



Answer the following 4 questions:

(Total marks: 80)

1st Question: Choose the Correct Answer

marks: 20

1. Quality Assurance methods are usually considered:

- a. Detective
- b. Preventive
- c. Corrective
- d. Protective

2. Quality has been defined as:

- a. Meeting requirements
- b. Customer satisfaction
- c. Zero defects
- d. All of the above
- e. (a) and (b)

3. Which of the following is a form of functional testing?

- a. Boundary value analysis
- b. Usability testing**
- c. Performance testing
- d. Security testing

6. A deviation from the specified or expected behavior that is visible to end users is called:

- a. an error
- b. a fault
- c. a failure
- d. a defect

7. The cost of fixing a fault:

- a. is not important
- b. increases the later a fault is found
- c. decreases the later a fault is found
- d. can never be determined

8. A program with high cyclometric complexity is almost likely to be:

- a. Large
- b. Small
- c. Difficult to write
- d. Difficult to test

10. To test a function, the programmer has to write a_____, which calls the function and passes test data.

- a. Stub c. Proxy
- b. Driver d. None of the above

11. Which testing methods are used by end-users who actually test software before they use it.

- a. Alpha and Beta Testing
- b. White Box Testing
- c. Black Box Testing
- d. Trial and Error Testing

12. Maintenance testing is:

- a. updating tests when the software has changed
- b. testing a released system that has been changed
- c. testing by users to ensure that the system meets a business need
- d. testing to maintain business advantage

13- type of integration testing in which software elements, hardware elements, or both are combined all at once into a component or an overall system, rather than in stages.

- a. System Testing
- b. Big-Bang Testing**
- c. Integration Testing
- d. Unit Testing

14- Testing should be stopped when:

- a. all the planned tests have been run
- b. all faults have been fixed correctly
- c. both a) and b)
- d. it depends on the risks for the system being tested

15- Order numbers on a stock control system can range between 10000 and 99999 inclusive. Which of the following inputs might be a result of designing tests for only valid equivalence classes and valid boundaries:

- a. 1000, 5000, 99999
- b. 9999, 50000, 100000
- c. 10000, 50000, 99999
- d. 10000, 99999

16- Given the following code, which is true about the minimum number of test cases required for full statement and branch coverage:

Read P

Read Q

IF P+Q > 100 THEN

Print "Large"

ENDIF

If P > 50 THEN

Print "P Large"

ENDIF

- a. 1 test for statement coverage, 3 for branch coverage
- b. 1 test for statement coverage, 2 for branch coverage
- c. 1 test for statement coverage, 1 for branch coverage
- d. 2 tests for statement coverage, 2 for branch coverage

17- Beta testing is:

- a. Performed by customers at their own site
- b. Performed by customers at their software developer's site
- c. Performed by an independent test team
- d. Performed as early as possible in the lifecycle

18- The difference between re-testing and regression testing is

- a. re-testing ensures the original fault has been removed; regression testing looks for unexpected side-effects
- b. re-testing looks for unexpected side effects; regression testing is repeating those tests
- c. re-testing is done after faults are fixed; regression testing is done earlier
- d. re-testing is done by developers, regression testing is done by independent testers

19-When testing a grade calculation system, a tester determines that all scores from 90 to 100 will yield a grade of A, but scores below 90 will not. This analysis is known as:

- a. Equivalence partitioning
- b. Boundary value analysis
- c. Decision table
- d. Hybrid analysis

20. Function Point Analysis (FPA) measures the size of a software product in terms of:

- a. The functionality it delivers
- b. The Line of code it delivers
- c. The Interface it delivers
- d. (a) and (b)

2nd Question:

marks: 10

A software project of application generator category with estimated 60KLOC has to be developed. The scale factor (B) has high precedentness, low development flexibility and high team cohesion. Other factors are nominal. The early design cost drivers like platform difficult (PDIF) and Personnel Capability (PERS) are high and others are nominal. Calculate the effort of the project in person months (Pm nominal and Pm adjusted) using COCOMO2

3rd Question

marks: 20

Q1) In discussing the funds transfer software with the national bank, it appears that some small changes in design the system could reduce the adaptation effort. The resulting estimates of the adaptation parameter are (10 marks):

$$DM=40,$$

$$CM=60,$$

$$IM=80$$

Calculate the corresponding estimated effort required for these changes

Q2) Consider a database system needed for an office automation system. CAD CO, Inc. wants to develop the system. They are a new company and a bit inexperienced in CAD systems. Initial analysis of the problem leads to requirements calling for 4 major modules with the following size (10 Marks):

Data Entry	0.6 KDSI
Data Updates	0.6 KDSI
Query	0.8 KDSI
Report Generation	1.0 KDSI

The adjusted factors are estimated as follows:

Data base size	high
Product Complexity	very high
Main storage	very high
Execution time constraints	high
Programmer capability	very high
All other characteristics rate Nominal	

- a) Estimates the total effort, duration of the project, the no of developers needed, using the intermediate COCOMO model
- b) Calculate the cost required to develop the system if the MM cost is 5000\$ and the product required high reliability. Then calculate the cost again if the product required low reliability.
- c) As a software engineer decide which option to choose (low reliability or high reliability) and why?
- d) Calculate the annual effort and cost required to maintain this system if the annual Traffic is 15%, reliability is low and the MODP is high.

4th Question:

marks: 30

Q1) A course taught at a university has two components: exam and coursework; to pass the course with a Grade C, a student must score at least 50% in the exam, and 50% in the coursework. They pass in the course with Grade B, if they score at least 60% and 50% in the coursework. If, in addition to this, the average of the exam and the coursework is at least 70%, then they are awarded a Grade A (*20 Marks*)

```

10 main()
20 {
30 float exam, course, average;
40 char trailer[20];
50 cout <<"Input exam mark: "<<endl;
60 cin>>exam;cin.ignore(80,'n');
70 cout <<"Input course mark: "<<endl;
80 cin>>course;cin.ignore(80,'n');
90 if ( (exam<0.0) || (exam>100.0) || (course<0.0) || (course>100.0) )
110 cout<<"Marks out of range - program terminating"<<endl;
120 else {
130 if ( (exam<50.0) || (course<50.0))
140 cout<<"Fail"<<endl;
150 else if (exam < 60)
160 cout<<"Pass - Grade C"<<endl;
170 else if ( ((exam+course)/2.0) >= 70.0) {
180 average = (exam+course)/2.0;
190 cout<<"Pass - Grade A with an average mark of "<< average <<endl;
200 }
210 else
220 cout<<"Pass - Grade B"<<endl;
230 }
240 }
```

For the exam grade classification program given above, draw a Control Flow Graph (CFG). Using this, justify how you would derive test cases and test data for the following White Box Software testing methods:

i. Branch Testing ii. Mutation Testing for line 130 (test cases for dead and alive cases)
Then, derive test cases and test data for the following Black Box Software testing methods:

i. Boundary Value Analysis ii. Equivalence class partitioning

Q2) Design a decision table for a shop owner allows credit facility to his customers if they satisfy any one of the following conditions (*10 marks*)

1. Holding the present job for more than 3 years and residing in the same place for more than 5 years
2. Monthly salary exceeds 1500\$ and holding the present job for more than 3 years
3. Residing in the same place for more than 5 years and monthly salary exceeds 1500\$, the facility is rejected for all other customers

With My Best Wishes,

Rasha Ismail

Effort(MM)	Small 2 KDSI	Intermediate 8 KDSI	Medium 32 KDSI	Large 128 KDSI	Very Large 512 KDSI
Organic	5.0	21.3	91	392	
Semidetached	6.5	31	146	687	3250
Embedded	8.3	44	230	1216	6420
Productivity (DSI/MM)	Small 2 KDSI	Intermediate 8 KDSI	Medium 32 KDSI	Large 128 KDSI	Very Large 512 KDSI
Organic	400	376	352	327	
Semidetached	308	258	219	186	158
Embedded	241	182	139	105	80
Schedule (months)	Small 2 KDSI	Intermediate 8 KDSI	Medium 32 KDSI	Large 128 KDSI	Very Large 512 KDSI
Organic	4.6	8	14	24	
Semidetached	4.8	8.3	14	24	42
Embedded	4.9	8.4	14	24	41
Average Personnel(FSP)	Small 2 KDSI	Intermediate 8 KDSI	Medium 32 KDSI	Large 128 KDSI	Very Large 512 KDSI
Organic	1.1	2.7	6.5	16	
Semidetached	1.4	3.7	10	29	77
Embedded	1.7	5.2	16	51	157

Effort Distribution		Size				
Mode	Phase	Small 2 KDSI	Intermediate 8 KDSI	Medium 32 KDSI	Large 128 KDSI	Very Large 512 KDSI
Organic	Plans & requirements(%)	6	6	6	6	
	Product design	16	16	16	16	
	Programming	68	65	62	59	
	Detailed Design	26	25	24	23	
	Code & unit test	42	40	38	36	
	Integration & test	16	19	22	25	
Semidetached	Plans & requirements(%)	7	7	7	7	7
	Product design	17	17	17	17	17
	Programming	64	61	58	55	52
	Detailed Design	27	26	25	24	23
	Code & unit test	37	35	33	31	29
	Integration & test	19	22	25	28	31
Embedded	Plans & requirements(%)	8	8	8	8	8
	Product design	18	18	18	18	18
	Programming	60	57	54	51	48
	Detailed Design	28	27	26	25	24
	Code & unit test	32	30	28	26	24
	Integration & test	22	25	28	31	34
Schedule Distribution		Size				
Mode	Phase	Small 2 KDSI	Intermediate 8 KDSI	Medium 32 KDSI	Large 128 KDSI	Very Large 512 KDSI
Organic	Plans & requirements(%)	10	11	12	13	
	Product design	19	19	19	19	
	Programming	63	59	55	51	
	Integration & test	18	22	26	30	

Schedule Distribution		Size				
Mode	Phase	Small 2 KDSI	Intermediate 8 KDSI	Medium 32 KDSI	Large 128 KDSI	Very Large 512 KDSI
Semidetached	Plans & requirements(%)	16	18	20	22	24
	Product design	24	25	26	27	28
	Programming	56	52	48	44	40
	Integration & test	20	23	26	29	32
Embedded	Plans & requirements(%)	24	28	32	36	40
	Product design	30	32	34	36	38
	Programming	48	44	40	36	32
	Integration & test	22	24	26	28	30

TABLE 8-7 RELY Maintenance Effort Multipliers

Very Low	Low	Nominal	High	Very High
1.35	1.15	1.00	0.98	1.10

TABLE 8-8 MODP Maintenance Effort Multipliers

Product Size (KDSI)	Rating				
	Very Low	Low	Nominal	High	Very High
2	1.25	1.12	1.00	0.90	0.81
8	1.30	1.14	1.00	0.88	0.77
32	1.35	1.16	1.00	0.86	0.74
128	1.40	1.18	1.00	0.85	0.72
512	1.45	1.20	1.00	0.84	0.70

Cost Drivers	Rating					
	Very Low	Low	Nominal	High	Very High	Extra High
Product Attributes						
RELY Required software reliability	.75	.88	1.00	1.15	1.40	
DATA Data base size		.94	1.00	1.08	1.16	
CPLX Product complexity	.70	.85	1.00	1.15	1.30	1.65
Computer Attributes	Very Low	Low	Nominal	High	Very High	Extra High
TIME Execution time constraint			1.00	1.11	1.30	1.66
STOR Main storage Constraint			1.00	1.06	1.21	1.56
VIRT Virtual machine volatility		.87	1.00	1.15	1.30	
TURN Computer turnaround time		.87	1.00	1.07	1.15	
Personnel Attributes	Very Low	Low	Nominal	High	Very High	Extra High
ACAP Analyst capability	1.46	1.19	1.00	.86	.71	
AEXP Applications experience	1.29	1.13	1.00	.91	.82	
PCAP Programming capability	1.42	1.17	1.00	.86	.70	
VEXP Virtual machine experience	1.21	1.10	1.00	.90		
LEXP Programming language experience	1.14	1.07	1.00	.95		
Project Attributes	Very Low	Low	Nominal	High	Very High	Extra High
MODP use of modern programming practices	1.24	1.10	1.00	.91	.82	
TOOL Use of software tools	1.24	1.10	1.00	.91	.83	
SCED Required development schedule	1.23	1.06	1.00	1.04	1.10	

Scaling Factors	Very Low	Low	Nominal	High	Very High	Extra High
Precedentness	6.20	4.96	3.72	2.48	1.24	0.00
Development Flexibility	5.07	4.05	3.04	2.03	2.03	0.00
Architecture/ Risk Resolution	7.07	5.65	4.24	2.83	1.41	0.00
Team Cohesion	5.48	4.38	3.29	2.19	1.10	0.00
Process Maturity	7.80	6.24	4.68	3.12	1.56	0.00

RCPX	0.73	0.81	0.98	1.0	1.30	1.74	2.38
RUSE	-	-	0.95	1.0	1.07	1.15	1.24
PDIF	-	-	0.87	1.0	1.29	1.81	2.61
PERS	2.12	1.62	1.26	1.0	0.83	0.63	0.50
PREX	1.59	1.33	1.12	1.0	0.87	0.71	0.62
FCIL	1.43	1.30	1.10	1.0	0.87	0.73	.620
SCED	-	1.43	1.14	1.0	1.0	1.0	



Answer the following:

Question 1:

(13 Marks)

- Explain what is meant by service primitives and parameters and their relation with confirmed services. How the concept of confirmed services has been implemented in the OSI model
- What is the relation between data rates, signal-to-noise ratio, signal levels in both noiseless & noisy channels. What is the appropriate signal-to-noise ratio and number of corresponding signal levels of a 1-MHz bandwidth channel to provide 6 Mbps?
- Write a taxonomy and classification of signal encoding techniques. Drive the relation between modulation rate and data rate for digital data to digital signal encoding. Compute the modulation rates if a Manchester encoding applied to 1 Mbps bit streams all of one's
- Compare between different guided transmission Media

Question 2:

(13 Marks)

- Briefly explain the following:
 - How DNS queries are resolved in the DNS system
 - The E-mail components, protocols, and message formats
- Explain how CSMA/CD works. Derive the condition of collision detection
A CSMA/CD network consists of the stations A, B, C, and D connected using a bus topology. The data rates is 10 Mbps, the distance between station A and C is 400 m, and the distance between C & D is 100 m, and the propagation speed is $4 \cdot 10^6$ m/sec. Station A starts sending a long frame at time $t_1 = 0$, station C starts sending a long frame at time $t_2 = 3 \mu\text{sec}$ station D starts sending a long frame at time $t_3 = 4 \mu\text{sec}$. The size of the frame is long enough to guarantee the detection of collision by both stations.
Draw a time sequence diagram for the network activities and find the number of bits station D and C sent before detecting the collision
- Drive an expression for the link utilization for stop and wait error free channel
Consider an error free selective repeat channel of 10 Mbps data rate and one-way propagation delay 20 msec transmitting frames of length 10000 bits and ACK of length 100 bits . *what will be the minimum number of sequence bits required for 100% utilization*

Question 3:

(12 Marks)

- Explain by drawing the changes made to IPv4 header to get the IPv6 header and the transition strategies to the next generation IP
- The FCIS of Ain Shams University is granted a block of addresses with the beginning address is 14.24.74.0 /24. The faculty needs to have 11 subnets as follows:
 - 2 subnets each with 64 addresses
 - 2 subnets each with 32 addresses
 - 3 subnets each with 16 addresses
 - 4 subnets each with 4 addresses*Design the subnets assuming that all 0s and all 1s subnets ID are allowed*
- Given that the router forwarding table is as given:

Mask	Network address	Next Hop	Interface
/26	180.7.65.192	-	M2
/25	180.7.65.128	-	M0
/24	201.4.22.0	-	M3
/22	201.4.16.0	-	M1
Default	Default	180.7.65.200	M2

How the router forward a packet with the destination address 180.7.65.140

Question 4:**(15 Marks)**

- a) Consider OSPF: explain briefly how it works, how LSA update, OSPF packet format, and its advanced features not included in traditional link state routing protocols
- b) TCP opens a connection using an initial sequence number 1400. The other party opens the connection with ISN of 4000
- 1- Draw time sequence diagram showing the connection establishment three segments
 - 2- Calculate retransmission time out value RTO after the connection establishment phase knowing that RTO initially was set to 6 seconds and the measured RTT after SYN+ACK segment was 1.5 seconds
 - 3- After connection establishment phase first party sends 100 bytes in one segment that is lost, and timed-out and retransmitted and acknowledged
 - a) Complete the time sequence diagram for the data transfer phase
 - b) Calculate the RTO for associated with these events
 - 4- Complete the time sequence diagram for the connection termination phase segments

- c) Explain briefly how BGP works. How BGP routes are selected

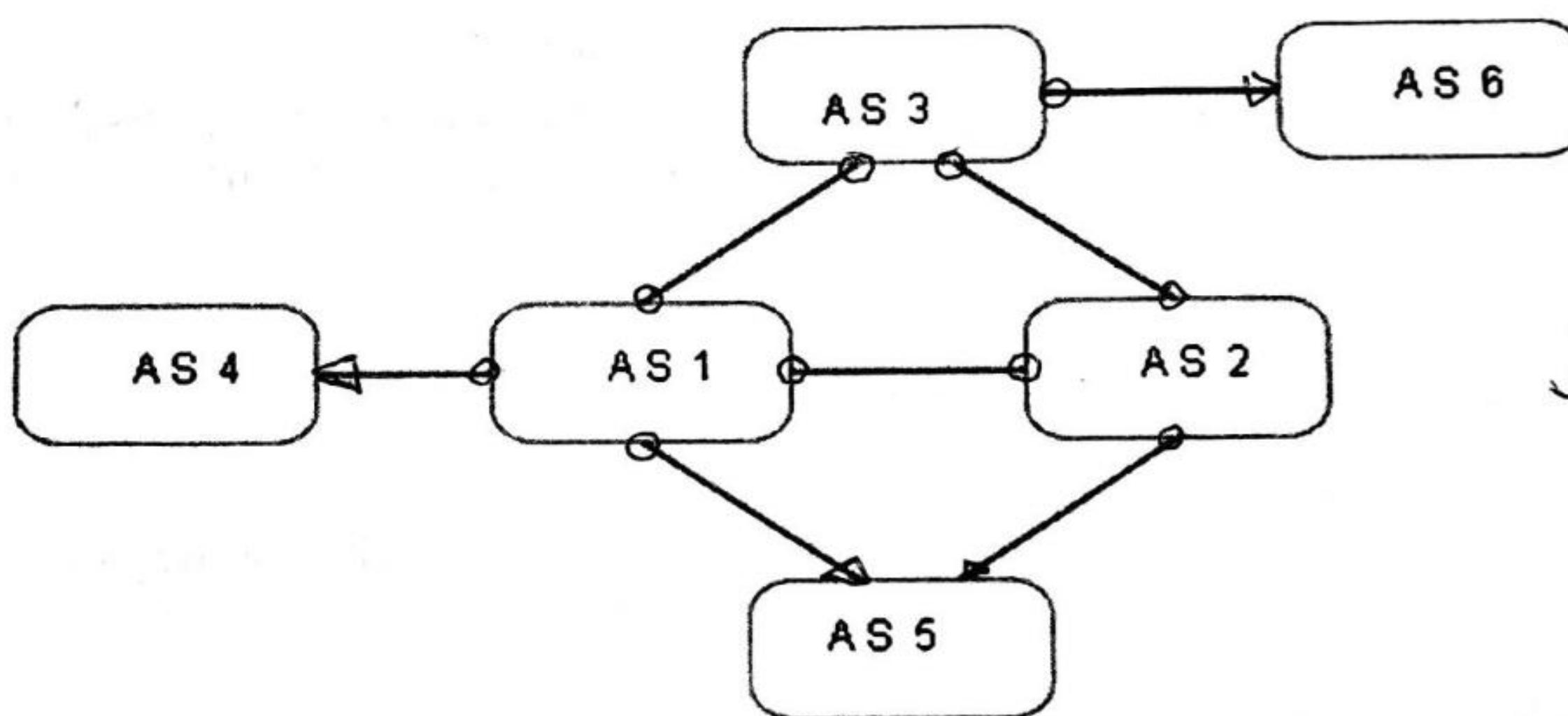
Consider a network with six autonomous systems with the following properties:

- AS1, AS2, AS3 are transit with peer relations between them
- AS4, AS5, AS6 are stub and customer to the provider AS1, AS2, AS3 as shown
- AS4 owns the prefix 10.0.1.0/24 and sends an advertisement to AS1 with the following attributes: 10.0.1.0/24, ORIGEN{AS4}, AS-PATH{AS4}

Explain how the other autonomous systems process and disseminate for the prefix 10.0.1.0/24.

Indicate which autonomous systems advertise the prefix to their neighboring autonomous systems.

Provide the ORIGIN and AS-PATH attributes used in the advertisement

**Question 5:****(12 Marks)**

- a) What is the role of checksum, sequence numbers, acknowledgement, and timers in TCP error control .Explain with the aid of drawing receiver actions on detecting lost and out-of-order segments
- b) Explain by drawing only the relation between send , receive buffers, send ,receive window, congestion window, flow control , congestion control. And explain how the windows sizes and sending rates controlled
- c) Write a pseudo code for a Reno TCP congestion control
- The *ssthresh* value for a Reno TCP station is set to 8 MSS . The station begins a connection in the slow-start with *cwnd* = 1 MSS and *ssthresh* = 8 MSS. Show the values *cwnd*, *ssthresh* , and the state of the station after the following events occur: 7 consecutive nonduplicate ACKs arrived, followed by a time out, followed by 5 consecutive nonduplicate ACKs , followed by 3 duplicate ACKs followed by 2 nonduplicate ACKs followed by a time out and finally 2 nonduplicate ACKs



Answer ALL the following FOUR questions:

(Total marks: 65)

1st Question

marks: 25

Please answer the following questions.

(5 marks each)

- Discuss the steps required in process migration in order to achieve logical mobility. What are the main advantages of process migration?
- What are the main challenges of Mobile Ad-hoc Networks? Clarify four. Which is more complex?
- How does the mobile IP work, in terms of discovery, registration, and tunneling activities, in order to transmit processed data to and from a mobile node? What enhancements have been then introduced to avoid triangular routing?
- Clarify the requirements for Mobile Agent Systems.
- Discuss the GPRS network operations provided by the GSM cellular networks.

2nd Question

marks: 16

Draw the following Diagrams.

(4 marks each)

- The MMS transaction flows diagram.
- The Bluetooth state transition diagram.
- The GPRS architecture.
- The Value Added Services architecture.

3rd Question

marks: 16

Clarify the main components and associated changes that have been added to the GSM architecture to allow the following services:

(2 marks each)

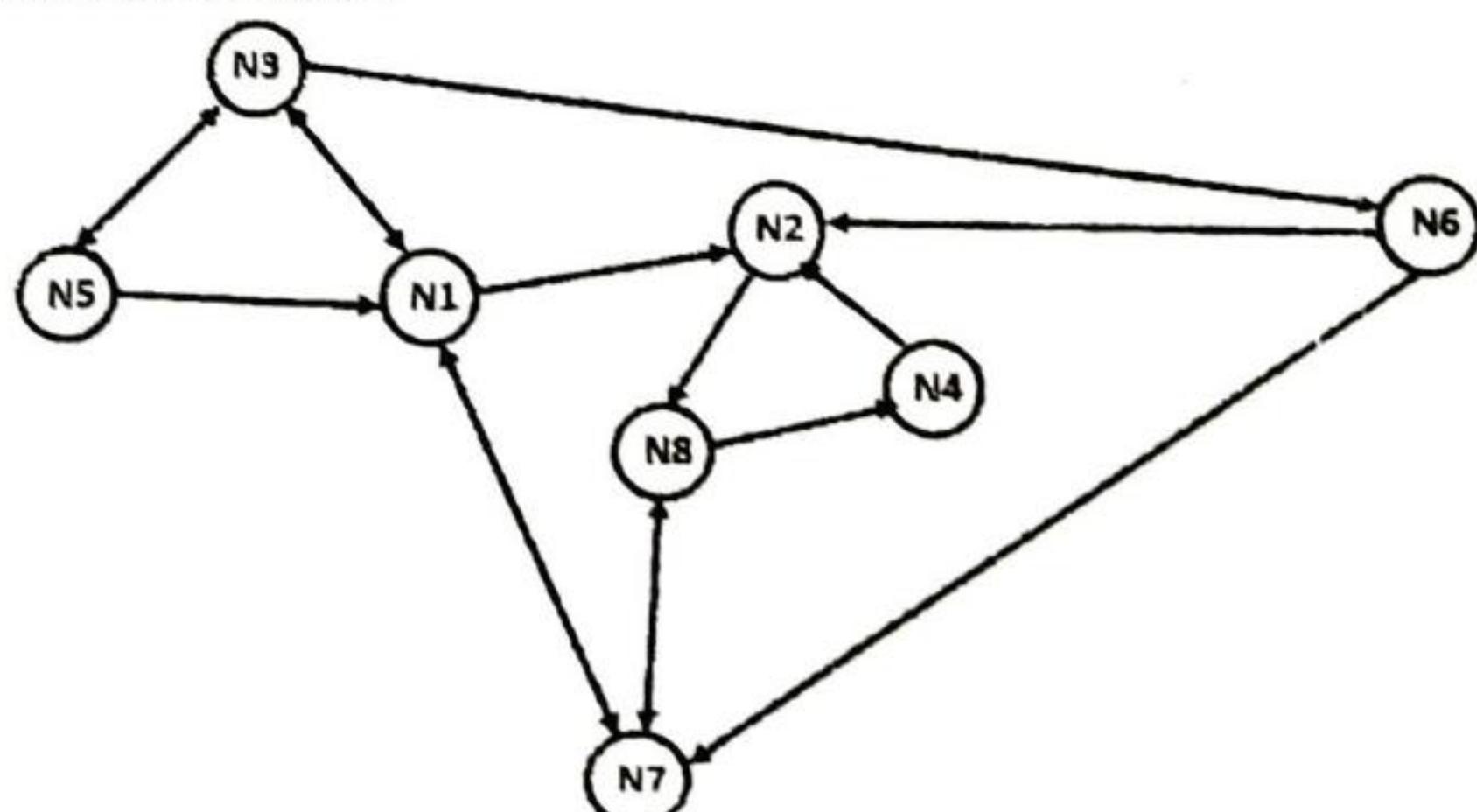
- General Radio Packet Services.
- Short Messaging Services.
- Wireless Application Protocol.
- Multimedia Messaging Services.
- Roaming.
- Value-added Services.
- Cellular IPs.
- Mobile TCPs.

4th Question

marks: 8

~~X~~ Given the following MANET topology at a certain instance of time:

- If the weight of the cost metric "minimum hop-count" is 0.6 and the weight of the cost metric "least interference" is 0.4, discover the optimum route(s) from N3 to N4 considering the two cost-metrics.
(5 marks)
- Create the routing table of N1 at this instance of time.
(3 marks)



With My Best Regards,

Dr. Sherin Moussa



(Total marks: 65)

Answer the following questions:

marks: 10

1st Question

- 1- Support-Vector Machines (SVMs) approach classification problem by finding the widest possible bar that fits between points of two different classes. (True/False)
- 2- In the use phase, k-means classifies new instances by finding the k most similar training instances and applying a combination function to the known values of their target variables. (True/False)
- 3- A maximal frequent itemset is always a closed itemset. (True/False)
- 4- An outlier can result from noise, but that is not always the case. (True/False)
- 5- SVM is particularly effective for categorical data compared to other techniques such as decision trees. (True/False)
- 6- KNN techniques are computationally efficient in the “use” phase of predictive modeling. (True/False)
- 7- Which of the following does not describe SVM (support vector machine)?
 - a- SVMs are based on supervised learning
 - b- SVM chooses the line to minimize the margin between two classes
 - c- SVM can be applied when the data are not linearly separable
- 8- Similarity measures are most essential for
 - a- Naïve Bayes
 - b- Tree Induction
 - c- Hierarchical Clustering
 - d- Logistic Regression
- 9- What is the confidence for the rules $\emptyset \rightarrow A$ and $A \rightarrow \emptyset$?
- 10-Suppose the confidence of the rules $A \rightarrow B$ and $B \rightarrow C$ are larger than some threshold, minconf. Is it possible that $A \rightarrow C$ has a confidence less than minconf?

2nd Question

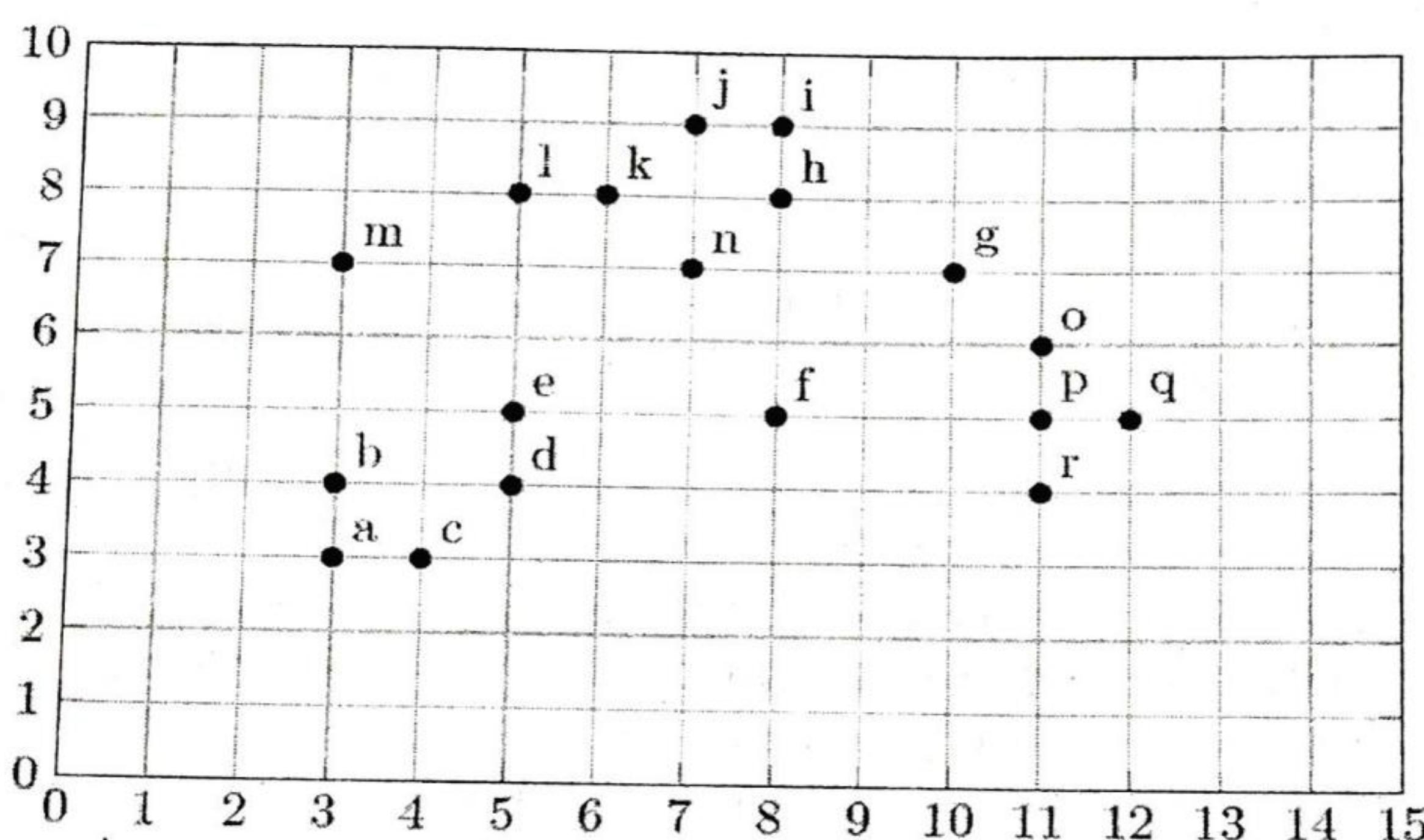
marks: 15

- 1- Suppose that the data mining task is to cluster points (with (x, y) representing location) into three clusters, where the points are
 $A_1(2, 10), A_2(2, 5), A_3(8, 4), B_1(5, 8), B_2(7, 5), B_3(6, 4), C_1(1, 2), C_2(4, 9)$.
The distance function is Euclidean distance. Suppose initially we assign A_1 , B_1 , and C_1 as the center of each cluster, respectively. Use the k-means algorithm to show only
 - a- The three cluster centers after the first round of execution.
 - b- The final three clusters.
- 2- Consider the following set of frequent 3-itemsets: $\{1, 2, 3\}, \{1, 2, 4\}, \{1, 2, 5\}, \{1, 3, 4\}, \{1, 3, 5\}, \{2, 3, 4\}, \{2, 3, 5\}, \{3, 4, 5\}$. Assume that there are only five items in the data set.
 - a- List all candidate 4-itemsets obtained by a candidate generation procedure using the $F_{k-1} \times F_k$ merging strategy.
 - b- List all candidate 4-itemsets obtained by the candidate generation procedure in Apriori.
 - c- List all candidate 4-itemsets that survive the candidate pruning step of the Apriori algorithm.

- 1- Consider the training examples shown in the following table for a binary classification problem.

Customer ID	Gender	Car Type	Shirt Size	Class
1	M	Family	Small	C0
2	M	Sports	Medium	C0
3	M	Sports	Medium	C0
4	M	Sports	Large	C0
5	M	Sports	Extra Large	C0
6	M	Sports	Extra Large	C0
7	F	Sports	Small	C0
8	F	Sports	Small	C0
9	F	Sports	Medium	C0
10	F	Luxury	Large	C0
11	M	Family	Large	C1
12	M	Family	Extra Large	C1
13	M	Family	Medium	C1
14	M	Luxury	Extra Large	C1
15	F	Luxury	Small	C1
16	F	Luxury	Small	C1
17	F	Luxury	Medium	C1
18	F	Luxury	Medium	C1
19	F	Luxury	Medium	C1
20	F	Luxury	Large	C1

- a- Compute the Gini index for the overall collection of training examples.
 b- Which attribute is better, Gender, Car Type, or Shirt Size?
 2- Consider the data in the Figure below. Answer to the following questions assuming that we are using Euclidean distance, that $\epsilon = 2$, and minpts = 3.



- a- List all the core points.
 b- Is r directly density-reachable from q?
 c- Is f density-reachable from a? Show the complete chain or where it breaks.
 d- Is g density-connected to r? Show the intermediate points that make them density-connected or that break the condition.

4th Question

marks: 15

- a- What are the major differences among three typical sequential pattern mining algorithms: GSP, PrefixSpan and SPADE?
- b- Suppose a sequence database contains three sequences of transactions as follows.

Seq-ID	Sequence
1	a (bc) d (ef)
2	(acd) (bc) g
3	(ab) d e g f

Suppose the minimum support is 2. What is the complete set of sequential patterns? What is the complete set of closed sequential patterns?

- c- Assume we have an association rule, if Drink_Tea and Drink_Coffee then Smoke that has a lift of 2. What does say about the relationship between smoking, and drinking coffee, and drinking tea? Moreover, the support of the above association rule is 1%. What does this mean?
- d- Explain the difference between Distance-Based vs. Density-Based Outlier Detection. Give an Example.
- e- How do k-nearest neighbor-based outlier detection techniques determine the degree to which "an object in a dataset is believed to be an outlier"
- f- Assume you run DBSCAN with MinPoints=6 and epsilon=0.1 for a dataset and we obtain 4 clusters and 20% of the objects in the dataset are classified as outliers. Now we run DBSCAN with MinPoints=5 and epsilon=0.1. How do expect the clustering results to change?



Answer all of the following FOUR questions:

(Total marks: 65)

1st Question

marks: 7

Complete the missing term(s) in the following statements.

Note: write only your answer in proper index and order and not the full statement

- a) In designing multimedia search engines, the paradigm is considered a subjective and time-consuming approach for search and retrieval.
- b) Examples for the descriptive features in Content-based retrieval (CBR) are: , and
- c) A Samsung Galaxy Tab S2 Tablet of screen size 2048 x 1536 has an aspect ratio =
- d) and are examples for file extensions for image and audio respectively.
- e) In the eye retina, receptors are light sensitive to the presence of high intense light and are responsible for colored vision.
- f) is a method that processes histograms by spreading the frequencies along the brightness scale, as a way for enhancing image contrast.
- g) is a mask matrix that stores transparency values in images.
- h) In Information Theory, the self-information is defined by the equation , where the maximum value for self-information occurs at a probability equal to
- i) Manipulating the image pixel values by division will make the image appear
- j) Image preprocessing is considered a/an level of processing. [low-intermediate-high]. choose the correct answer

2nd Question

marks: 18 [4,10,4]

Answer the following questions:

- (I) A pinhole camera has a focal length of 100 mm. If the object is located at a distance of 5 meters, at which distance would the image occur behind the lens? What is the displacement and direction that will be made by the lens to adjust a sharp image?
- (II) A gray-level image has a pixel depth value of 5. It is required to reduce the pixel depth to 3 using image quantization masking operations.
 - (a) What type of resolution are we changing?
 - (b) Write down the masks and the corresponding operations that can be used for the reduction operations.
 - (c) Referring to your answers in (b), mention the new mapped integer values for the color codes 15 and 21.
 - (d) Explain a common problem that might rise due to the pixel depth reduction?

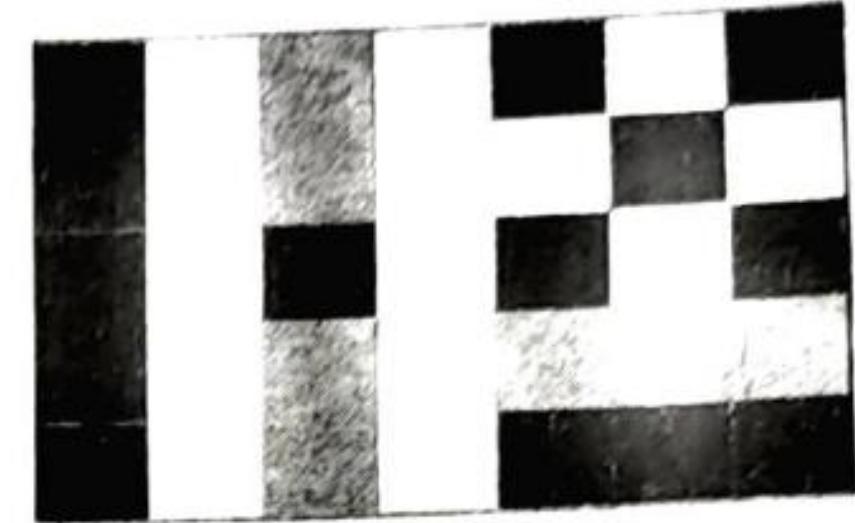
- (III) If the speech signal frequency occurs in the range between 300 to 4000 Hz. What is the minimum sampling rate for digitization? If the signal is quantized at 8 bits per sample, calculate the bitrate?

3rd Question

marks: 20 [14,6]

- (I) The given image array has the 3 gray-level intensity values (Black=0, gray=4, white=8):

- (a) Draw the image corresponding histograms (classic and cumulative).
(b) What is the image pixel depth?
(c) Which feature are we extracting by using histograms?
Mention the advantage and disadvantage of using such feature extraction method.
(d) Draw the equivalent GLCM (grey-level co-occurrence matrix), with the standard operator P = "one pixel right", and calculate the values of Energy
 $E = \sum_{k=0}^{n-1} x_k^2$ & Entropy $H = -\sum_{k=0}^{n-1} x_k \log_{10} x_k$.
(e) Which feature are we extracting by using GLCM?



- (II) For the same image in part (I), process the image using convolution with the filter

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

. Highlight the pixels changed.

4th Question

marks: 20

For the given 25-character message data: COMPRESSION_IS_IMPRESSIVE

- (a) Construct the Huffman tree to generate the Huffman codes for the symbols.
(b) Write down the code for each symbol. Are those codes generated uniquely decodable?
(c) Calculate the average symbol length after coding.
(d) Calculate the entropy for the source producing such symbol set.
(e) Calculate the Efficiency of the coding.
(f) What is the size of the coded data message in bits?
(g) What is the size of the message if NO compression is used with a standard 8-bit per symbol applied for encoding?
(h) Calculate the compression ratio.

Approximate your calculations to the nearest 2 decimal places.

With My Best Regards,

Dr. Sherine Rady