10})$ and (3) $(d_1, d_2, c_3, ..., d_9, d_{10})$, where $a_1 \neq d_1$, $b_2 \neq d_2$ and $c_3 \neq d_3$. The measure of the cube is count. (20)

i) How many nonempty aggregate (i.e., nonbase) cells will a full cube contain?
ii) How many nonempty aggregate cells will an icebery cube contain if the condition of the icebery cube is "count ≥ 2" ?
iii) (iii) How many closed cells are in the full cube?

b) Briefly explain the computation technique of Icebery cubes from the Apex cuboid downward. (10)

SECTION B

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

4. a) What do you mean by closed frequent pattern? (02)
   b) A databse has six transactions described in the following table. Let min-sup = 60% and min-conf = 80%. (13)

TID	Item Bought
T1000	M, O, N, K, E, Y
T2000	D, N, K, E, Y
T3000	M, O, N, K, E, Y, L
T4000	M, U, C, K, O
T5000	C, O, O , K, E, S
T6000	M, A, K, E

   Now perform the followings:
   (i) Find the frequent itemset using apriori principle.
   (ii) List all the strong association rules.
   c) How does a boxplot visualization identify outliers? Explain with examples. (08)
   d) What is Laplace correction? What are the advantages and disadvantages of Naïve Bayesian classification? (07)

5. a) Define Clustering Feature (CF) and CF tree for BIRCH clustering. Apply BIRCH algorithm for the following dataset: <2, 5>, <3, 2>, <4, 3>, <5, 6>, <6, 8>, <8, 10> (09)
   b) Differentiate between agglomerative and divisive methods of clustering. (05)
   c) Explain how a ROC curve works. Draw ROC curve for the below given table: (10)

Tuple Number	Class	Probability
1	Class 1	0.78
2	Class 1	0.90
3	Class 2	0.61
4	Class 1	0.50
5	Class 2	0.62

   d) What is the class imbalance problem? How will you evaluate the performance of a classifier having class imbalance problem? (06)

6. a) "Active learning is a supervised learning" – justify this statement. (08)
   b) Define paremetric method for outlier detection. How can you detect univariate outlier for normal distributed data? Find outlier from the following data (if any). 148.9, 148.9, 149.0, 140.0, 149.1, 149.1, 149.2, 149.4, 149.3, 149.6. (10)
   c) What is cluster-based local outlier factor? How can we use it for outlier detection purpose? Explain. (06)
   d) What are the differences between traditional learning and transfer learning? (06)

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY
B.Sc. Engineering 4th Year 2nd Term Examination, 2020
Department of Computer Science and Engineering
CSE 4241
Biomedical Engineering

TIME: 1.5 hours FULL MARKS: 120

N.B. i) Answer **ANY TWO** questions from each section in separate scripts.
ii) Figures in the right margin indicate full marks.

SECTION A

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

1. a) "Biomedical Engineering is an interdisciplinary field" – justify the statement and hence, mention the contributional scopes of CSE students/graduates in the diverse fields of Biomedical engineering. (14)
   b) Define action and resting potential. Describe the generation process of action potential in the living excitable cells using suitable diagram. (16)

2. a) Write down the full form of ECG, EEG, EMG, and EOG. Distinguish these biosignals interms of their originating organ, voltage and frequency ranges. (09)
   b) What are the ECG signal processing steps of Pan-Tompkins algorithm for real time heart beat (QRS) detection? (06)
   c) What do you mean by computed tomography (CT)? Mention the technical features of different generations of CT. (15)

3. a) Write down the components of X-ray tube along with their functions. Why are the collimator and grid used in X-ray imaging? (10)
   b) Draw the general networking systems for developing integrated interconnectivity among various units/workstations inside and outside of a healthcare organization. (12)
   c) Define eHealth. List different standards used in eHealth and explain any one of them in brief. (08)

SECTION B

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

4. a) Define Bioinformatics and mention its sub-disciplines. What is the main role of a bio informatician in present biological research and development area? (09)
   b) Consider the following two strings: (12)

      ALGO
      TEST

      (i) What is the optimal alignment?
      (ii) What is the cost/score of the optimal alignment?
   c) Given is a set of multiple aligned sequences. Compute the sequence profile for this set. (09)

      ATAATAC
      ATAATAG
      ATAATTC
      ATATTAC
      ATAATAA

5. a) "Protein act as messenger" – justify the statement with example(s) (10)
   b) What is sequence Motif and PROSITE? Write down the steps to find the motif on PROSITE? (12)
   c) What is BLAST and FASTA? Make a relationship between FASTA and BLAST. (08)

6. a) Define term propensity value. What steps does Chou-Fasman method performs to predict the secondary structure of protein? (13)
   b) Explain the difference between homology and similarity. How is similarity used to infer homology? (10)
   c) What is meant by gene prediction? What is the ultimate goal of gene prediction? (07)

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY
B.Sc. Engineering 4th Year 2nd Term Examination, 2020
Department of Computer Science and Engineering
CSE 4221
Natural Language Processing

TIME: 1.5 hours FULL MARKS: 120

N.B. i) Answer **ANY TWO** questions from each section in separate scripts.
ii) Figures in the right margin indicate full marks.

SECTION A

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

1. a) "Pattern matching by regular expression is greedy" – Explain the term taking example to match the word 'There' from a string. (10)
   b) Given two strings $S1$ and $S2$ and three operations ADD, DELETE and EDIT. How many minimum operations are needed to convert $S1$ to $S2$? Derive the algorithm to answer the question. Apply the algorithm for $S1$ = "*abcdef*" and $S2$ = "*azced*". If each of the operation costs 1 unit, find the total cost to convert $S1$ to $S2$. (20)

2. a) "Minimizing perplexity is same as maximizing the probability" – Explain the term. (12)
   b) Consider the following corpus: (12)

   $\langle S\rangle abc\langle /S\rangle$
   $\langle S\rangle bac\langle /S\rangle$
   $\langle S\rangle abc\langle /S\rangle$

   Calculate the bigram probabilities $P(a|\langle S\rangle)$, $P(c|b)$ and $P(b|a)$.
   c) How do you handle the stop words for text classification? (06)

3. a) Given the following text classification for C_1 and C_2 (14)

   $PQR \rightarrow C_1$
   $PPS \rightarrow C_1$
   $PT \rightarrow C_1$
   $UVP \rightarrow C_2$

   Compute the most likely class for the text "*PPPUV*".
   b) Consider the following grammar in CNF (16)

   $S \rightarrow AB \mid BC$
   $A \rightarrow BA \mid a$
   $B \rightarrow CC \mid b$
   $C \rightarrow AB \mid a$

   Is '*baaba*' in $L(G)$? Explain your answer using CYK algorithm.

SECTION B

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

4. a) "POS tagging is a disambiguation task" – Justify the statement with example. (10)
   b) Consider the sentence: (08)
   "*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 "race".

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

   c) What is prosody? Discuss about the three aspects of prosody. Discuss the effect of zero situation. (12)

5. a) Given a sequence of ice-cream observations (1 1 3) and an HMM $\lambda = (A, B)$ in the following figure, find the best hidden weather sequence $Q(like\ H\ H\ H)$. (12)

A		B	
$Start \rightarrow HOT$	0.25		
$Start \rightarrow COLD$	0.75	$P(1 \mid HOT) = 0.2$	$P(1 \mid COLD) = 0.5$
$HOT \rightarrow HOT$	0.60	$P(2 \mid HOT) = 0.35$	$P(2 \mid COLD) = 0.4$
$HOT \rightarrow COLD$	0.40	$P(3 \mid HOT) = 0.45$	$P(3 \mid COLD) = 0.1$
$COLD \rightarrow COLD$	0.55		
$COLD \rightarrow HOT$	0.45		

b) Define the term odds for logistic regression. Show that the observation should be labeled true if $\sum_{i=0}^{N} w_i f_i > 0$. (12)

c) What are the components of Markov chain? (06)

6. a) Discuss about the steps of text to waveform mapping in speech synthesis using Hourglass metaphor. (10)

b) What is Homograph disambiguation? What are the problems of CMU? How does UNISYN overcome the problems of CMU? (10)

c) What are the use of Question Answering System (QAS)? Discuss about the main phases of QAS architecture. (10)