KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY
B.Sc. Engineering 3rd Year 1st Term Examination, 2024
Department of Computer Science and Engineering
CSE 3101
Operating Systems

TIME: 3 hours FULL MARKS: 210

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

SECTION A

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

1. a) Including terms, e.g., 'resource allocator', 'control program', 'kernel', etc., discuss the definition of the Operating System (OS). Discuss how bootstrap program loads OS kernel. (10) [CO1]
   b) Discuss benefits of virtual memory system. Suppose you have reference string: 1, 2, 3, 4, 1, 2, 5, 1, 2, 3, 4, 5, and 4 frames. Depict the scenarios in case of FIFO, Optimal, and LRU algorithms. (13) [CO2]
   c) How Domain Structure and Domain Implementation are organized? Discuss about Access Matrix Implementation. (12) [CO3]

2. a) "Dynamic loading and dynamic linking affect the main memory" – how? Discuss. Using necessary figures, show how "dynamic storage allocation problem" can be solved? (14) [CO2]
   b) Define and discuss the purpose of page table. How paging hardware is implemented with TLB? Depict the necessary diagram also. (13) [CO3]
   c) Explain the failure detection system in distributed environment. (08) [CO3]

3. a) In case of multiple resource instances, explain data structures to manage deadlock avoidance. Also using Safety Algorithm, show how to decide that a system is safe or unsafe. (12) [CO2]
   b) Explain the purpose of "wait-for-graph" in deadlock detection. (07) [CO3]
   c) Discuss various categories of security violations. "Attackers can make a computer system vulnerable through program threats" – how? Discuss various program threats. (10) [CO3]
   d) Explain distributed file system structure. (06) [CO1]

4. a) Differentiate between network operating systems and distributed operating systems. Explain various migration happening in distributed operating systems. (12) [CO1]
   b) Make a comparative study among various routing strategies required for distributed communication structure. (08) [CO1]
   c) "There are distinctions between stateful and stateless services" – clarify them. (08) [CO2]
   d) "A Boot-sector Computer Virus is enough to make the OS out-of-order" – explain how it would be possible. (07) [CO3]

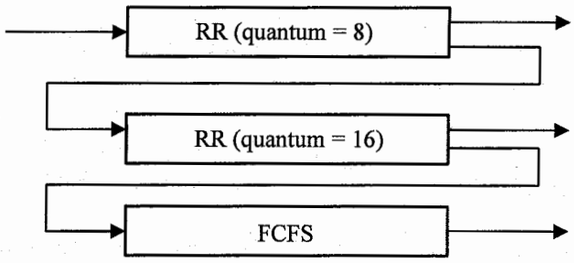
SECTION B

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

5. a) Suppose, there are two processes named A and B. How does MS-DOS OS execute the two processes? Briefly describe the steps. (10) [CO4]
   b) Discuss about your choice of a parameter passing method in a system call design. (10) [CO1]
   c) Say, you are tasked with designing an OS for a computer system, where many interactive processes may run concurrently, on a single core processor. Additionally, a process in this system may have many multiple concurrent worker threads in them. Decide and comment on the information you may choose to include in the process control block (PCB). Discuss about the strength and weakness of your PCB design choice in terms of system performance. (15) [CO2]

6. a) "Disabling interrupts to achieve mutual exclusion in inter process communication is a pretty bad idea" – do you agree with the statement? Explain your answer. (10) [CO2]
   b) In a multithreaded system, when a thread calls "`fork()`", how does the system decide between creating a copy of the whole process and only the calling thread? (10) [CO2]
   c) Consider the following systems: (i) Air traffic control, (ii) communication system like voice over IP, (iii) medical system, and (iv) multimedia transmission and reception. To develop them, what will you use among hard and soft real time systems and why? (15) [CO4]

7. a) "In paging, the chosen frame location is irrelevant in placement policy" – do you agree with this statement? Justify your answer. (10) [CO1]
   b) "Using virtual address space greatly helps multitasking features at the cost of decline in system performance" – justify the statement. (10) [CO1]
   c) Suppose you devised a strategy for CPU scheduling using two priority level for foreground (priority = 0) and background (priority = 1) tasks. The foreground tasks are scheduled using a multilevel feedback queue describe by the following figure. The background tasks are scheduled with FCFS algorithm. Say, several processes arrive with the description given in the following Table. Calculate and show the Gantt chart of the process scheduled by your said algorithm. Note that, lower priority value denotes higher priority. (15) [CO3]



Process	Arrival Time	Burst Time	Priority
P1	0	32	0
P2	4	4	0
P3	16	30	0
P4	28	10	1
P5	30	5	1

8. a) Discuss the trade-off between prioritizing load balancing and processor affinity in CPU scheduling for a multiple processor computer system. (10) [CO3]
   b) Describe the reasonings behind a fixed size boot block in every partition of a disk regardless of the fact that it may or may not contain an OS. (10) [CO1]
   c) You are given an array of integers of size *N*. You are asked to add these numbers in parallel with *M* threads. Then implement it using the `Pthread` library. The function name must be "`Summation`" which will be used to do the task. (15) [CO4]

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY
B.Sc. Engineering 3rd Year 1st Term Examination, 2024
Department of Computer Science and Engineering
CSE 3105
Embedded Systems and Internet of Things

TIME: 3 hours FULL MARKS: 210

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

SECTION A

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

1. a) What are the distinguishable characteristics of an embedded system? (07) [CO1]
   b) How can embedded systems help human being by working in harsh environment? (06) [CO1]
   c) "Adding designers can decrease individual productivity and at some point can actually lengthen the project completion time" – justify the statement mathematically with suitable examples. (12) [CO3]
   d) What are the different types of IC technologies for designing an embedded system? Show the comparisons among them. (10) [CO1]

2. a) What are the challenges that are faced in designing embedded systems? (07) [CO1]
   b) Illustrate the ideal top-design process of embedded systems with a neat sketch. (09) [CO4]
   c) Write short notes on: (i) ASSP, and (ii) ASIC. (06) [CO4]
   d) The NRE cost and unit cost for the three IC technologies for a particular product are shown as follows: (13) [CO4]
      i) FPGA ($2,000, $100).
      ii) ASIC ($30,000, $30).
      iii) VLSI ($100,000, $2).
      Determine precise volume for which each technology yields the lowest total cost.

3. a) Suppose, you want to design and develop an automated chocolate vending machine. (12) [CO2]
      i) Draw the diagrammatic representation of this system.
      ii) Show the software requirement for this system.
      iii) Briefly explain the hardware components that are required in this system.
   b) What are the differences between sensors and transducers? Why are actuators required in embedded systems? (09) [CO3]
   c) How do 'Internet of Underground Things' work? Explain with a neat sketch. (07) [CO3]
   d) Illustrate Internet of Things life cycle in brief. (07) [CO3]

4. a) What is Hardware-Software co-design in embedded systems? How do we perform this task? Explain with a neat sketch. (09) [CO3]
   b) Suppose, you want to implement an embedded system. What are the criteria to choose a board for this task? (07) [CO2]
   c) A system consists of three modules: M_1, M_2, and M_3. It was analyzed, and the following reliability expression was derived: (07) [CO3]
      $$R_{system} = R_1R_3 + R_2R_3 - R_1R_2R_3$$
      Draw the reliability block diagram for this system.
   d) Show the hardware design and algorithms for obstacle avoidance robotic car using Arduino Uno and ultrasonic sensor. (12) [CO2]

SECTION B

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

5. a) What are the key factors you should consider in designing hard realtime and soft realtime systems? Explain with application examples. (10) [CO1]
   b) What is the main difference between co-operative scheduling and preemptive scheduling? Show the feasibility of scheduling for the following tasks by using Earliest Deadline First (EDF) algorithm. (13) [CO1]

Task	Release Time	Compute Time (C_i)	Deadline	Period (T_i)
T1	0	1	4	4
T2	0	2	6	6
T3	0	3	8	8

   c) Discuss RTOS kernel architecture and the modes of operation. Illustrate with neat sketches. (12) [CO1]

6. a) What are the security issues of operating systems in embedded systems? (08) [CO4]
   b) For a microcontroller based system, explain the role of interrupt latency and how it affects real-time behavior. (07) [CO2]
   c) Explain why interrupt-driven I/O is generally preferred over polling (with example). (08) [CO3]
   d) Provide the hardware and software design of a smart card system with necessary diagram(s) and table(s). (12) [CO4]

7. a) What are the different testing methods using software engineering approaches? (10) [CO1]
   b) What is shared data problem and how can you solve it? (10) [CO1]
   c) Design an IoT-based Room Temperature Monitoring and Control System using an ESP32 and a DHT22 sensor. The fan should start when temperature goes above 30°C, speed gradually increasing from 50% to 100% as temperature rises from 30°C to 50°C or above, using PWM control. Use Blynk IoT platform for remote temperature and fan monitoring. Provide necessary connection diagram, coding, and IoT server-node setup steps. (15) [CO3]

8. a) Explain Digital Twin for an Operational Technology (OT) system. How does MODBUS protocol work in industrial automation and control system along with SCADA and HMI (include figures). (13) [CO3]
   b) What is Zigbee and how can it be used for large-scale sensor deployment for effective remote monitoring with mesh network topology? (10) [CO3]
   c) How can you design a custom acceleration device using hardware descriptor language? (12) [CO2]

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY
B.Sc. Engineering 3rd Year 1st Term Examination, 2024
Department of Computer Science and Engineering
CSE 3107
Applied Statistics and Queuing Theory

TIME: 3 hours FULL MARKS: 210

N.B. i) Answer **ANY THREE** questions from each section in separate scripts.
ii) Figures in the immediate right column of the questions indicate full marks.
iii) The rightmost column indicates the course outcomes.
iv) *t*-table, χ^2-table, *F*-table, *z*-table will be provided if required.
v) A "*Formula Chart*" is provided on the last page of the question paper.

SECTION A

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

1. a) Explain how probability acts as a bridge between descriptive and inferential statistics with an example. (10) [CO3]

b) A software engineering course instructor has the project scores of students shown in the following table based on when they submitted their assignments relative to the deadline. (18) [CO4]

Submission Time	Scores (out of 100)
Early submission	95, 93, 90, 92, 96, 91, 94, 97, 88, 89
On-time submission	86, 84, 82, 88, 90, 79, 85, 83, 87, 81
Late submission	70, 68, 65, 63, 75, 72, 60, 67, 71, 74

i) Create a boxplot for each submission group. Show median, quartiles, IQR, and outliers.
ii) Compare the boxplots by skewness, spread, and median. Discuss how submission timing impacts project performance.

c) Suppose that the four inspectors at a film factory are supposed to stamp the expiration date on each package of film at the end of the assembly line. John, who stamps 20% of the packages, fails to stamp the expiration date once in every 200 packages; Tom, who stamps 60% of the packages, fails to stamp the expiration date once in every 100 packages; Jeff, who stamps 15% of the packages, fails to stamp the expiration date once in every 90 packages; and Pat, who stamps 5% of the packages, fails to stamp the expiration date once in every 200 packages. If a customer complains that her package of film does not show the expiration date, what is the probability that it was inspected by John? (07) [CO2]

2. a) In a networking performance test, the latency (in milliseconds) and data transfer rate (in Mbps) were measured for 10 test runs of a cloud server. The dataset is provided in the following table. (11) [CO3]

X (Latency, in milliseconds)	15	25	20	30	12	18	28	22	35	18
Y (Data Transfer Rate, in Mbps)	90	75	85	70	95	88	75	80	65	80

i) Rank the values and compute the Spearman's rank correlation coefficient for the dataset.
ii) Interpret the correlation result in terms of how latency and data transfer rate are related.

b) The following measurements were recorded for the drying time, in hours, of a certain brand of latex paint: (12) [CO2]

3.4, 2.8, 4.4, 2.5, 3.3, 4.0, 4.8, 5.6, 5.2, 2.9, 3.7, 3.0, 3.6, 2.8, 4.8

Assuming that the measurements represent a random sample from a normal population, find a 95% confidence interval for the drying time of this paint. Explain how confidence level affects the interval width and precision.

c) When computing sample variance, why do we divide $\sum_{i=1}^{n}(x_i - \bar{x})^2$ by $(n-1)$ instead of n (sample size) to get an unbiased estimation of the population variance? The lifetime of batteries is normally distributed with a mean of 30 months and a standard deviation of 2 months. A random sample of 1600 batteries are taken. Find the probability that the sample mean lifetime lies within two standard deviation of the population mean. (12) [CO1]

3. a) The frequency distribution of battery lives is approximately determined by a normal distribution with mean, $\mu = 3.5$ and standard deviation, $\sigma = 0.7$. The expected frequencies are obtained by computing the areas under the hypothesized normal curve that fall between the various class boundaries. Decide whether or not the normal distribution provides a good fit for the distribution of battery lives using χ^2-test. Here, $\alpha = 0.05$. (20) [CO

Class Boundaries	Observed Frequency
1.45–1.95	2
1.95–2.45	1
2.45–2.95	4
2.95–3.45	15
3.45–3.95	10
3.95–4.45	5
4.45–4.95	3

b) Calculate the regression coefficient and obtain the lines of regression for the following data. (10) [CO2]

X	10	12	13	12	16	15
Y	40	38	43	45	37	43

c) Illustrate Regression to the Mean with proper diagram. (05) [CO1]

4. a) In a medical experiment, the effectiveness of different medicines in reducing blood pressure was analyzed. For this purpose, three groups of patients, each group consisting of 6 patients were randomly selected; where each group was administrated a different drug medicine such as – medicine A, medicine B, medicine C for a period of four weeks. After the treatment period, the reduction is systolic blood pressure (in mmHg) was measured for each patient. The data obtained are shown the following table. (25) [CO4]

Medicine – A (mmHg)	6, 9, 7, 8, 10, 6
Medicine – B (mmHg)	10, 12, 11, 13, 14, 8
Medicine – C (mmHg)	4, 5, 6, 5, 7, 9

Now using 5% level if significance perform the followings:

i) Apply ANOVA technique for testing the significance of the difference between group means.

ii) If (i) shows significant difference then perform t-tests to find out pairs of groups that are significantly different from one another.

b) Prove that $Var(X) = E(X^2) - (E(X))^2$ where, X is a random variable, Var is variance and E is expected value. (10) [CO3]

SECTION B

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

5. a) Derive the Chapman-Kolmogorov equation for discrete time Markov chain. (11) [CO3]

b) Parkinson's disease progression in patients can be modeled as a continuous time Markov chain with 3 states, $X = 0$ (Early stage), $X = 1$ (Advanced stage), and $X = 2$ (Severe stage). Transition from $X = 0$ to $X = 1$ occurs with a rate μ_1, from $X = 1$ to $X = 2$ occurs with a rate μ_2, from $X = 2$ to $X = 1$ occurs with a rate μ_3. Show that, steady state value for state, $X = 1$ is equivalent to $\frac{\mu_3}{(\mu_2 + \mu_3)}$. (11) [CO2]

c) In a small community clinic, there are two patients who are vulnerable to a chronic viral infection. Each healthy patient becomes infected after an average of 22 days (exponentially distributed). Once a patient becomes infected, they receive treatment and recover after an average of 10 days (exponentially distributed). Answer the followings: (13) [CO1]

i) Formulate a birth-death model for this situation.

ii) Determine the fraction of the time when no patients are affected.

iii) Determine the fraction of the time when both patients are affected.

6. a) Briefly explain the significance of the "memory property" in modeling the arrival and service process in a queue system. (09) [CO1]

b) From the time a request for data is submitted until the request is fulfilled, a database takes an average of 3 seconds to respond to a request for data. The database is found to be idle around 20% of the time. Answer the following questions assuming that the database can be modeled as an M/M/1 queue system. (14) [CO4]

i) What is the average service time per database query?

ii) What is the average number of queries in the system?

iii) What is the probability that 5 or more queries are present?

c) Determine which of the following chains is "ergodic". (12) [CO1]

$$P_1 = \begin{bmatrix} 0.4 & 0 & 0.6 \\ 0.3 & 0.3 & 0.4 \\ 0 & 0.5 & 0.5 \end{bmatrix} \quad P_2 = \begin{bmatrix} 0.7 & 0 & 0 & 0.3 \\ 0.2 & 0.2 & 0.4 & 0.2 \\ 0.6 & 0.1 & 0.1 & 0.2 \\ 0.2 & 0 & 0 & 0.8 \end{bmatrix}$$

7. a) A radiology department processes medical images arriving randomly for diagnosis. The arrivals follow a Poisson process with an average of 30 image scans per 8-hour shift. The service time per scan (i.e., the time to process and analyze an image) is uniformly distributed between $a = 10$ and $b = 14$ minutes. If the mean and variance for any uniform distribution can be computed using $\frac{b+a}{2}$ and $\frac{(b-a)^2}{12}$ respectively, find the followings: (14) [CO4]
   i) The expected fraction of the time, the processing station is idle.
   ii) The expected number of scans waiting at any random time during the shift.
   iii) The expected time a scan spends in the system.

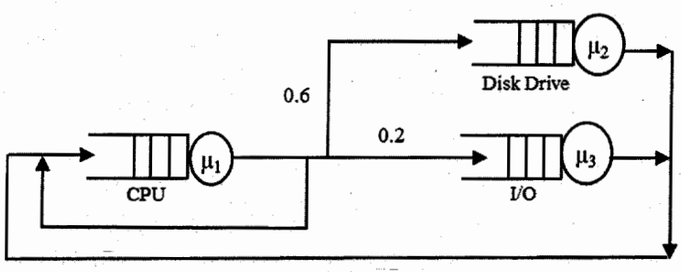
   b) Explain why the M/D/1 queue typically has lower average waiting time than an M/M/1 queue with the same arrival rate and mean service time. (08) [CO1]

   c) A machine learning inference server processes prediction requests. The server can handle an average of 20 predictions per hour (processing times are exponentially distributed). On average, 30 prediction requests per hour arrive at the server (interarrival times are exponentially distributed). Due to memory limitations, the maximum capacity of the server is 30 predictions (including the one being processed). Answer the followings: (13) [CO4]
   i) On average, how many prediction requests successfully enter the system each hour?
   ii) What is the probability that the inference server is busy processing a request?

8. a) Derive the equation to compute the mean response time of an M/M/s/GD/∞/∞ queuing system for various traffic intensities. The symbols have their usual meanings. (12) [CO3]

   b) Consider a circular closed queuing network of two queues where the departure of Queue 1 goes into Queue 2 and that of Queue 2 goes into Queue 1. If there are 4 customers circulating in the network and the service times at both queues are exponentially distributed with rates μ_1 and μ_2 respectively, draw the state transition diagram. Can you simplify the network by removing the feedback? (10) [CO3]

   c) Consider the simple model of a computer system shown below, where Queue 1 – the CPU, Queue 2 – disk drive, and Queue 3 – I/O. Given $\mu_1 = 10$, $\mu_2 = 5$, $\mu_3 = 1$, $K = 4$ jobs. Find the normalization constant and actual throughputs of each queue. (13) [CO4]



Formula Chart: Applied Statistics and Queuing Theory

- Queuing Models

Parameters	M/M/1	M/M/s	M/M/1/GD/n/∞	M/G/1/GD/∞/∞
π_0	$1-\rho$, $\rho = \lambda/\mu$	$\left(\sum_{j=0}^{s-1}\frac{(s\rho)^j}{j!}+\frac{(s\rho)^s}{s!(1-\rho)}\right)^{-1}$ $\rho = \lambda/(s\mu)$	• $\frac{1-\rho}{1-\rho^{n+1}}$ $for\ \rho \neq 1$ • $\frac{1}{n+1}$ $for\ \rho = 1$	$1-\rho$
π_j	$\rho^j(1-\rho)$	• $\frac{(s\rho)^j\pi_0}{j!}$ for$(j = 1,2,\dots,s)$ • $\frac{(s\rho)^j\pi_0}{s!s^{j-s}}$ for $(j = s, s+1, s+2, \dots)$	$\rho^j\pi_0$ $for\ j = 1,2,3,\dots\dots,n$ and $for\ any\ \rho$	$\rho^j(1-\rho)$
L	$\frac{\rho}{1-\rho}$	$\frac{\rho}{1-\rho}P_d + s\rho$	• $\frac{\rho[1-(n+1)\rho^n+n\rho^{n+1}]}{(1-\rho^{n+1})(1-\rho)}$ $for\ for\ \rho \neq 1$ • $\frac{n}{2}$ $for\ \rho = 1$	$L_q + \rho$
L_q	$\frac{\rho^2}{1-\rho}$	$\frac{\rho}{1-\rho}P_d$	$L-(1-\pi_0)$	$\frac{\lambda^2\sigma^2+\rho^2}{2(1-\rho)}$
Extra	-	delay probability: $P_d = P(j \geq s)$ $= \frac{(s\rho)^s\pi_0}{s!(1-\rho)}$	Saturation Probability, • $\pi_n = \frac{1-\rho}{1-\rho^{n+1}}\rho^n$ $for\ \rho \neq 1$ • $\pi_n = \frac{1}{n+1}$ $for\ \rho = 1$ • $\lambda_{lost} = \lambda\pi_n$ • $\lambda_{eff} = \lambda - \lambda_{lost} = \lambda(1-\pi_n)$	Erlang Distribution: Mean: k / λ. Variance: k / λ^2. Where k: shape parameter and λ: rate parameter

- Waiting Time Distribution

M/M/1/FCFS/∞/ ∞	M/M/s/FCFS/∞/ ∞
• $w(t) = (\mu-\lambda)e^{-(\mu-\lambda)t}, t > 0$ • $P(t > W) = 1 - \int_0^W w(t)dt$	• $P(W > t) = e^{-\mu t}\left\{1 + P(j \geq s)\frac{1-\exp[-\mu t(s-1-s\rho)]}{s-1-s\rho}\right\}$
• $w_q(t) = \rho(\mu-\lambda)e^{-(\mu-\lambda)t}, t > 0$	• $P(W_q > t) = P(j \geq s)\exp[-s\mu(1-\rho)t]$, $P(j \geq s) = \frac{(s\rho)^s\pi_0}{s!(1-\rho)}$

- Queuing Network

Closed Queuing Network Parameters	• $\pi(\mathbf{n}) = \frac{1}{G(K,m)}\prod_{i=1}^{m}\rho_i^{n_i}$ where $\rho_i = \frac{\lambda_i}{\mu_i}$ • $G(K,m) = G(K,m-1) + \rho_m G(K-1,m)$ $G(0,m) = 1, \quad m = 1,2,\dots\dots,M$ $G(k,1) = \rho_1^k, \quad k = 0,1,2,\dots\dots,K$ • $L_i = \frac{1}{G(K,M)}\sum_{k=1}^{K}\rho_i^k G(K-k,M)$ • $e_i = \lambda_i\frac{G(K-1,M)}{G(K,M)}$, $W_i = \frac{L_i}{e_i}$ • $P(n_i \geq k) = \rho_i^k\frac{G(K-k,M)}{G(K,M)}$ • $\rho_{e_i} = \frac{e_i}{\mu_i}$

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY
B.Sc. Engineering 3rd Year 1st Term Examination, 2024
Department of Computer Science and Engineering
CSE 3109
Database Systems

TIME: 3 hours FULL MARKS: 210

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

SECTION A

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

1. a) Define data model with example. How can you differentiate logical and physical design of a database? Explain. (08) [CO1]
   b) In an educational institute, a teacher teaches one or more courses. Each of the course is taught by one or more teachers. A teacher is associated with ID and name. A course is associated with code and content. Sketch an E-R diagram of the system. Hence develop a physical schema. (10) [CO1]
   c) Consider the following INSTRUCTOR entity. Convert it to an appropriate physical schema. How can you handle a multivalued attribute? (10) [CO1]

```
INSTRUCTOR
ID
Name
        first-name
        middle-name
        last-name
Address
        Street
                no.
                name
                apt-number
        City
        State
        Zip
{Phone Number}
Date-of-Birth
age()
```

   d) How can you represent a weak entity set in a database? Explain with an E-R diagram. (07) [CO1]

2. a) Make a classification of storage devices. When do you choose Storage Area Network (SAN) and Network Attach Storage (NAS), and why? (10) [CO2]
   b) Give a comparison among (i) primary index and secondary index, (ii) dense index and sparse index. (10) [CO2]
   c) For a B+ tree, the key value is 18 bytes and pointer (both page and record pointer) is 14 bytes. If the page size is 4KB then how many key values you can enter in the leaf and non-leaf node of the B+ tree? (07) [CO2]
   d) How does Bitmap indexing perform indexing? When do you like Bit indexing to apply in a database? (08) [CO2]

3. a) What are the goals of a good database design? How does database normalization help to achieve the goals? Explain. (09) [CO1]
   b) Consider two tuples t_1 and t_2 of a relation r. If $t_1 \neq t_2$ and $t_1.k \neq t_2.k$ then what decision(s) can you take about k, where k is an attribute of r? Explain with example. (08) [CO2]
   c) A relation $R(A, B, C)$ is decomposed into two relations $R_1(A, B)$ and $R_2(B, C)$. How can you determine whether the decomposition is lossless or not? (10) [CO1]
   d) In a database schema $R(buyer, product, company)$ there are 20 *buyers*, 50 *products* and 100 *companies* exist. Redesign R so that total number of records are minimal. (08) [CO2]

4. a) What is closure of attributes? What are the applications of it? Given $R = (A, B, C, G, H, I)$ with $F = \{A \rightarrow B, A \rightarrow C, CG \rightarrow H, CG \rightarrow I, B \rightarrow H\}$, Do you think (AG) and (BG) are the primary keys? (10) [CO1]

b) Make a comparison between structured, semi-structured and unstructured data with example. (08) [C
c) What are the advantages of NoSQL systems over SQL? (07) [CO3]
d) Explain 'wide column representation' and 'sparse column representation' for semi-structured data. How are nested datatypes handled in JSON data format? Explain with example. (10) [CO3]

SECTION B

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

5. a) Consider the following relational database. (20) [CO2]

```
doctor(did, dname, specialty)
patient(pid, pname, disease)
appointment(pid, did, date)
ward(wid, wname, capacity)
admission(pid, wid, admit-date)
```

Draw the schema diagram and write SQL statements for each of the following queries:

i) Find the names of patients who have appointments with cardiologists.
ii) Find the names of doctors who treated more than 5 patients.
iii) Find the names of patients admitted in wards with `capacity > 50`.

b) "View definition is not the same as creating a new relation by evaluating the query expression" – justify the statement. (09) [CO3]
c) Briefly discuss the ACID properties required for maintaining data integrity. (06) [CO3]

6. a) Consider the following schema: (15) [CO2]

```
account(acc-id, acc-name, br-name, balance)
branch(br-name, br-city, asset)
```

Write relational algebra expressions for each of the following queries:

i) Find the account id and name of the accounts whose balance is greater than the balance of the account with `acc-id = 'A-1212'`.
ii) Find the branch city of the account with `acc-id = 'A-101'` and rename the output relation as `temp` with the attribute name as `branch-city`.
iii) Find the `account` who has the minimum balance.

b) Consider the following schema: (09) [CO3]

```
students(roll, name, cgpa)
```

Create a cursor to find the `roll` and `name` of the `students` who have `cgpa` more than 3.5 and display them inside a PL/SQL block.
c) Explain SQL injection with an example. (06) [CO2]
d) Discuss how precedence graph is used for determining conflict serializability. (05) [CO3]

7. a) You are given two tables: Result and Distinction. The Result table contains the grades of all students in all subjects. The Distinction table contains the conditions for a student to pass with distinction, i.e., a student must get the same grades specified in the distinction table for all subjects to be a distinction holder. (16) [CO2]

Table1: Result

ID	Name	Subject	Grades
1	Abe	Japanese	A
2	Imai	Japanese	C
5	Ono	Japanese	B
5	Ono	English	A
1	Abe	English	A
2	Imai	English	B

Table2: Distinction

Subject	Grade
Japanese	A
English	A

i) Identify the relational operator with which you can find out the IDs and names of the students who passed with distinction, and show the output table.
ii) Formulate an SQL query to do the operation of (i).

b) Explain Thoma's write rule in concurrency control. (06) [CO4]
c) You are given the following schema: (13) [CO4]

```
students(s-id, s-name, total-credits)
courses(c-id, title, credits)
takes(s-id, c-id)
```

Write a PL/SQL trigger that automatically updates the `total-credits` of a student in the `students` table whenever a new record is inserted into the `takes` table. The trigger should add the `credits` value of the taken course to the student's `total-credits`. If the `credits` value of the taken course more than 3, the insertion shouldn't take place.

8. a) "Every cascadeless schedule is recoverable" – Explain the statement. (08) [CO4]
   b) Discuss the key ideas behind Multiversion schemas in concurrency control. Explain how the Multiversion Timestamp Ordering Protocol (MTSO) handles `read()` and `write()` operations. (12) [CO4]
   c) Explain how write conflicts are handled in Snapshot Isolation. (06) [CO4]
   d) Consider the following concurrent schedule: (09) [CO4]

T_{25}	T_{26}
read(B)	
	read(B)
	B := B-50
	write(B)
read(A)	
	read(A)
display(A+B)	
	A := A+50
	write(A)
	display(A+B)

Determine if the schedule is valid under Timestamp Ordering Protocol (TSO).

KHULNA UNIVERSITY OF ENGINEERING & TECHNOLOGY
B.Sc. Engineering 3rd Year 1st Term Examination, 2024
Department of Computer Science and Engineering
CSE 3119
Information Systems Design

TIME: 3 hours FULL MARKS: 210

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

SECTION A

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

1. a) What is an information system? How does it support business operation? (10) [CO1]
   b) What are the elements of a system? Can you have a viable system without feedback? Explain. (13) [CO2]
   c) When does an analyst terminate a project? How does it tie with post implementation? Explain. (12) [CO3]

2. a) Define the system development life cycle (SDLC). What are its main phases? (08) [CO1]
   b) Distinguish between initial investigation and feasibility study. In what way are they related? (13) [CO2]
   c) Describe the role of a system analyst. What skills are essential for this role? (14) [CO3]

3. a) What do you mean by input design in information systems? Why is it important? (12) [CO1]
   b) What is form design? What principles should be followed when designing input form. (13) [CO2]
   c) Discuss the use of validation checks in input design. Explain it with an example. (10) [CO3]

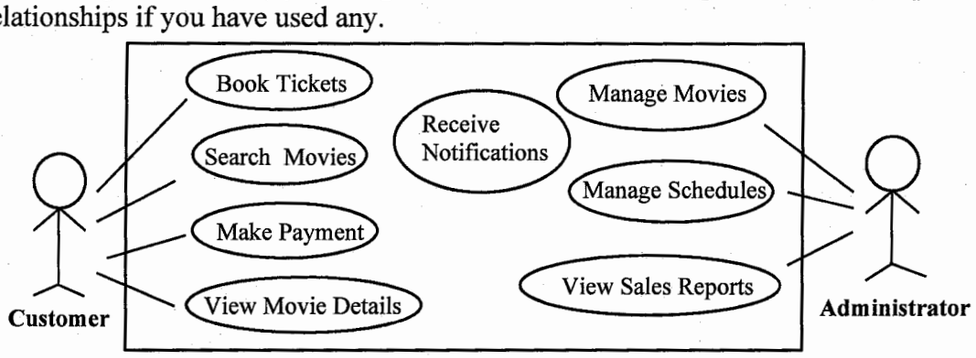
4. a) What is the difference between analysis and design? Can one begin to design without analysis? Why? (10) [CO1]
   b) Why is a database important in MIS? Explain. (10) [CO2]
   c) Write a short essay on the concept and uses of DSS. Include a brief discussion on the relationship between DSS and decision making. (15) [CO2]

SECTION B

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

5. a) "X" hospital is developing a new online appointment system. The IT department has fixed set of requirements. Hospital administrators demand a quick interface mock-up for early visualization before final development. Now, answer the followings: (12) [CO3]
      i) Identify the most suitable SDLC model for the IT team and administrators.
      ii) As a developer, which model would you prefer and why?
   b) Compare six sigma and Total Quality Management (TQM) in the context of system development. (10) [CO2]
   c) An e-commerce platform sells a wide range of products and manages user interactions. Design a class hierarchy for the e-commerce platform. Include a base class `'Product'` and at least three derived class (`'Electronics'`, `'Clothing'`, `'Book'`). Explain how inheritance is used in your design to promote code reuse and extensibility. (13) [CO1]

6. a) Consider the following simplified use case diagram for the online movie Ticket Booking System which is also incomplete. Draw the complete diagram including actions for (i) Booking tickets and (ii) sending notifications after booking tickets. Also, explain the relationships if you have used any. (13) [CO1]



b) When using unobtrusive methods to analyze historical order forms at "X" company, an analyst needs to determine an appropriate sample size. An initial check suggests 5% of forms have errors. If the analyst requires the sample estimate to be within ±2% of the true proportion with 95% confidence, calculate the required sample size. (10) [CO2]

c) "X" university is launching a new Online Course Registration System consisting of modules such as `UserAuth`, `CourseCatalog`, `RegistrationManager`, and `PaymentHandler`. Before deployment, the team must ensure that all modules and their interactions work correctly. Explain how the testing process would be applied to this system. (12) [CO3]

7. a) Draw a sequence diagram for an Online Food Delivery System that handles customer orders, restaurant processing, delivery logistics, and payment transactions. (10) [CO1]

b) The following table lists the activities, durations (in Days), and dependencies for a project management scenario. (10) [CO2]

Activity	Duration (Days)	Dependencies
A	5	–
B	7	A
C	4	A
D	2	B
E	3	C
F	6	D, E
G	5	F

Now, answer the followings:

i) Draw the PERT diagram for this project.

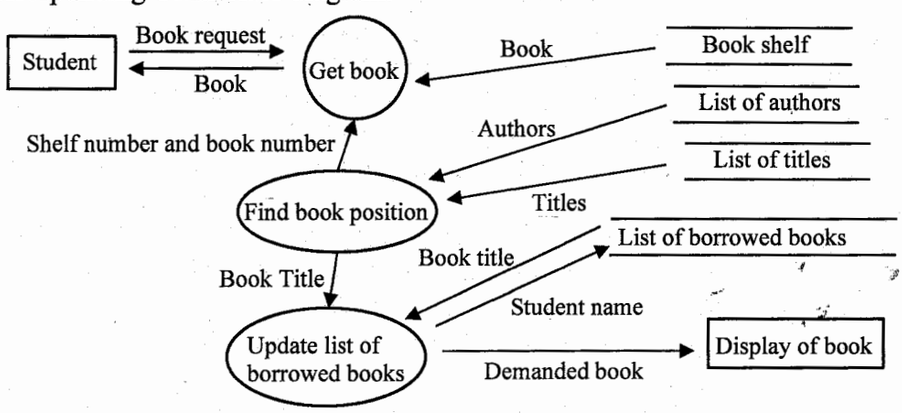
ii) Identify the critical path and calculate the total project duration.

c) A company installs monitoring software that tracks every employee's screen and keystrokes in real-time, without notifying them. Management claims it is to "improve productivity". Identify the ethical conflict in this situation. How can the organization maintain a balance between security and employee privacy? (08) [CO3]

d) Describe the three fundamentals structures for arranging questions in interactive methods for information gathering. (07) [CO4]

8. a) You are designing a data-entry procedure for a new user registration form. The business rules state that a password must be at least 8 characters long and contains at least one uppercase letter, one lowercase letter, and one digit. Create a regular expression that could be used to validate that a password meets these requirements. (10) [CO1]

b) Based on the principles of Understanding Organizational Style, a Level-1 DFD for a Library Management System is provided in the following. Using this DFD, create a corresponding UML class diagram. (17) [CO2]



c) A company stores all data on one server. One day, a hardware crash leads to total data loss. Identify which security weakness caused this issue. Propose two disaster recovery measures to prevent this in the future. (08) [CO3]