INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

Stamp / Signature of the Invigilator

EXAMINATION (End Semester)													SEMESTER (Spring 2023)	
Roll Number										Section		Name		
Subject Number	C	S	6	0	0	0	2	Subject Name					*Distributed Systems*	
Department / Center of the Student													Additional sheets	

Important Instructions and Guidelines for Students

1. You must occupy your seat as per the Examination Schedule/Sitting Plan.
2. Do not keep mobile phones or any similar electronic gadgets with you even in the switched off mode.
3. Loose papers, class notes, books or any such materials must not be in your possession, even if they are irrelevant to the subject you are taking examination.
4. Data book, codes, graph papers, relevant standard tables/charts or any other materials are allowed only when instructed by the paper-setter.
5. Use of instrument box, pencil box and non-programmable calculator is allowed during the examination. However, exchange of these items or any other papers (including question papers) is not permitted.
6. Write on both sides of the answer script and do not tear off any page. **Use last page(s) of the answer script for rough work.** Report to the invigilator if the answer script has torn or distorted page(s).
7. It is your responsibility to ensure that you have signed the Attendance Sheet. Keep your Admit Card/Identity Card on the desk for checking by the invigilator.
8. You may leave the examination hall for wash room or for drinking water for a very short period. Record your absence from the Examination Hall in the register provided. Smoking and the consumption of any kind of beverages are strictly prohibited inside the Examination Hall.
9. Do not leave the Examination Hall without submitting your answer script to the invigilator. **In any case, you are not allowed to take away the answer script with you.** After the completion of the examination, do not leave the seat until the invigilators collect all the answer scripts.
10. During the examination, either inside or outside the Examination Hall, gathering information from any kind of sources or exchanging information with others or any such attempt will be treated as **'unfair means'**. Do not adopt unfair means and do not indulge in unseemly behavior.

Violation of any of the above instructions may lead to severe punishment.

Signature of the Student

To be filled in by the examiner

Question Number	1	2	3	4	5	6	7	8	9	10	Total
Marks Obtained											

Marks obtained (in words)	Signature of the Examiner	Signature of the Scrutineer

Write the answers in the boxes only. You can use the designated spaces for rough works. This question has 14 pages including the space for rough works.

1. Answer the following questions briefly. Marks will be deducted for unnecessary descriptions. No marks will be given if the answers are not explained and only Yes/No answer is given.

[15x3=45 Marks]

(a) Assume that there are three processes. Any consecutive pair of messages received on any of these three processes follows the causal delivery principle. Does this ensure that the message delivery between any pair of processes across the above three processes is FIFO? If yes, deduct a formal proof; if no, give a counter-example.

(b) Does Lamport timestamp ensure that all the message delivery are causal across different processes?

(c) Define a stable message. Does FIFO delivery across processes ensure that messages are stable?

(d) Consider two events e_i from process p_i and e_j from process p_j, such that $i \neq j$. When do we say that e_i and e_j are pairwise inconsistent? Can we use vector clocks to detect pairwise inconsistent events? Give an example.

(e) Define the interactive consistency problem. Can we use an agreement protocol to solve the interactive consistency problem? Explain your answer.

(f) Say, there can be a maximum of f number faults in a system. What are the minimum numbers of nodes required to achieve safety in a consensus algorithm for the following two cases – (i) Synchronous CFT, and (ii) Asynchronous CFT? Explain your answers.

```
Synchronous CFT:



Asynchronous CFT:
```

(g) Consider a Paxos proposer P1 who has received an Accept message with ID 40 and value *Apple*. In the same Paxos round, what will happen when a second Paxos proposer P2 sends a Prepare message with (i) ID 20, and (ii) ID 50?

```
(i) P2 sends Propose with ID 20:



(ii) P2 sends Propose with ID 50:
```

(h) Does RAFT ensure liveness for an asynchronous CFT? Explain your answer.

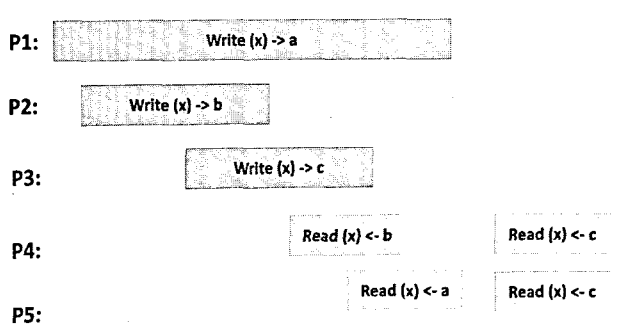
(i) Why does a PBFT replica need to send the information about the last stable checkpoint along with a set of $2f+1$ valid checkpoint messages (for tolerating a maximum of f number of faults) with the VIEW-CHANGE message?

(j) PBFT clients need to receive only $f+1$ replies from the replicas to decide the consensus output, then why do PBFT replicas need $2f+1$ checkpoint messages to ensure a consensus with a maximum of f number of faulty nodes?

(k) Why is it not possible to design a deterministic leader election algorithm over an anonymous ring? Explain your answer.

(l) Consider a tree-structured quorum protocol with processes P_1 to P_7. Write down the set of all feasible quorums. Assume that process P_3 fails. What are the feasible quorums at this case.

(m) Consider the following set of operations across five different processes, where the x-axis indicate the timeline. Are these events (i) linearizable?, (ii) sequential consistent?



Linearizabe? (Yes/No)

Sequential Consistent? (Yes/No)

(n) What is the trade-off between active replication and passive replication? Explain a scenario when passive replication is a better design choice over active replication.

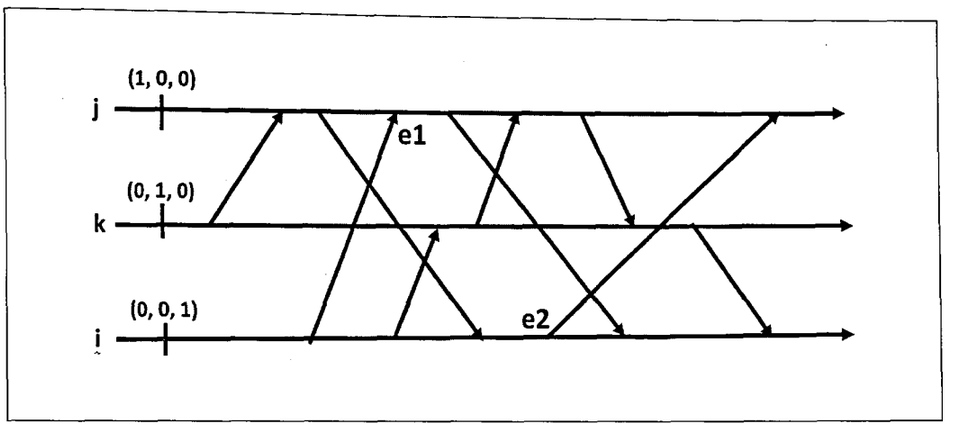
(o) Assume that you want to design a web button with a counter, called *like*, as follows. If the button is clicked, it is considered as a "like", and the counter value is incremented by one. If the button is long-pressed, it is considered as a "superlike", and the counter value gets doubled up. Can you design a CmRDT to implement this counter with multiple replicas supporting eventual consistency? Explain your answer.

2. **Distributed Clocks**

[Total 15 Marks]

(a) In the following diagram, there are three processes *i*, *j* and *k*. The arrows indicate the message passing across different processes. Put the vector clocks for all the events in the diagram.

[5 Marks]



(b) When do we say two events are parallel, as per the vector clock notation (define formally)? In the above diagram, can we say that event e1 and e2 are parallel? Explain your answer.

[2+2 Marks]

(c) Define a true clock. With an example, show and explain why Lamport clock is not a true clock, but vector clock is a true clock. Your example should consider a set of processes and communications among them through message passing. Then put up Lamport clock values and Vector clock values for all the events, and explain. [No marks will be given without the example.]

[2+4 Marks]

3. **Consensus** **[Total 10 Marks]**

(a) Define a Byzantine dissemination quorum. **[2 Marks]**

(b) Why does PBFT need a weak synchrony assumption for the view change protocol? Explain which step of the view change protocol will not work under a pure asynchronous environment. **[2 Marks]**

(c) Consider a RAFT instance with 5 replicas R1 to R5. Say, the leader R1 is elected at Term 4, and the latest (index, term) for the five replicas are as follows. `R1:(10,4)`, `R2:(8,4)`, `R3:(5,4)`, `R4:(9,4)`, `R5:(4,3)`. Assume that all the logs are consistent. (a) What is the first index that R1 has served as a leader? (b) What is the last index upto which the operations can be considered to be committed? **[1+1 Marks]**

(d) Consider a RAFT log in the form `(index,term)||Operation`. Consider the following logs for a RAFT leader (that has been elected at Term 4) and a follower.

Leader	Follower
`(1,1)|| x := 3`	`(1,1)|| x := 3`
`(2,1)|| y := 5`	`(2,1)|| x := 5`
`(3,1)|| x := 8`	`(3,1)|| x := 8`
`(4,2)|| y := 3`	`(4,1)|| y := 5`
`(5,3)|| z := 5`	

Now, say the leader broadcast a log message as `(6,4)|| x := 15`. Explain the steps that will be followed in RAFT to ensure the log consistency between the leader and the above follower.

[4 Marks]

4. **Distributed Algorithms** **[Total 10 Marks]**

(a) Let there be four processes – P1, P2, P3, and P4. Following Maekawa's algorithm for distributed mutual exclusion, (i) write down a set of feasible quorums for all the four processes. (ii) Say, P1 wants to execute a critical section C. However, assume that P4 is currently executing C. Write down the steps through which P4 will release the critical section, and P1 will start executing the critical section. **[2+4 Marks]**

(b) Consider the distributed mutual exclusion algorithm based on Lamport's clock. Does this algorithm support fairness? **[2 Marks]**

(c) Consider Chang and Robert's algorithm for distributed leader election over a unidirectional ring. What is the worst case message complexity for this algorithm? Explain your answer. **[2 Marks]**

5. **Consistency in Distributed Applications** **[Total 10 Marks]**

(a) Consider a data structure called a 2P-set, where you can add or remove elements dynamically. Let $\mathbb{P}$ be a 2P-set, you can add an element e, $\mathbb{P} = \mathbb{P} \cup \{e\}$, or remove an element e', $\mathbb{P} = \mathbb{P} \setminus \{e'\}$. Design a CvRDT to implement the 2P set over multiple replicas. Show the execution of the add and remove operations on this set through a diagram, with three replicas. **[10 Marks]**