UNIVERSITY OF LONDON

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 1996

BEng Honours Degree in Computing Part II
MEng Honours Degrees in Computing Part II
for Internal Students of the Imperial College of Science, Technology and Medicine

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

PAPER 2.4

ARCHITECTURE II

Wednesday, May 8th 1996, 2.00 - 3.30

Answer THREE questions

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

1a i) Provide a formula showing how the execution time of a program can be affected by factors such as the number of instructions for that program.

ii) Suggest three ways of improving the execution time for a given program.

b Two machines, M1 and M2, implement the same instruction set with two classes of instructions, A and B. M1 takes $A1$ cycles to run each Class A instruction and $B1$ cycles to run each Class B instruction, while M2 takes $A2$ cycles to run each Class A instruction and $B2$ cycles to run each Class B instruction. M1 runs at $C1$ Hz, while M2 runs at $C2$ Hz.

i) State the peak performance, in number of instructions per second, of M1 and M2.

ii) State the execution time of M1 and M2 for running programs with m Class A instructions and n Class B instructions.

iii) For M1, a program P has $u1$ Class A instructions and $v1$ Class B instructions, while another program Q has $x1$ Class A instructions and $y1$ Class B instructions. The corresponding parameters for M2 are respectively $u2$, $v2$, $x2$ and $y2$.
If P is run k times more frequently than Q, derive a value for k such that the total execution time for M1 will be less than that for M2.

iv) For the two programs in part (iii), what is the ratio of the clock speed for the two machines if P is run as frequently as Q, and the total execution time is the same for both machines?

The two parts carry, respectively, 30% and 70% of the marks.

2a Give three characteristics of Load-Store architectures.

b Suggest two advantages and two disadvantages of Load-Store architectures. Mention two factors concerning technology advance in the last decade which contributed to the popularity of Load-Store architectures.

c The Load Immediate Signed instruction of a 32-bit Load-Store architecture is divided into four fields: a 6-bit opcode field, two 5-bit register fields, and a 16-bit constant field.

i) Explain what is *sign extension*, and why it is necessary in implementing the Load Immediate Signed instruction.

ii) Provide a circuit diagram of an appropriate sign extension circuit for this Load-Store architecture. Label the inputs in0, in1, in2, ..., and the outputs out0, out1, out2,..., least significant bit first.

iii) Provide another circuit diagram of a sign extension circuit that can deal with both unsigned representations and two's complement representations, given that a control signal c indicates which representation is required in every cycle.

The three parts carry, respectively, 15%, 30% and 55% of the marks.

3a Provide formulae for computing the number of *read stall* cycles and *write stall* cycles for a given program executing on a given machine.

b Explain briefly why read stalls and write stalls occur.

c A machine takes on average c cycles per instruction without considering memory stalls, and has a miss penalty p for all misses. Let i and d be respectively the instruction cache miss rate and data cache miss rate for a program with n instructions and f memory references per instruction. Assume that the hit time is not a factor in determining cache performance.

i) How many cycles are required to run this program, when memory stall is taken into account?

ii) If we want to improve the execution time of this program by a factor of k, how much do we need to reduce the number of cycles that each instruction takes?

iii) Another way to improve the execution time is to reduce the cycle time of the machine. If we want to improve the execution time by a factor of k, how much do we need to improve the cycle time?

The three parts carry, respectively, 15%, 15% and 70% of the marks.

Turn over...

4a Cache misses can be classified into three categories: *compulsory misses*, *capacity misses* and *conflict misses*.

i) Explain briefly what they are.

ii) How are they affected by increasing cache size or by increasing the degree of associativity? Are there negative effects on performance?

b Show how to configure a twelve-block cache as

i) direct mapped,

ii) three-way set associative,

iii) fully associative.

Suggest a disadvantage of increasing the degree of associativity.

c Consider three caches, each consisting of six one-word blocks. Cache D is direct mapped, Cache S is two-way associative, and Cache F is fully associative. Least Recently Used replacement policy is used.

For each of these caches, given the sequence of block addresses 0, 6, 0, 8, 6,

i) show the cache content after each block access,

ii) count the number of hits and misses,

iii) classify each miss according to the three categories in part (a).

The three parts carry, respectively, 35%, 20% and 45% of the marks.

End of paper