UNIVERSITY OF LONDON

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 1996

MEng Honours Degrees in Computing Part IV
MSc Degree in Foundations of Advanced Information Technology
for Internal Students of the Imperial College of Science, Technology and Medicine

This paper is also taken for the relevant examinations for the
Diploma of Membership of Imperial College
Associateship of the City and Guilds of London Institute

PAPER 4.37

DISTRIBUTED ALGORITHMS

Wednesday, May 1st 1996, 10.00 - 12.00

Answer THREE questions

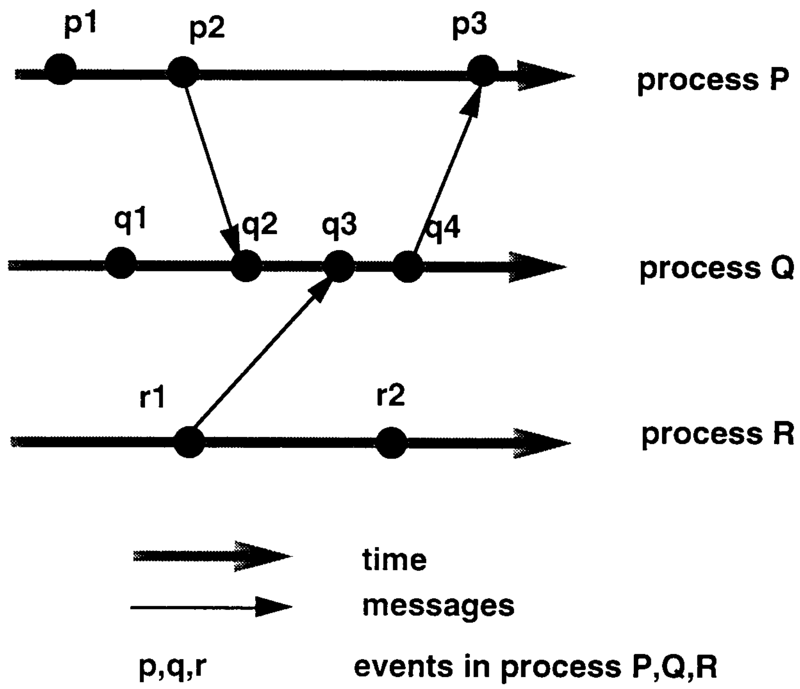
For admin. only: paper contains
4 questions
3 pages (excluding cover page)

1 a i) Given any two events a,b in a distributed system, define what is meant by the causal ordering relation a → b, i.e. "a precedes b".

ii) Describe how Lamport's logical clocks can be used to provide partial ordering consistent with causal ordering, i.e.
a → b implies that $C(a) < C(b)$ for logical timestamps C and events a, b.

iii) Starting from C=0, annotate the space/time diagram below with logical timestamp values for each event in each process.

iv) Explain why $C(a) < C(b)$ does not necessarily imply that a → b?



b i) Describe how vector clocks can be used to provide partial ordering which reflects causal ordering precisely, i.e.
a → b iff $VT(a) < VT(b)$ for vector timestamps VT and any events a, b.

ii) Starting from VT=(0,0,0), annotate the space/time diagram above with vector timestamp values for each process.

iii) It is asserted that, at any instant,
$\forall i, \forall j$: VT[i] in process i $\geq$ VT[i] in process j.
Is this true and why?

The two parts carry, respectively, 60%, 40% of the marks.

2 *Communication Deadlock Detection:*

Consider an environment in which there are a number of communicating processes Pi. Processes can be blocked waiting for message transmission to or message receipt from other processes. For some blocked process Pi we can define the *Dependence Set, DSi*, of Pi as the set of all processes from which Pi is expecting communication.

a Using dependence sets, define the conditions necessary for some set S of processes to be in a communication deadlock.

b Describe a diffusing algorithm for *communication deadlock detection* (Chandy Misra and Haas) which propagates probes to the dependence set of each process. Ensure that you describe the following:

i) the form of the probe and reply,

ii) the information kept by the processes to maintain the diffusion tree (*engager* and *num*), to discard old probes/replies (*latest*), and to record changes of state (*wait*),

iii) an outline of the algorithm used by each process to initiate, discard or propagate probes and handle replies, and

iv) the condition for detection of deadlock.

(You may assume no loss of messages and preservation of sequencing i.e.no overtaking).

c Comment on the symmetry of the algorithm in part b.
Is it possible for more than one process to detect deadlock? What do you think is the implication for deadlock resolution?

The three parts carry, respectively, 15%, 75%, 10% of the marks.

Turn over...

3a Explain briefly why it is impossible to design a correct atomic broadcast algorithm for asynchronous systems subject to crash failures of processes. Do not attempt to prove any general impossibility results you use, simply state them.

b Given a reliable broadcast algorithm which satisfies the Validity, Agreement and Integrity properties for broadcasts:

i) State the additional property required of a Timed Reliable Broadcast.

ii) State the ordering property required of an Atomic Broadcast.

iii) Develop a Timed Atomic Broadcast algorithm which satisfies the properties of Atomic Broadcast and is implemented using Timed Reliable Broadcast.

c Explain briefly how the ISIS system enforces the ordering property required of Atomic Broadcasts in its ABCAST protocol.

d Comment on why ABCAST does not contradict the impossibility mentioned in part a).

The four parts carry, respectively, 20%, 30%, 35% , 15% of the marks.

4a Define the *Byzantine Generals Problem* in terms of the *interactive consistency* conditions.

b Show informally that a solution to the Byzantine Generals Problem is not possible with only three generals one of whom may be a traitor.

c State the message passing assumptions which are usually made in arriving at a solution to the Byzantine Generals Problem. Briefly describe what each of these assumptions implies in the construction of reliable computing systems.

d Show that by restricting the types of failures which can occur, a solution can be achieved for three generals with one traitor. State explicitly any additional assumptions you make.

e The computer company Paranoid PLC decides to build a reliable multi-processor system which can tolerate Byzantine failures. How many messages are required to reliably send a message from one processor to the rest in a system with ten processors which is required to tolerate up to three simultaneous failed processors?

The five parts carry, respectively, 10%, 25%, 25%, 25%, 15% of the marks.

End of paper