L-4/T-1/CSE **Date : 25/09/2022**

BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY, DHAKA

L-4/T-1 B. Sc. Engineering Examinations 2020-2021

Sub : **CSE 405** (Computer Security)

Full Marks : 210 Time : 3 Hours

The figures in the margin indicate full marks.

USE SEPARATE SCRIPTS FOR EACH SECTION

SECTION – A

There are **FOUR** questions in this section. Answer any **THREE**.

1. (a) What was the main vulnerability of WannaCry attacks? Why was it successful? Was it a zero-day attack? Why or why not? **(12)**

   (b) Suppose an Internet service provider (ISP) has voice over IP (VoIP) telephone system that it manages and sells. Suppose further that this ISP is deliberately dropping 25% of the packets used in its competitors' VoIP system when those packets are going through this ISP's routers. What kind of an attack is this? **(12)**

   (c) Why is compromise recording so important in computer security? What tools are used for such purpose? **(11)**

2. (a) What is role-based access control? Explain with a figure. What are its benefits and drawbacks? **(12)**

   (b) Consider the following security measures for airline travel. A list of names of people who are not allowed to fly is maintained by the government and given to the airlines; people whose names are on the list are not allowed to make flight reservations. Before entering the departure area of the airport, passengers go through a security check where they have to present a government-issued ID and a boarding pass. Before boarding a flight, passengers must present a boarding pass, which is scanned to verify the reservation. Show how someone who is on the no-fly list can manage to fly provided boarding passes can be printed online. Which additional security measures should be implemented in order to eliminate this vulnerability? **(12)**

   (c) Why should the forensic investigator make two copies of a digital evidence? What precautions are needed for this purpose? How is chain of custody maintained? **(11)**

3. (a) Why would it be bad to mix the stack and heap segments of memory in the same segment? Give one illustrative example how this mixture can be manipulated. **(12)**

   (b) How can you find out the vulnerabilities of a popular Web Server? How such vulnerabilities can be exploited? Give step by step approaches that can be followed with required tools you may use. **(12)**

   (c) Explain epidemic model for worm propagation with necessary figure. **(11)**

Contd P/2

CSE 405

4. (a) Explain possible ways how ARP spoofing can be prevented. Also, list the methods how a network administrator can monitor the network to identify the attacking host in the internal network. **(12)**

   (b) Why is it beneficial for an attacker's perspective to install an agent in the victim PC inside the corporate network? What methods can be used for such task? How can a network administrator prevent such intrusion attempt? Can the firewall help? Why or why not? **(12)**

   (c) What is base-rate fallacy? Explain with an example. How does it impact the incident handlers in their day-to-day activities? **(11)**

SECTION – B

There are **FOUR** questions in this section. Answer any **THREE**.

5. (a) Bob created a web app for a free streaming service. Eavesdropping is not an issue in this case. Therefore, he decides to use UDP and hence, cannot use TLS (and HTTPS). However, he needs to ensure data integrity, so that any other person cannot **impersonate** the sender and **alter** or **replay** his messages. How can you achieve this using Cookies? Explain your construction. **(15)**

   (b) Bob is writing a web server in C++. To avoid SQL injection attacks, he ensures one query statement runs exactly one SQL query in the database. Alice mentions that people can still do SQLi attacks with this, like changing password for everyone. Perform this attack in the following html form in Bob's website. How does the attack work? **(10)**

   Old Password [ ]

   New Password [ ]

   (c) Briefly describe four countermeasures against return-to-libc vulnerability. **(10)**

6. (a) In a social network website NS (`networksocial.com`), Trudy wants to make as many friends as possible. To do this, he needs to create a malicious website which he plans to name `trudymalicious.com`. This website should do the following once a person A enters: **(15)**

   (i) Send a friend request to Trudy (in NS)

   (ii) Set the homepage of A (in NS) such that any user visiting that page sends a friend request to trudy (in NS).

CSE 405

Contd ... Q. No. 6

Explain how Trudy can create this website in order to execute the attacks. Assume NS doesn't have security measures for CSRF or XSS.

The following GET request sends a friend request to Trudy.

```
http://www.networksocial.com/action/friends/add?friend=20
```

(b) Explain the Elgamal Signature Scheme. Why shouldn't you reuse the same random number, k, on two different signatures? **(10)**

(c) Using diagrams, show how you can figure out the placement of the argument for `system()` function while performing the return-to-libc attack. **(10)**

7. (a) Bob created a cipher that first applies a Caesar cipher and then a Hill cipher over the plaintext to get the ciphertext, which he calls a *CH network*. He thinks that if a substitution-permutation network satisfies Shannon's confusion and diffusion properties, so should a CH network. However, Alice mentions that it is very easy to break this cipher. Explain how a CH network can be broken. **(15)**

(b) The following cookies are all set by *docs.example.com*:

Cookie 1	Cookie 2	Cookie 3
name=file value=main domain=.example.com path=/ secure	name=file value=test domain=docs.example.com path=/test non-secure	name=file value=print domain=.example.com path=/print secure

For each of the following URLs list which cookies are available. **(10)**

(i) http://print.example.com/print

(ii) https://docs.example.com/

(iii) http://example.com/test

(iv) https://example.com/print/test

(v) https://docs.example.com/test/print

(c) What is StackGuard? Describe how it counters buffer overflow attacks. Use diagrams where necessary. **(10)**

8. (a) In the following C program try.c, figure out which attacks you can perform. Briefly explain how the attacks work in this program. Write down any assumption about the environment that you need to make to perform these attacks. **(15)**

CSE 405

Contd ... Q. No. 8

```
#include <stdio.h>
#include <unistd.h>

int main() {
   char *fn = "tmp/XYZ";
   char buffer[60];
   FILE *fp;

   scanf("%s", buffer);

   fp = fopen(fn, "a+");
   if (!access(fn, W_OK))
      fwrite(buffer, sizeof(char), strlen(buffer), fp);
   else
      printf("No permission");
   fclose(fp);

   return 0;
}
```

(b) 'A b-bit hash function provides b/2-bit security' – prove this statement. **(10)**

(c) Bob created a website where the authentication cookie expires after 1 year. If a user changes the password, it is updated in the database. Alice told Bob that while resetting the password he should reset the authentication cookie too. Explain the reason behind this suggestion. **(15)**

L-4/T-1/CSE **Date : 16/10/2022**

BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY, DHAKA

L-4/T-1 B. Sc. Engineering Examinations 2020-2021

Sub : **CSE 409** (Computer Graphics)

Full Marks : 210 Time : 3 Hours

The figures in the margin indicate full marks.

USE SEPARATE SCRIPTS FOR EACH SECTION

SECTION – A

There are **FOUR** questions in this section. Answer any **THREE**.

1. (a) Show that procedure of scan conversion of the curve, $x^2 + y^2 + 4x - 2y - 4 = 0$, including the steps of optimization with second order differences. **(25)**

   (b) If you have triangular pixels in a display device, what geometric shapes would you choose as filters for unweighted and weighted area sampling antialiasing techniques respectively and why? **(10)**

2. (a) Suppose a ray is cast towards (4, 5, 6) from a camera located at (2, 0, –1). Calculate its intersections with the geometric object expressed by the equation, $x^2 + y^2 + z^2 - 10x - 20y - 8z + 60 = 0$. **(25)**

   (b) Which hidden surface removal algorithm can you use to scan convert curved surfaces? Does it incur additional cost? Justify your answer. **(10)**

3. (a) Suppose you are designing a First-Person Shooter (FPS) game and you want to model the event of load-shedding at night time. In an FPS game, the player gets a view of what an actual person would see in the game and so, the view changes as he moves. Human eyes have mainly two different types of cells associated with vision. The rod cells work in very low light and create grayscale images while the cone cells work in sufficient light and create colored images. Usually one of the aforementioned two types of cells dominate the image generation process of our vision. If a suitably illuminated environment gets dark suddenly, our rod cells are activated and cone cells are deactivated. But this transition, take some time and we see absolute darkness in between, until our eyes get adjusted to the darkness. Similarly, if a dark environment gets illuminated suddenly, our rod cells are deactivated and cone cells are activated. This transition also takes some time and we keep blinking in between, until our eyes get adjusted to the lights. Now, devise a strategy based on your knowledge of illumination to model load-shedding in your FPS. You should consider both the situations when electricity goes off and comes back. **(20)**

   (b) Illustrate the following with appropriate 2D examples. **(15)**

   (i) Impact of camera position, view direction and the sequence of polygons considered in the construction of a BSP-tree.

   (ii) Impact of camera position and view direction in displaying polygons from an already constructed BSP-tree.

Contd P/2

CSE 409

4. (a) Show the steps of clipping the line, y =2|x| according to Cyrus-Beck line clipping algorithm. Assume the sides of the clip rectangle are axes parallel and the coordinates of its bottom-left and top-right corners are (–2,2) and (4,5) respectively. **(20)**

(b) Consider the fractional shown in the **Figure for Q. 4.2(b).** Calculate the dimension, length at n-th generation, and propose a string production rule of this fractal. **(15)**

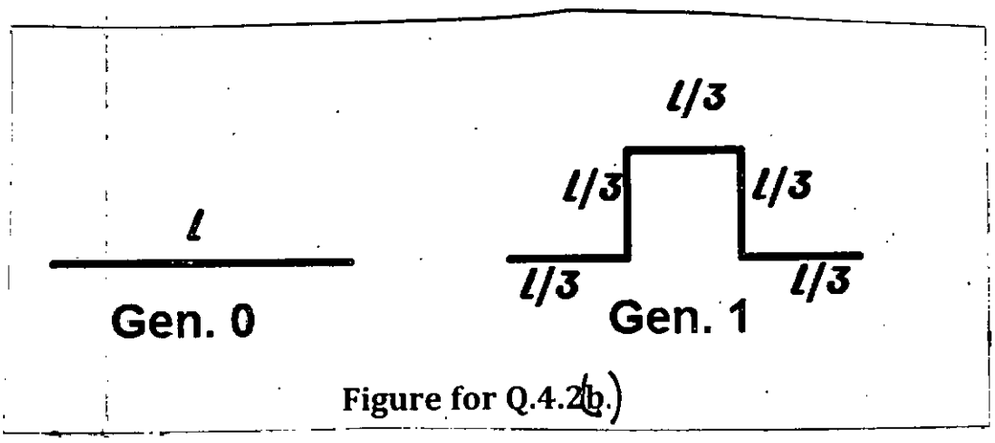

Figure for Q.4.2(b)

SECTION – B

There are **FOUR** questions in this section. Answer any **THREE**.

5. (a) Consider the equations below: **(5+5+2)**

$$P_1 : u + v + w + z = 6$$
$$P_2 : u + w + z = 4$$
$$P_3 : u + w = 2$$

Each of the equation represents a plane in a 4-dimensional space. Now, answer the following questions below:

(i) Determine whether the intersection of these three planes (P_1, P_2, P_3) is a line or a point or an empty set.

(ii) P_4 : u = –1; is another equation of a plane in a 4-dimensional space. Again, determine whether the intersection of all of these four planes (P_1, P_2, P_3 P_4) is a line or a point or an empty set.

(iii) Give a fourth equation of a plane in 4-dimensional space that leaves us with no solution with the initially given three planes (P_1, P_2, P_3)

(b) Given two equations of lines: **(12+5)**

$$L_1 : x = t + 2, y = t + 11, z = t - 2$$
$$L_2 : x = 5t - 8, y = 2t - 8, z = t + 2$$

Figure out the smallest distance these two lines, you **must** mention the points (one on each line) that are the closest point among all possible points of L_1 and L_2. Also, give a third line intersecting both lines at right angles.

(c) Define **affine** and **convex** combination of vectors. **(6)**

Contd P/3

CSE 409

6. (a) Look carefully at the given figures below: **(18)**

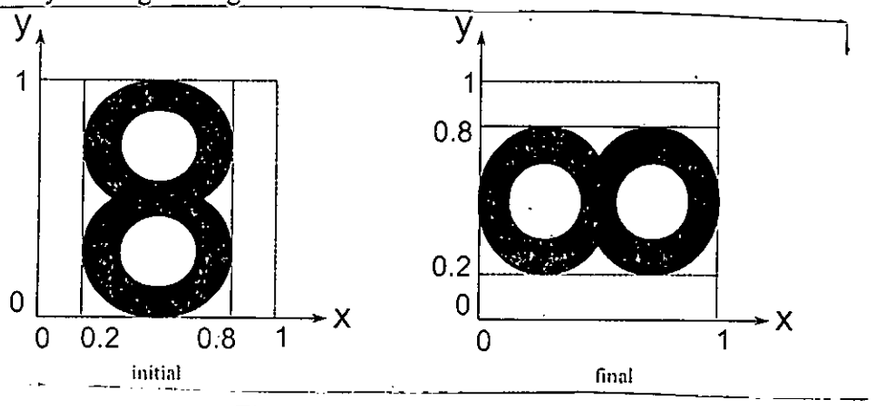

Figure for Question 6(a)

There are four different minimal composite transformations that can transform the figure from **initial** to **final**. You need to determine at least three of the composite transformations matrices. (You are required to express all the transformations by 3 × 3 homogeneous transformation matrices. You **don't** need to perform any matrix multiplication.)

(b) Determine the matrix for mirror reflection with respect to the plane passing through the point, $\mathbf{P_0(1, 2, 3)}$ and having a normal vector whose direction is N = 2i – 3j – k. The rotation matrix A_v that aligns vector V with z axis (vector k) can be calculated as follows. **(12)**

$$A_v = \begin{bmatrix} \frac{\lambda}{|V|} & \frac{-ab}{\lambda|V|} & \frac{-ac}{\lambda|V|} & 0 \\ 0 & \frac{c}{\lambda} & \frac{-b}{\lambda} & 0 \\ \frac{a}{|V|} & \frac{b}{|V|} & \frac{c}{V} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(c) State the differences between perspective and parallel projections. **(5)**

7. (a) The projection plane for perspective projection is given as follows: $\mathbf{P_{per}}(2, 4, -5)$ and $\mathbf{N_{per}}(3, -2, 1)$. The center of projection located at **COP**(– 1, 3, –2). Given a circle with center located at $C_0(3, 3, 3)$ and radius r = 3. Let, **L** be the equation of a straight line that passes through the center of the circle and intersects two distinct points of the circle. Let those intersection points be $\mathbf{C_1}$ and $\mathbf{C_2}$. Where, **L** is defined as **L**:(7, –1, 1) + t(–4, 4, 2). Measure the length of the line segment ($\mathbf{C_1}$ $\mathbf{C_2}$) after projecting (perspective projection) the circle into the given projection plane. **(20)**

(b) (i) Derive the general form of transformation matrix for an oblique projection onto the *xy* plane. You must show the calculation(s). Use figure(s) if necessary.

(ii) Using the transformation matrix you derived in (i), determine the transformation matrix for cabinet projection. You only need to calculate the final transformation matrix this time. No need to show explicit derivation. **(7+2)**

Contd P/4

CSE 409

Contd...Q. No. 7(c)

(c) Given the figure below, illustrate parallel projections for each of the viewing directions **A**, **B**, **C**. **(6)**

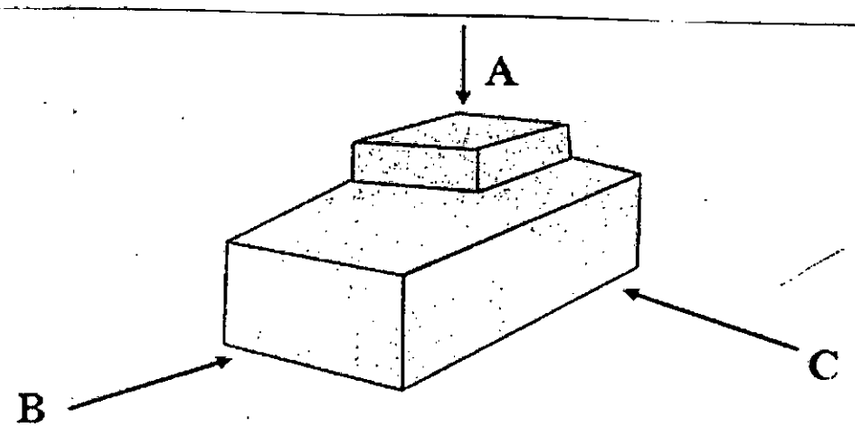

Figure for Question 7(c)

8. (a) Given two planes:

$$\mathbf{P}_1 : x + y + z = 1$$
$$\mathbf{P}_2 : x - 2y + 3z = 1$$

determine (i) the line of intersection and (ii) the angle between two planes $\mathbf{P}_1$ and $\mathbf{P}_2$. **(8+4)**

(b) A camera in $3D$ is located at point (2, 3, 1) it looks at point (4, 5, −5). The up direction of the camera is aligned with the negative Y axis. Derive the viewing transformation matrix. You **must** show all the calculation(s) necessary to derive the final viewing transformation matrix. **(12)**

(c) Consider the following matrix R in homogenous form that represents a $3D$ rotation.

$$R = \begin{bmatrix} \frac{2}{3} & \frac{-1}{3} & \frac{2}{3} & 0 \\ \frac{2}{3} & \frac{2}{3} & \frac{-1}{3} & 0 \\ \frac{-1}{3} & \frac{2}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

From the given rotation matrix R determine the axis and the angle of the rotation. You **must** show all the calculation(s) necessary. **(11)**

L-4/T-1/CSE **Date : 29/10/2022**

BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY, DHAKA

L-4/T-1 B. Sc. Engineering Examinations 2020-2021

Sub : **CSE 421** (Basic Graph Theory)

Full Marks : 210 Time : 3 Hours

The figures in the margin indicate full marks.

USE SEPARATE SCRIPTS FOR EACH SECTION

SECTION – A

There are **FOUR** questions in this section. Answer any **THREE**.

1. (a) Define the diameter, the radius and the center of a graph in terms of the eccentricities of vertices. Can you remember any other parameter of a graph which has relationship with the diameter of a graph? Explain the relationship. **(6+4)**

   (b) Prove or disprove: if a tree has a 1-factor, then 1-factor is unique. **(10)**

   (c) Construct the Prüffer's code for the tree in Fig. 1 showing every step. **(10)**

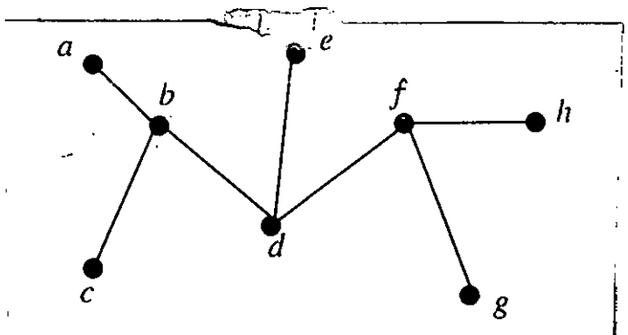

Figure 1: The tree for Question Number 1(c).

   (d) Why are we interested in counting the number of spanning trees in a graph? **(5)**

2. (a) Define the chromatic number and the chromatic index of a graph. Find chromatic numbers of P_5, W_5, K_5 and $K_{2,3}$. What is a k-critical graph? **(4+4+2)**

   (b) Let G be a bipartite graph with the maximum degree Δ. Then derive a proof for $\chi'(G) = \Delta$. **(10)**

   (c) Design an algorithm to find a 5-coloring of a planar graph. Analyze the time complexity of your algorithm. **(10)**

   (d) Design a graph theoretic model to preserve the scores of a round-robin tournament among 5 teams. Assume that no ties are allowed. **(5)**

Contd P/2

CSE 421

3. (a) What do you mean by a perfect matching? Explain Hall's matching condition with a positive example and a negative example. **(4+6)**

   (b) Show that the number of separating triangles in a triangulated plane graph of n vertices is at most $n - 4$. **(10)**

   (c) Design an efficient algorithm to find a maximum independent set in a chordal graph. **(10)**

   (d) Develop a graph theoretic model for designing a campus backbone network connecting the minimum number of routers. **(5)**

4. (a) Describe Kuratowski's theorem for planar graphs. Why is Petersen graph in Fig. 2 nonplanar? Explain using Kuratowski's theorem. **(4+6)**

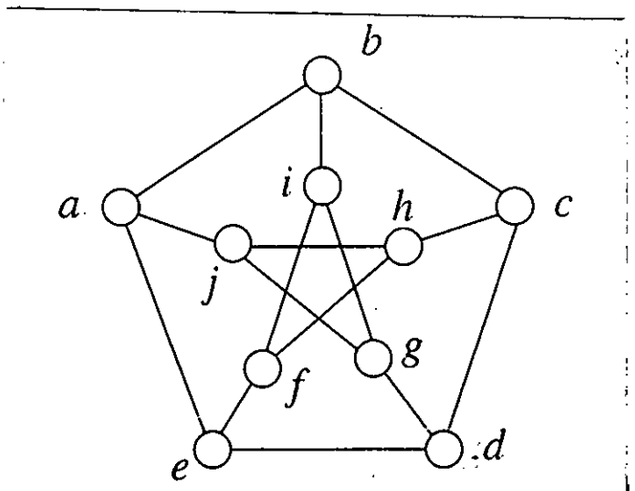

Figure 2: The graph for Question Number 4(a).

   (b) Derive a proof for the proposition: every maximal planar graph of four or more vertices has at least four vertices of degree five or less. **(10)**

   (c) Develop an algorithm for finding a straight line drawing of a planar graph? **(10)**

   (d) Explain a graph theoretic properties of a circuit which is related to the number of PCBs required for implementing the circuit? **(5)**

SECTION – B

There are **FOUR** questions in this section. Answer any **THREE**.

5. (a) What is the complement of a graph? Give an example of a self-complementary graph. **(4)**

   (b) A regular graph with an odd degree cannot have an odd number of vertices – Prove or disprove. **(6)**

Contd P/3

CSE 421

Contd ... Q. No. 5

(c) Suppose in a party, there are h hosts and g guests. Each person shakes hands with each other except that no host shakes hands with any other host. Transform the scenario into a graph problem and find the total number of handshakes. **(10)**

(d) Let G be a simple graph on n vertices. If G has k components, then prove that the number m of edges of G, satisfies $n - k \leq m \leq \frac{(n-k)(n-k+1)}{2}$. **(15)**

6. (a) An industry has 900 square meter rectangular area on a floor of a building where it needs to establish four processing units A, B, C, D and E. Processing units A, B and C require 100 square meter area each whereas D and E require 300 square meter each. Furthermore, the following adjacency requirements must be satisfied: A, B, C and E should be adjacent to D; A should be adjacent to B and D; B should be adjacent to D and C. Draw a graph with the given specification. Now construct a floor layout where the space for each processing unit will be a rectangle. Propose a suitable layout in your justification. **(10)**

(b) What is a cut edge? Suppose a connected graph G has n vertices. How many cut edges should G have if G has no cycle? **(2+5)**

(c) Prove that a simple graph G is a tree if and only if for each pair of vertices u and v of G, there is a unique (u, v) path in G. **(12)**

(d) Define binary tree. What is the minimum and maximum number of leaves a rooted tree with an odd number of vertices can have? **(6)**

7. (a) Give an efficient algorithm to check whether a graph is bipartite or not. **(10)**

(b) Define a complete bipartite graph. What is the maximum number of edges a complete bipartite graph with n vertices can have where n is even? **(2+6)**

(c) Define graph isomorphism with an example. Suppose you are studying a large community where few people know each other. You need to store this data and query time to time about whether two people know each other or not. Your primary concern is storage, that is, you want to use the least amount of memory as possible. Which graph representation will you use for this and why? **(10)**

Contd P/4

CSE 421

Contd ... Q. No. 7

(d) **(7)**

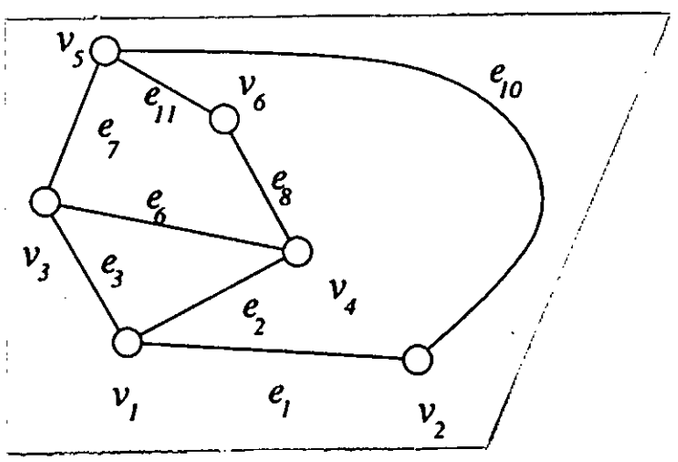

Figure for Q.7(d)

Select one edge in the above graph which if contracted results in a graph whose degree sequence is no more a graphic sequence. Draw the resultant graph ad write down its degree sequence.

8. (a) Define Euler path and Euler circuit with examples? Suppose a connected graph G is not Eulerian. With the help of the degree sum formula, find out the minimum number of edges we need to add in G to make it Eulerian. **(4+4)**

(b) If a graph is Hamiltonian then it cannot have any cut vertex. Prove or disprove. **(5)**

(c) Prove that a graph G of three of more vertices is 2-connected if and only if there are two internally disjoint paths between every pair of vertices in G. **(12)**

(d) **(10)**

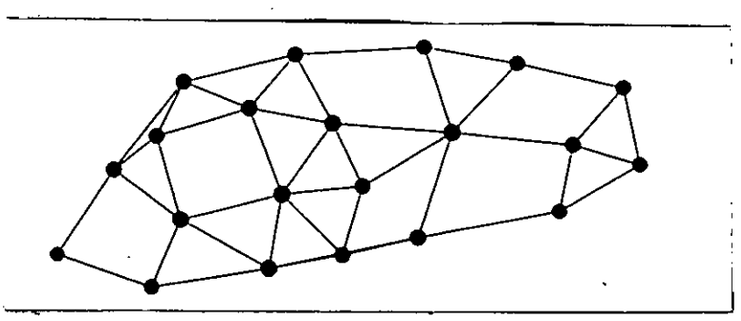

Figure for Q.8(d)

What is the relationship between connectivity and edge connectivity in a cubic graph? Find connectivity, edge connectivity and minimum degree of the above graph.

L-4/T-4/CSE **Date: 28/09/2022**

BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY, DHAKA

L-4/T-1 B. Sc. Engineering Examinations 2020-2021

Sub: **CSE 423** (Fault Tolerant System)

Full Marks: 210 Time: 3 Hours

USE SEPARATE SCRIPTS FOR EACH SECTION

The figures in the margin indicate full marks

SECTION – A

There are **FOUR** questions in this section. Answer any **THREE** questions.

1. (a) For hardening a memory system, a redundancy scheme needs to be applied with spare rows and columns. Design an efficient replacement scheme so that the spare rows and columns can be properly utilized to replace faulty portions of the memory system. Present necessary diagram(s) and example(s) to elaborate your design. **(20)**

   (b) Present and analyze - (i) an empirical defect-size distribution model, and (ii) survival probability of an electronic component. How the Burn-in testing be motivated from the survival probability of an electronic component? **(15)**

2. (a) Suppose a computation's running time T is twice the MTTF of the machine used to execute it. Ignoring the checkpointing time overhead, compare the probability of completing the computation in 2T time: **(20)**

   (i) Assuming no checkpointing, and

   (ii) Assuming checkpointing at regular intervals of T/2.

   (b) You are given a software module with detailed specifications and asked to harden the overall system. Develop methodologies to do so through utilizing the concepts of Recovery Block, N-Version Programming, and N Self-Checking Programming. **(15)**

3. (a) To escape from being a fail-hard system, a fail-soft system is designed that follows a state transition diagram of a 3-state repairable system. However, the system has an imperfect coverage. **(20)**

   Consider all failure rates to be λ and all repair rates to be μ. Besides, consider the coverage parameter to be C.

   With the above considerations, you need to formulate the steady-state probabilities of each of the states in the transition diagram.

   (b) What are the key aspects of robust electronic voting? What are the testing that we should conduct to make sure escaping defective vote cast in electronic voting? **(15)**

Contd P/2

CSE 423

4. (a) Draw the state transition diagram for a 2-state simple repairable system having a failure rate λ and a repair rate μ. To analyze the system, formulate time-dependent availability of the system with all necessary derivations. **(20)**

(b) For the purpose of defect avoidance, we may first need to identify the most vulnerable component(s) in a system. To do so, a generally-adopted approach is to utilize the notion of Fault Tree. **(15)**

You need to analyze how a Fault.Tree could facilitate identifying the most vulnerable component(s) in a system. Process necessary figures(s) and elaboration that are needed for the analysis.

SECTION – B

There are **FOUR** questions in this section. Answer any **THREE.**

5. (a) For a distributed data storing system having all-connected five nodes, analyze availability of single copy storing, data duplication, data triplication, and data dispersion. Here, consider the probability of a link or connection being active as 0.95 and the probability of a node being active as 0.99. **(20)**

Evaluate the trade-off among all these different types of data storing systems.

(b) There are two different production systems for UPS - System A and System B. System A produces 10000 units each day where 1% of the produced UPSs are defective. System B produces 20000 units each day where 1% of the produced UPSs are defective. **(15)**

Department of CSE, BUET purchases 50 units of UPSs from each of System A and System B. Between the two systems, which will have the higher probability to deliver 5 defective UPSs? Besides, for each of the systems, what is the probability of picking the 10th element of the purchased lot having 50 units as the first defective?

6. (a) What is meant by passive redundancy using majority voting? Analyze the different variants of TMR that can achieve passive redundancy. Evaluate applicability of a TMR from the perspectives of its advantages and challenges. **(20)**

(b) "Amdahl's Law is applicable for reliability while considering the performance metric Reliability Improvement Index, however, not for Reliability Improvement Factor" - Justify or refute this statement with necessary arguments and derivations. **(15)**

Contd P/3

CSE 423

7. (a) Distinguish between probability density function and instantaneous hazard rate. Differentiate different types of instantaneous hazard rates during lifetime of a typical computer system. **(20)**

Answer the above with necessary derivation(s) and figure(s).

(b) In a passenger plane, a failure of cabin pressurizing controlling system occurs in each 100000 hours on an average. Besides, a failure of Oxygen mask controlling system occurs in each 100000 hours on an average. The Oxygen mask system is required only when cabin pressurizing system fails. **(15)**

For the plane, you need to evaluate the fatality probability of a flight from Dhaka to London, which needs 10 hours flight time. Here, you need to consider the following separately for two different solutions-

(i) Failures of Oxygen mask system and cabin pressurizing system in the same hour result in fatality, and

(ii) Failures of Oxygen mask system and cabin pressurizing system in the same flight result in fatality.

8. (a) A manufacturing system has two components - X and Y. X and Y fail for 5 and 2 times respectively in each 10000 hours. The time required for repairing X is 100 hours. For Y, a periodic check occurs in each 1000 hours and it is made operational at the end of the checking period. The system fails if both x and Y fail. In such a case, an immediate repair takes place.

You need to analyze the system with a state transition diagram and need to determine the overall average failure rate of the manufacturing system from the state transition diagram. **(20)**

(b) Evaluate the impact of "coverage" for a system having one or more standby spare(s). Present all necessary derivations and graphs in the process of the evaluation. **(15)**

L-4/T-1 CSE

~~L-1/T-2/CSE~~ **Date: 26/10/2022**

BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY, DHAKA

L-4/T-1 B. Sc. Engineering Examinations 2020-2021

Sub: **CSE 453** (High Performance Database System)

Full Marks: 210 Time: 3 Hours

The figures in the margin indicate full marks

USE SEPARATE SCRIPTS FOR EACH SECTION

SECTION – A

There are **FOUR** questions in this section. Answer any **THREE** questions.

1. (a) Define strong consistency, weak consistency and eventual consistency models for data replication. **(9)**

(b) Explain the four Replica Update Management Strategies with illustrative figures. Briefly describe the pros and cons of each of them. **(12)**

(c) A company has several geographically distributed warehouses storing and selling products. Consider the following partial database schema: **(14)**

ITEM(ID, ItemName, Price, . . .)
STOCK(ID, Warehouse, Quantity, . . .)
CUSTOMER(ID, CustName, Address, CreditAmt, . . .)
CLIENT-ORDER(ID, Warehouse, Balance, . . .)
ORDER(ID, Warehouse, CustID, Date)
ORDER-LINE(ID, ItemID, Amount, . . .)

The database contains relations with product information (ITEM contains the general product information, STOCK contains, for each product and for each warehouse, the number of pieces currently on stock). Furthermore, the database stores information about the clients/customers, e.g., general information about the clients is stored in the CUSTOMER table. The main activities regarding the clients are the ordering of products, the payment of bills, and general information requests. There exist several tables to register the orders of a customer. Each order is registered in the ORDER and ORDER-LINE tables. For each order/purchase, one entry exists in the order table, having an ID, indicating the customer-id, the warehouse at which the order was submitted, the date of the order, etc. A client can have several orders pending at a warehouse. Within each order, several products can be ordered. ORDER-LINE contains an entry for each product of the order, which may include one or more products. CLIENT-ORDER is a summary table that lists, for each client and for each warehouse, the sum of all existing orders.

(i) The company has a customer service group consisting of several employees that receive customers' orders and payments, query the data of local customers to write bills or register paychecks, etc. Furthermore, they answer any type of requests which the customers might have. For instance, ordering products changes (update/insert) the CLIENT-ORDER, ORDER,

Contd P/2

CSE 453

Contd... for Q. No. 1(c)

ORDER-LINE, and STOCK tables. To be flexible, each employee must be able to work with any of the clients. The workload is estimated to be 80% queries and 20% updates. Since the workload is query oriented, the management has decided to build a cluster of PCs each equipped with its own database to accelerate queries through fast local access. How would you replicate the data for this purpose? Which replica control protocol(s) would you use to keep the data consistent?

(ii) The company's management has to decide each fiscal quarter on their product offerings and sales strategies. For this purpose, they must continually observe and analyze the sales of the different products at the different warehouses as well as observe consumer behavior. How would you replicate the data for this purpose? Which replica control protocol(s) would you use to keep the data consistent?

2. (a) Consider two relations R(A,B,C,D,E) and S(A,F,G,H). Assume there is a clustered index on attribute A for each relation. Assuming a database cluster with full replication, for each of the following queries, determine whether Virtual Partitioning can be used to obtain intraquery parallelism and, if so, write the corresponding subquery and the final result composition query. **(20)**

(i)
```
SELECT B, SUM(D), AVG(E)
FROM R, S
WHERE R.A=S.A
GROUP BY B
HAVING COUNT(*)>50
```

(ii)
```
SELECT B, MAX(D)
FROM R, S
WHERE C = (SELECT SUM(G) FROM R, S WHERE S.A=R.A)
GROUP BY B
```

(b) Distinguish between Uniform Memory Architecture (UMA) and Non-Uniform Memory Architecture (NUMA) for shared memory architectures with illustrative figures. Define Cache Coherent NUMA (CC-NUMA). **(8)**

(c) Distinguish between Shared disk (SD) and Shared nothing (SN) architectures for Parallel DBMS. **(7)**

3. (a) Define Big Data. Discuss the four characteristics of Big Data with examples. **(10)**

(b) Define MapReduce programming model. What are the functions of Mappers and Reducers? Illustrate MapReduce Processing framework and execution flow with illustrative figures. **(12)**

(c) What are the shortcomings of Map-Reduce programming model that are overcome by the Spark data processing framework and how does Spark overcome those. Describe Spark Program Flow with an illustrative flow chart. **(13)**

Contd P/3

CSE 453

4. (a) What are the limitations of relational DBMSs with SQL that led to the advent of NoSQL systems? **(7)**

(b) How is data partitioning done across nodes in DynamoDB? Explain with a figure. **(7)**

(c) Briefly describe the column family data model in Bigtable. Distinguish between hash partitioning and range partitioning for table creation. **(7)**

(d) How is NewSQL different from NoSQL and Relational SQL? **(7)**

(e) Compare the Google F1 NewSQL system with a standard parallel relational DBMS, e.g., MySQL Cluster, in terms of data model, query language and interfaces, consistency, scalability, and availability. **(7)**

SECTION – B

There are **FOUR** questions in this section. Answer any **THREE** questions.

5. (a) Given relation EMP as in Figure for Question 5(a), let p_1: TITLE <"Programmer" and p_2: TITLE > "Programmer" be two simple predicates. Assume that character strings have an order among them, based on the alphabetical order. **(20)**

(i) Perform a horizontal fragmentation of relation EMP with respect to $\{p_1, p_2\}$.

(ii) Explain why the resulting fragmentation (EMP_1, EMP_2) does not fulfill the correctness rules of fragmentation.

(iii) Modify the predicates p_1 and p_2 so that they partition EMP obeying the correctness rules of fragmentation.

EMP

ENO	ENAME	TITLE
E1	J. Doe	Elect. Eng.
E2	M. Smith	Syst. Anal.
E3	A. Lee	Mech. Eng.
E4	J. Miller	Programmer
E5	B. Casey	Syst. Anal.
E6	L. Chu	Elect. Eng.
E7	R. Davis	Mech. Eng.
E8	J. Jones	Syst. Anal.

ASG

ENO	PNO	RESP	DUR
E1	P1	Manager	12
E2	P1	Analyst	24
E2	P2	Analyst	6
E3	P3	Consultant	10
E3	P4	Engineer	48
E4	P2	Programmer	18
E5	P2	Manager	24
E6	P4	Manager	48
E7	P3	Engineer	36
E8	P3	Manager	40

PROJ

PNO	PNAME	BUDGET	LOC
P1	Instrumentation	150000	Montreal
P2	Database Develop.	135000	New York
P3	CAD/CAM	250000	New York
P4	Maintenance	310000	Paris

PAY

TITLE	SAL
Elect. Eng.	40000
Syst. Anal.	34000
Mech. Eng.	27000
Programmer	24000

Figure for Question 5(a)

Contd P/4

CSE 453

Contd... for Q. No. 5

(b) Distinguish between scale up and scale out High Performance Database architectures. **(8)**

(c) How do we measure the correctness of a fragmentation algorithm? Briefly describe each property. **(7)**

6. (a) Let $Q = \{q_1, \cdots, q_5\}$ be a set of queries, $A = \{A_1, \cdots, A_5\}$ be a set of attributes, and $S = \{S_1, S_2, S_3\}$ be a set of sites. The matrices below describe the attribute usage values and give the application access frequencies. Assume that $ref_i(q_k) = 1$ for all q_k and S_i and A_1 is the key attribute. Use the bond energy and vertical partitioning algorithms to obtain a vertical fragmentation of the set of attributes in A. **(30)**

	A_1	A_2	A_3	A_4	A_5
q_1	0	1	1	0	1
q_2	1	1	1	0	1
q_3	1	0	0	1	1
q_4	0	0	1	0	0
q_5	1	1	1	0	0

(a)

	S_1	S_2	S_3
q_1	10	20	0
q_2	5	0	10
q_3	0	35	5
q_4	0	10	0
q_5	0	15	0

(b)

(b) Draw the mediator/wrapper architecture for multi-database system. **(5)**

7. (a) Consider the sample database **(20)**

EMP(ENO, ENAME, TITLE)

ASG(ENO, PNO, RESP, DUR)

PROJ(PNO, PNAME, BUDGET)

Assume that relations PROJ and ASG of the sample database are horizontally fragmented as follows:

$PROJ_1 = \sigma_{PNO \leq "P2"}(PROJ)$

$PROJ_2 = \sigma_{PNO > "P2"}(PROJ)$

$ASG1 = \sigma_{PNO \leq "P2"}(ASG)$

$ASG2 = \sigma_{"P2" < PNO \leq "P3"}(ASG)$

$ASG3 = \sigma_{PNO > "P3"}(ASG)$

Transform the following query into a reduced query on fragments, and determine whether it is better than the fragment query:

```
SELECT RESP, BUDGET
FROM ASG NATURAL JOIN PROJ
WHERE PNAME = "CAD/CAM"
```

Contd P/5

CSE 453

Contd... for Q. No. 7

(b) Consider the join graph of the Figure for Question 7(b) and the following information: size (EMP) = 100, size(ASG) = 200, size(PROJ) = 300, size(EMP ⋈ ASG) = 300, and size(ASG ⋈ PROJ) = 200. Describe an optimal join program for joining the three relations based on the objective function of total transmission time. **(8)**

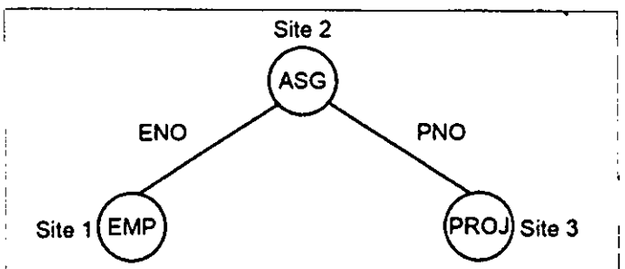

Figure for Question 7(b)

(c) Describe 2-phase locking protocol. How can 2-phase locking protocol ensure conflict serializability of transactions. Explain with a figure. **(7)**

8. (a) Consider the following transaction histories of three transactions in four nodes: **(20)**

$H_1 = \{W_2(x); W_1(x); R_3(x); R_1(x); W_2(y); R_3(y); R_3(z); R_2(x)\}$

$H_2 = \{R_3(z); R_3(y); W_2(y); R_2(z); W_1(x); R_3(x); W_2(x); R_1(x)\}$

$H_3 = \{R_3(z); W_2(x); W_2(y); R_1(x); R_3(x); R_2(z); R_3(y); W_1(x)\}$

$H_4 = \{R_2(z); W_2(x); W_2(y); W_1(x); R_1(x); R_3(x); R_3(z); R_3(y)\}$

(i) Which of the above histories are conflict equivalent?

(ii) Which of the above histories are serializable?

(b) Describe the two-phase commit protocol for distributed atomic commitment of transactions. How can a transaction be blocked under two-phase commit? How can we overcome the problem of blocking? **(8)**

(c) Briefly describe Distributed Deadlock Detection algorithm with illustrative figures. **(7)**

L-4/T-1/CSE Date: 20/10/2022

BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY, DHAKA

L-4/T-1 B. Sc. Engineering Examinations 2020-2021

Sub : **CSE 463** (Introduction to Bioinformatics)

Full Marks : 210 Time : 3 Hours

The figures in the margin indicate full marks.

USE SEPARATE SCRIPTS FOR EACH SECTION

SECTION – A

There are **FOUR** questions in this section. Answer any **THREE**.

1. (a) Briefly describe splicing and alternative splicing. **(10)**

   (b) With necessary diagrams briefly describe the process of DNA replication in living cells. In this context explain what "leading strand", "lagging strand", and "okazaki fragment" are. **(10)**

   (c) What do you understand by circadian clock? How is the circadian clock controlled by LCY, CCA1 and TOC1 genes in plants? **(10)**

   (d) What is the reverse complement of the genomic sequence "ACAGCGTGGC"? **(5)**

2. (a) Write down the pseudocode for randomized motif search algorithm. Explain the algorithm with a suitable example. **(25)**

   (b) State and prove the cycle theorem. **(10)**

3. (a) Formulate the *Soft Decoding Problem*. Solve it using the forward-backward algorithm. State relevant equations, draw necessary diagrams, and provide necessary explanations. **(20)**

   (b) Describe the steps of *Viterbi Learning* for HMM parameter estimation. **(10)**

   (c) Describe, with necessary diagrams, the formation process of the Philadelphia chromosome. **(5)**

4. (a) Compute the prefix array (used in KMP algorithm) for the pattern P = *ababaaba*. Explain with an example how the prefix array helps in improving pattern matching running time. **(10)**

   (b) What is a suffix array? Construct the suffix array for the string CGACTGTG$. From the suffix array, construct the Burrows-Wheeler transform (BWT). Using these 2 data structures, identify whether the pattern "GAC" exists in the original string and if so, then in which position. Properly explain different steps in this process. **(25)**

Contd P/2

CSE 463

SECTION – B

There are **FOUR** questions in this section. Answer any **THREE**.

5. (a) What do homologous sites mean? Suppose you are going to permute the sites (columns) of a given multiple sequence alignment (MSA). For example, Fig. 1 shows an MSA and Fig. 2 is the permuted MSA where the 2nd and the 4th sites have been interchanged (highlighted in grey). If you construct two phylogenetic trees from these two alignments using a particular method, are they going to be the same or different? Briefly justify your answer. **(12)**

A	T	C	C	A	T	G
A	-	C	A	-	T	T
A	T	-	C	A	-	G
A	-	A	C	A	C	G

Fig. 1: A multiple sequence alignment (MSA)

A	C	C	T	A	T	G
A	A	C	-	-	T	T
A	C	-	T	A	-	G
A	C	A	-	A	C	G

Fig. 2: The MSA with 2nd and 4th site (column) interchanged

(b) Find the optimal local alignment of the following two sequences using -2 as the gap penalty, -1 as the mismatch penalty, 2 as the score for a match. You have to use the Smith-Waterman algorithm, and show the corresponding dynamic programming (DP) table. Please show the alignment and mark the path that corresponds to the alignment in the DP table. **(15)**

CCAGTGC

CTAT

(c) Briefly discuss the progressive alignment technique, its drawbacks and how these drawbacks can be overcome. **(8)**

6. (a) For a given gene tree gt = (((((b,a),c),d),e),f), and a species tree ST = ((((a,(b,c)),e),f),d), explain the discordance between gt and ST using:

   (i) Deep coalescence (DC)

   (ii) Gene duplications and losses (GDL).

You have to find the optimal reconciliations, meaning that the number of extra lineages and duplications and losses should be minimized.

Contd P/3

CSE 463

Contd...Q. No. 6(a)

Please show the reconciliations with appropriate figures (separate figures for DC and GDL) indicating the deep coalescence, duplication, and loss events. Report the numbers of extra lineages, duplications and losses. **(20)**

(b) Suppose you are trying to construct a species tree on 15 different species. You have sampled 200 genes from each of these 15 species. Your supervisor has asked you to use a method called GT-est for constructing trees from sequence alignments, and SP-est (which is a summary method) for estimating species tress from gene trees. **(7)**

(i) How many times do you need to run GT-est and SP-est to estimate a species tree by summarizing gene trees?

(ii) How many times do you need to run GT-est and SP-est to estimate a species tree using "Combined Analyses"?

(c) Given an unrooted tree T with n taxa, how many neighboring trees you can generate that are one Nearest Neighbor Interchange (NNI) move away from T? **(8)**

7. (a) Given a true and an estimated tree, False Negative (FN) rate can be defined as the proportion of the missing edges (ratio between the total numbers of FN edges in the estimated tree to the total number of internal edges in the true tree). Similarly False Positive (FP) rate is the ratio between the total numbers of false positive edges in the estimated tree to the total number of internal edges in the true tree. If we allow the estimated tree to be non-binary, FP rate suffers from a major limitation. What do you think is the major limitation of FP rate when estimated trees are non-binary? Briefly justify your answer. **(10)**

(b) (i) Consider the following distance matrix on four taxa. Is this additive? If this is additive, draw the corresponding unrooted tree with branch lengths.

	a	b	c	d
a	0	4	5	4
b		0	3	4
c			0	5
d				0

(ii) Consider the following set *CS* of clusters: {bc, fg, abc, dfg, edfg, bcdefg}. Trivial clusters (i.e., the clusters containing only one taxon and the cluster with all the taxa) are not included in *CS*. Is this a compatible set of clusters? If yes, show the corresponding tree. If not, explain why. **(6+6=12)**

(c) Construct a tree on the leaf set {a, b, c, d, e} which is consistent with each of the following tiplets. You have to show the intermediate steps of your algorithm. **(13)**

Contd P/4

CSE 463

Contd...Q. No. 7(c)

a|bc	a|de
a|bd	b|cd
a|be	b|ce
a|cd	b|de
a|ce	c|de

8. (a) Consider the following five reads. Construct the overlap graph with minimum overlap 2 (meaning that there is an edge between two reads if the overlap between them is at least two). Find the shortest common superstring from the overlap graph. **(18)**

ATGCGTTA

GGTTAT

TAGATGG

TTAGAT

ATAGAT

(b) Construct the De Bruijn graph from the following read using 5-mers. **(10)**

ATGCATTTGCATGCG

(c) Suppose you are going to assemble a genome of a species which is believed to have a lot of repetitive regions. Which of the following options do you think would be the best for assembling the genome of this species? Briefly justify your answer. **(7)**

Options:

(i) Long reads and De Bruijn graph

(ii) Short reads and De Bruijn graph

(iii) Long reads and Overlap graph

(iv) Short reads and Overlap graph

BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY, DHAKA

L-4/T-1 B. Sc. Engineering Examinations 2020-2021

Sub : **HUM 475** (Engineering Economics)

Full Marks : 210 Time : 3 Hours

The figures in the margin indicate full marks.

Symbols have their usual meaning.

USE SEPARATE SCRIPTS FOR EACH SECTION

SECTION – A

There are **FOUR** questions in this section. Answer any **THREE**.

1. (a) Clarify the concept of diminishing marginal utility. Prove that consumer equilibrium conditions are identical in the Cardinalist approach and in the Ordinalist (Indifference curve) approach. **(20)**

   (b) 'People buy less when commodity price rises' – explain this phenomenon with the help of income effect and substitution effect and illustrate these two effects with numerical and graphical presentations. **(15)**

2. (a) Which factors would you consider constructing a comprehensive demand function for CC TV camera in Bangladesh? Give reasons in favour of these factors. Explain how you would construct a market demand curve for CC TV camera. **(20)**

   (b) What do you understand by 'change in demand' and 'change in quantity supplied'? The demand and supply functions of petroleum are given respectively

   $$Q_{pd} = 4550 - 20P_p \text{ and}$$
   $$Q_{ps} = 1550 + 12P_p$$

   Where, Pp is the price of petroleum. Find the equilibrium price and quantity of petroleum. If the Government provides 27% subsidy on price, what would be the new equilibrium price and quantity? What is the proportion of the subsidy that the consumers would enjoy? **(15)**

3. (a) What do you understand by revenue of a firm? Draw the average and marginal revenue curves of a firm facing a downward sloping demand curve. Describe the conditions for profit maximization of a firm. The average revenue (AR) and total cost (TC) functions of a firm are given by

   $$AR = 4350 - 13M$$
   $$TC = M^3 - 5.5M^2 + 150M + 675$$

   Where M refers to quantity of output.

   Find the profit maximizing level of output and maximum profit. **(20)**

   (b) What is monopolistic competition? Illustrate the short run and long run equilibrium of the firm under monopolistic competition. Evaluate the long run equilibrium between perfect competition and monopolistic competition. **(15)**

Contd P/2

HUM 475(CSE)

4. Write short Notes on any THREE of the following: **(35)**

   (a) The main challenges that every economy struggles to overcome

   (b) Optimum combination of factors of production

   (c) Oligopoly

   (d) Determinants and range of values of elasticity of demand.

SECTION – B

There are **FOUR** questions in this section. Answer any **THREE**.

5. (a) What is double counting problem in measuring GDP? How can we overcome it? **(10)**

   (b) Do you think that GDP is a good measure of economic well-being of a country? Why? **(10)**

   (c) Given the data in the following table: **(15)**

Market Basket	Prices in 2018	Prices in 2019	Prices in 2020	Prices in 2021
5 Pens	$2	$2.5	$4	$5
10 Eggs	$5	$7	$8	$8.25
15 Drinks	$1	$2	$3	$4

   Calculate consumer price index (CPI) in year 2018, 2019, 2020, and 2021 considering 2019 as base year. And calculate the inflation rate between 2018 & 2019, 2019 & 2020, and 2020 & 2021.

6. (a) Distinguish between economic growth and economic development. **(15)**

   (b) "Development is a multidimensional concept." Explain the statement. Given the following goal posts, **(20)**

Indicator	Assigned Maximum	Assigned Minimum
Life Expectancy (Years)	83.2	20
Mean Years of Schooling	13.2	0
Expected Years of Schooling	20.6	0
Per Capita National Income (PPP $)	108211	163

   Find the Human Development Index (HDI) of a hypothetical country if it has life expectancy of 72.60, mean years of schooling of 6.7, expected years of schooling of 10.2, and per capita national income of $5320 in PPP. Comment on the nation's level of human development.

Contd P/3

HUM 475(CSE)

7. (a) Briefly explain the different phases of project evaluation. **(15)**

   (b) Suppose a company has the following two possible investment opportunities. According to the net present value (NPV) criterion which project is more preferable and why if the market interest is 12%. **(12)**

Year	Cash Flow from Project A (Tk.)	Cash Flow from Project B (Tk.)
0	(10,00,0000)	(10,00,000)
1	1,00,000	3,00,000
2	2,00,000	3,00,000
3	6,00,000	3,00,000
4	3,00,000	3,00,000
5	3,00,000	3,00,000

   (c) Calculate the internal rate of return (IRR) of an investment project which involves a current outlay of Tk. 300,000 (three lakhs) and results in an annual cash inflow of Tk. 60,000 (sixty thousands) for 7 years. **(8)**

8. (a) Define short run and long run with reference to theory of production. Explain the law of diminishing marginal returns and show the short run cost curves in production with graphical presentations. **(15)**

   (b) How would you construct the long run average cost curve (LRAC)? Why is the curve called an envelope curve? **(10)**

   (c) What do you know about scale of production? Explain the reasons behind increasing returns to scale? **(10)**
