

Computer Science 773

Operating Systems

1997 EXAMINATION : ANSWERS AND NOTES.

GENERAL COMMENT : It was an easy paper (because I had no idea what sort of questions to set), but the difficulty in answering such a paper is to remember all the easy bits to say because they're too obvious.

If the results are anything to go by, you didn't learn much from the assignments. Perhaps it was the way I phrased the questions, but they look all right to me and the assessor made no comment.

PART A.

QUESTION 1.

Briefly describe PID (three-term) controllers and PLCs (Programmable Logic Controllers), pointing out in each case the type of control problem which the device is designed to solve.

A PID controller is designed for controlling continuous variables. It is a closed-loop controller, which receives a measure of the output of the plant which it controls. It provides a control signal which includes contributions from the error signal (the difference between the feedback signal and the required plant output), the integral of the error signal, and the derivative of the error signal. The proportions of the three components in the control signal are adjustable to suit the characteristics of the plant controlled.

[3 marks]

Continuous controllers are intended to apply control operations to a plant to keep some controlled variable at a fixed value (the set point) or following some predefined variation with time.

[2 marks]

A PLC is a programmable device which can receive inputs from a plant and return control signals to the plant. In early PLCs the signals were all binary, but newer models can commonly work with variables as well. The programme is declarative in nature, and defines the output in terms of the inputs using functions such as logical functions, comparisons of numeric values, and counters. The PLC must maintain the relationships specified in the programme.

[3 marks]

Sequencing controllers control the order of events, ensuring that a planned sequence of events take place in the correct sequence, at the correct times, and perhaps when certain conditions are satisfied.

[2 marks]

Generally well done. A surprising proportion of answers missed out things like getting input from the plant (particularly in the PLC case) and sending control signals back.

QUESTION 2.

Describe a closed-loop control system. Assuming that the behaviour of all components of the system is precisely known, and that a computer is used as the controller, what sort of analysis is necessary to determine the task which its programme must accomplish ? How is this task made more difficult by (a) time delays and (b) sampled data ?

In a closed-loop control system, the behaviour of the system is monitored by some sensor, and the sensor's output is used as information for the controller from which the control action needed to produce any necessary behaviour correction is calculated. The control action is then performed on the system, thus completing the loop.

Analysis of the behaviour of all the components of the system and their interactions determines the behaviour of the system as a whole. The control programme must be chosen so that this behaviour is the desired system behaviour.

[2 marks]

Time delays make this difficult because rapid response to changes in conditions becomes impossible. In a simple system, a correction signal might be applied for much longer than necessary before the result of the correction is available to the controller, by which time gross over-correction might have happened; the controller responds by applying a large correction in the opposite sense, and again receives no feedback until over-correction has again occurred. The result is oscillation. If the magnitude of the control signal is reduced to take account of this possibility, control response becomes much slower, and might not be sufficiently fast to maintain adequate control of the system.

[4 marks]

With sampled data, the feedback received by the controller is current when received, but is only received spasmodically. In designing the controller, one can then either assume that the feedback value remains constant over the interval between readings, thereby risking oscillation as with the delayed feedback, or (as in the Z-transform method) use the known behaviour of the system (in effect) to predict the behaviour while no input is received.

[4 marks]

Quite well done. Some people didn't say how you defined the task of the control programme; some described briefly what could go wrong but didn't say anything about how it affected the control task. Only one or two mentioned the possibility of interpolation with sampled data.

QUESTION 3.

What is a flexible manufacturing system (FMS) ? Identify the components of a FMS, and show how it differs from an assembly line and from a job shop, commenting on the differences between the components of the three manufacturing techniques.

A FMS is a manufacturing plant of several stages of manufacturing specialised to the production of a particular product, in which the progress of each product unit through the system is individually controlled and can in principle be varied from unit to unit.

[2 marks]

The components of a FMS are :

- 1 : Numerically controlled machine tools, possibly grouped into workcells;
- 2 : Materials handling and storage, for transport of product units through the plant;
- 3 : Computer control systems, to coordinate the operations of the other components.

[2 marks]

In an *assembly line*, manufacture is completely specialised to produce one product. All product items are identical; and there need be no provision for tracking individual items through the system. The machine tools used can be specialised for their tasks, and might not need control beyond starting and stopping for each unit; the materials handling system is very simple, and often reduced to a single conveyor belt; and there is little need for computer control.

[3 marks]

In a *job shop*, there is no specialisation; the available machines can be used at any time and in any order. All product items are assumed to be unique, so each must be programmed individually. For a computer-controlled job shop, numerically controlled tools are essential, workcells are less likely to be useful, versatile materials handling is essential, and computer control is necessary, and includes scheduling to a much greater degree than in a FMS.

[3 marks]

The first bits were well done, but not many people compared the components of the three methods.

QUESTION 4.

Explain how the nature of people's interaction with machines has altered with increasing automation. Give two examples of difficulties caused by this change, with brief explanations of their causes and possible remedies.

The direct effect of automation has been to replace people doing manual work by machinery. The people who remain in the systems must therefore deal with much more complex problems without the benefit of the expertise previously exercised by the manual workers.

[2 marks]

EXAMPLES :

Comparatively unskilled people must now operate complex devices. Clerical workers use photocopiers to accomplish tasks previously requiring skilled printers or photographers who knew what they were doing, and would sort out problems which arose. Clerical workers are unlikely to have the skills needed to sort out significant problems with photocopiers. Remedies within the control system are good diagnostic techniques, and an effective user interface (including manuals, perhaps) with which the operators can identify and perhaps correct the problem. (Within the broader system, an efficient service organisation is nice, but that isn't 773 stuff.)

[4 marks]

A plant controller in a factory can no longer rely on the distributed common sense of factory workers to identify or foresee failures in the process, and there is no one to notice that observable symptoms are linked with possible faults. A partial remedy is to convey more information about the current state of the plant to the controller. There is no obvious substitute for the common sense and experience of the shop-floor worker; though "intelligent" fault diagnosis programmes offer some hope, they are not sufficiently sophisticated to replace the vanished workers. The best that can be done is therefore to incorporate checks and procedures for foreseeable faults, and emergency procedures for inexplicable conditions.

[4 marks]

..... etc.

I should have defined "difficulties" more carefully. I'd thought of specific examples, as above; answers covered those, but also increased unemployment, social problems, etc. That was a fair interpretation of the question, and I marked them as well as I could. (Fortunately there's not much room to go into depth in a ten-minute answer !) Answers tended to be light on "possible remedies".

QUESTION 5.

Compare the task scheduling problems faced by general-purpose operating systems and by real-time systems, and show how differences between the two sorts of system can be exploited to develop a class of real-time scheduling techniques.

An operating system must schedule arbitrary tasks as they are submitted, ensuring that they do not interfere with each other, that operations within each task are carried out in the proper sequence, and that they are completed eventually.

A real-time system must schedule tasks from a previously defined set at defined times, ensuring that they do not interfere with each other, that operations within each task are carried out in the proper sequence, and that they are completed by defined deadlines.

[4 marks]

The much stricter temporal constraints on the real-time system demand much more precise task scheduling than is possible with a general-purpose operating system. On the other hand, because the set of tasks managed by the real-time system is fully known, more specific information about the performance of the tasks is available, and this can be used in scheduling.

Tasks to be scheduled can be classified in various ways according to their properties. For example, in rate-monotonic scheduling they are classified as periodic, sporadic, and aperiodic tasks, where periodic tasks must be executed regularly, spasmodic tasks must be executed at arbitrary times on request but have a known maximum load, and aperiodic tasks arrive at arbitrary times but can be delayed for a limited time. Given the known properties of tasks in these groups, a scheduling algorithm can (or can't) be constructed.

[6 marks]

The first part was a bit scrappy. I was looking for similarities (non-interference, proper sequence) and differences (known tasks, hard deadlines), and typically didn't get many. In the second part, I wanted stress on using information about the tasks; the refercne to "differences between the two sorts of system" was supposed to be a hint, but wasn't widely taken.

QUESTION 6.

Describe in outline a systematic method for producing a specification document for a system from a consideration of the required system functions. Explain the importance of this design step.

- 1 : Identify system inputs and outputs, defining the values required and any timing constraints.
- 2 : Record the characteristics of the input and output devices.
- 3 : Starting with the data required by the output devices and working backwards until all data are calculated :
- 4 : Identify each datum and the circumstances under which it is needed;
- 5 : Classify it (periodic, spasmodic, etc.);
- 6 :: Identify the function which calculates each datum;
- 7 : Tabulate the function for each relevant system mode.
- 8 : List possible faults which can affect each input, output, and function.

[6 marks]

The specification document :

- 1 : Demonstrates that the system can be constructed;
- 2 : Provides a basis for implementation;
- 3 : Provides a matrix in which any later changes in design and implementation can be recorded;
- 4 : Provides information on possible faults which can be used to develop additional fault-handling facilities.

[4 marks]

Not a popular question - only one person attempted it. (Great for marking !) He got full marks.

QUESTION 7.

List the features which you would hope to find in a programming language for real-time control systems, and which are not provided in conventional procedural languages. Comment on any difficulties you perceive in providing such facilities in a high-level language.

- 1 : Communication with devices through arbitrary interfaces.
- 2 : Timing functions : time delays, wait until, etc.
- 3 : Multiple processes.
- 4 : Execution on a distributed system.

[4 marks]

- 1 : Interfaces vary enormously from system to system. To insulate the code written for the "real" interface operation from the additional bits required to communicate through the interface, some means of encapsulating the peculiarities of the interface are required. This could take the form of an interface declaration (as in PEARL) or provision for writing specific interface-handling code when required.

- 2 : Either the language must provide its own operating system (or some equivalent thereof), or it is dependent on another operating system for timing services. A foreign operating system might incorporate its own, unknown, timing delays.
- 3 : The language must be constructed around a processing model which handles multiple processes and communications. The concurrent execution must be implemented either by the compiler or by a foreign operating system, which might not provide the detailed control which is necessary to manage the system effectively.
- 4 : The language must take account of communication systems, and the compiler should be able to compile complementary code for several, possibly different, machines.

[6 marks]

I thought I'd asked you to reproduce the Purdue Fortran recommendations, more or less, with a bit of distributed systems from PFL, and comment. Nobody did that. The result was mixed - some good answers, and some very scrappy ones.

QUESTION 8.

Identify the most urgent primary task of a real-time control system when a failure is detected, and describe it in terms of the system's states. Use the description to show what must be done to accomplish the task, and explain why it is difficult.

The most urgent task of a real-time control system after a failure is to regain control of the plant.

[2 marks]

A real-time system can (usually) be regarded as a finite-state automaton, with each state characterised by a certain set of processing activities which happen at some stage of the process being controlled. In each such state, control functions are provided to manage the system in an acceptable manner.

A failure can be regarded as a departure from the expected sequence of states, so that the control function needed to reestablish control is not known. To regain control, the controller must first identify the system's current state by measuring such parameters as are available. If this corresponds to a known state, the correct control functions can be used. If not, a means must be sought to move the system towards a known state by exercising any available operations; then when the known state is reached, control can be asserted again.

[5 marks]

This prescription depends on sufficient information being available to the controller to identify the state, the effects of operations on state variables being known with sufficient generality, the sensors working reliably after the failure, the operations working reliably after the failure, and so on. These conditions are satisfied in rather few real cases.

[3 marks]

Rather well done. Everybody knew the first answer, in practically the words I used. People who didn't talk about states didn't do so well.

QUESTION 9.

Define a robot, and explain how robots differ from other machines from the point of view of computer control. Distinguish between different elementary types of joint which might be found between links in a robot, and identify the joint coordinates.

A robot is a programmable machine which can move freely without continuous control from outside within some significant workspace.

[2 marks]

Robot control systems must provide for arbitrary motion in real space, while motions in other machines are almost invariably constrained by linkages, guides, jigs, or other devices to one dimension.

[2 marks]

The two basic elementary joint types are *revolute* and *prismatic*.

[1 mark]

In a revolute joint, the two links concerned can rotate relative to one another around the joint axis; that is the only degree of freedom. The joint coordinate is the angle of rotation measured from some arbitrary zero configuration.

In a prismatic joint, the two links concerned can translate relative to one another along the joint axis; that is the only degree of freedom. The joint coordinate is the distance of translation measured from some arbitrary zero configuration.

[5 marks]

This was the intended question. Unfortunately, at some stage in the processing the second sentence of the question disappeared. This gave me a very nasty shock when I came to mark it, and it was some little time before it occurred to me to check the question paper. Therefore, in the actual marking scheme the first two items above got five marks each - but I felt justified in requiring a fairly thorough answer for the full five marks.

Just "programmable" isn't enough; NC tools are programmable. I wanted something about free mobility for the first part, and multidimensional motion in the second.

People who tried both of the remaining questions (one forgot the second half) got reasonable marks.

QUESTION 10.

Describe the forward and inverse kinematics problems in robotic systems. Define the reachable and dexterous workspaces of the robot, and show how they are related to the existence of solutions to the inverse kinematic problem.

The forward kinematic problem is the problem of working out the position and orientation of a robot's end effector given the position of the base and the joint coordinates.

The inverse kinematic problem is the problem of working out the joint coordinates given the position of the base and the position and orientation of the end effector.

[3 marks]

The reachable workspace of a robot is the set of points in space at which the end effector can be placed in at least one orientation.

The dexterous workspace is the set of points at which the end effector can be placed in every orientation - or, more practicably, with some predefined set of orientations.

[3 marks]

A point is within the reachable workspace of a robot if the inverse kinematic problem has a solution for the coordinates of the point in at least one orientation.

A point is within the dexterous workspace of a robot if the inverse kinematic problem has a solution for the coordinates of the point in every orientation.

[4 marks]

Mostly well done; a few people missed bits out, and the relationship between workspaces and solution was not always clearly described.

QUESTION 11.

List some advantages of vision as a sensor for robotics. Describe the problems of machine vision with ambient light, and give an example to show how the use of structured light can simplify a problem.

Vision provides :

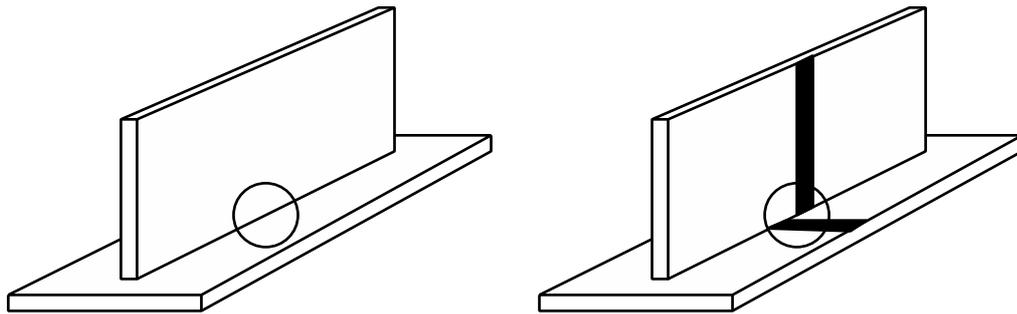
- 1 : A "universal" sensor, matching the general-purpose nature of robots;
- 2 : High precision, allowing precise control;
- 3 : Operation at a distance, so the sensor can be mounted on the robot.

[3 points]

The use of vision in ambient light is made difficult by the enormous volume of information in a picture. Without some artificial assistance, it is very difficult, and, therefore, time-consuming to identify objects in a scene, and - even if the identification can be managed - the time required is likely to be too long for practical applications of robotics.

[3 points]

Structured light can help by adding information which can isolate parts of the picture for specific attention, or in some way relate points in the received picture, assisting interpretation in both cases. An example is the use of a planar light source to show the width of a gap between the end of a sheet and a planar sheet; compare the views in the two circles in the diagram, where the dark stripe represents the illumination from the light source :



[4 points]

Most people knew more or less what this was about. Defects were giving only about one advantage, and not explaining the ambient light difficulties adequately, or at all. The problem isn't usually that you can't do it in ambient light; computer vision is really rather effective, and there are ways of overcoming an impressive variety of problems. The problems are speed. And there were a couple of answers which suggested that their authors hadn't understood how the planar light beam worked. (Everyone used the same example.)

QUESTION 12.

Present in outline an argument to the effect that it is impossible manually to programme a robot for realistic behaviour in the real world. Briefly describe the "subsumption" architecture, and say, with reasons, whether you believe it to be an improvement on conventional programming methods.

"Traditional" robot control systems are partitioned by function : there are more or less independent sections for perception, world view interpretation, planning, motion control, and so on. As robots deal with more and more complex environments, all these tasks become more demanding, and must be executed more precisely and faster, but there is little guidance as to how this should be done. Additional code must be written to cope with new circumstances as they arise, and to deal with a realistically complex world so much of this is necessary that the problem becomes insuperable.

[3 points]

In the "subsumption" architecture, this problem is to some degree circumvented by partitioning the system by behaviour rather than function. The programme is structured as a hierarchy of behaviours, each incorporating the specific perception, planning, etc. needed for its own operation. In action, higher behaviours depend on lower behaviours, interacting with them or overriding them when appropriate. In case of failure of a higher behaviour, lower behaviours will continue to operate, so catastrophic breakdown is unlikely.

[3 points]

(OPINION : any relevant and reasoned argument will do. For example :)

As a means of structuring a control programme, the "subsumption" approach is of interest, and is an alternative to more conventional means. As a way of avoiding complexity, it is less successful; to construct a robot which will operate in a realistic environment, so many behaviours will be necessary, with such complex interactions, that it is not at all clear that much is gained.

[4 points]

Too many answers to the first part reduced to "we can't do it because it's hard", which is a tautology; I wanted at least some indication of *why* it's hard for full marks. Most answers to the middle bit were all right. For the last part, I did want some reasons, and didn't always get them.

PART B.

QUESTION 13.

Describe the function of an interface adapter chip, such as the VIA. Explain in outline how it is used and how it is attached to the processor.

Interface adapter chips are devices which provide useful input and output operations for microprocessors.

To manage the communication, each adapter acts as a small and specialised independent processor. It deals with incoming signals in simple ways (typically by latching them in a register or by counting them) and makes the information available to the processor on request. Outgoing signals are written to the adapter and transmitted. Other functions might be provided, such as timing and interrupting the processor at significant times.

[3 marks]

To use the adapter, the processor must communicate with it in two ways : for communication, and for control.

In communication, the processor sends information to the adapter to be transmitted, or receives information which has arrived.

In control, the processor sends instructions to the adapter to set it up so that it performs certain required functions in certain ways, and reads status information from the adapter to determine what has happened.

[3 marks]

The adapter is attached to the processor by memory-mapping (or input-output mapping), which avoids using valuable pin connections. Communication occurs mainly through the microprocessor's address (or input-output) and memory buses. I shall assume memory-mapping from here on - input-output mapping is essentially identical except for the processor bus used.

The connection is made so that communication with the adapter is effected by using the processor's memory addressing functions. A memory range not used by real memory is chosen, and the address bits common to the memory combined by logic external to the adapter chip to generate a chip-select signal. The remaining address bits are presented to the adapter as instructions. For example, with a 16-bit address bus, it might be decided to reserve the memory addresses FFA? for the adapter; then the top 12 address bits would be combined to return a select signal when FFA is found, and the low four bits connected to the adapter's control pins.

The memory read-write signal is also connected to the adapter, and is used when the adapter function selected by the instruction requires the use of the data bus. Then the state of the read-write signal might determine whether information is taken from the data bus, or placed on the data bus.

[4 marks]

Disappointing, really. There were occasional mentions of being connected to a bus, but no detail - and nothing at all about how to use it.

QUESTION 14.

Explain why accurate timing is essential for reliable communication between two devices communicating through an RS-232 serial communications line. Describe one method by which such a device can determine the parameters used by a communicating partner, and make clear what conditions must be satisfied if the method is to work reliably. (Assume that the device must communicate through a conventional serial interface adapter; direct access to the line signals is not possible.)

RS-232 is a protocol which species how single characters are encoded for transmission, and each character is transmitted independently. On detecting the beginning of a character, the receiving device must start a clock and sample the signal at intervals depending on the transmission speed. The sequence of a number of values (depending on the communications parameters) is the received code which must be interpreted to reconstitute the character. If the receiver's clock runs more slowly than the transmitter's, it will detect only a fraction of the original character's signal, and the detection sequence might continue until the next character transmission has started, so characters might be lost. If the receiver's clock runs too quickly, it will finish sampling before the transmitter has finished sending, and further signals received as part of the original character are likely to be interpreted as the beginning of a new character, so too many characters are received.

[4 marks]

Parameter identification can be achieved in three stages.

- 1 : Determine the transmission speed : receive characters from the transmitter at different data rates. Ignore framing errors and parity errors. When you receive the expected characters, you have the right data rate.
- 2 : Determine the parity : Try different parity settings until parity errors disappear and the expected characters are received. Ignore framing errors.
- 3 : Determine how many stop bits : adjust the number until the framing errors disappear.

[4 marks]

For this method to work, it is necessary to be able to stimulate a response from the remote device when required, the response must be known, and it must contain characters of both odd and even parity.

[2 marks]

Again, a deep reluctance to explain why poor timing gave a bad result.

QUESTION 15.

Describe the components of a simple analogue computer, and explain why the output of an adder is inverted. Present a diagram to show how a suitable set of components can be connected to simulate the behaviour of a system defined by the equations

$$\begin{aligned} dx/dt + Ay &= 0 \\ dy/dt + Bx &= 0 \end{aligned}$$

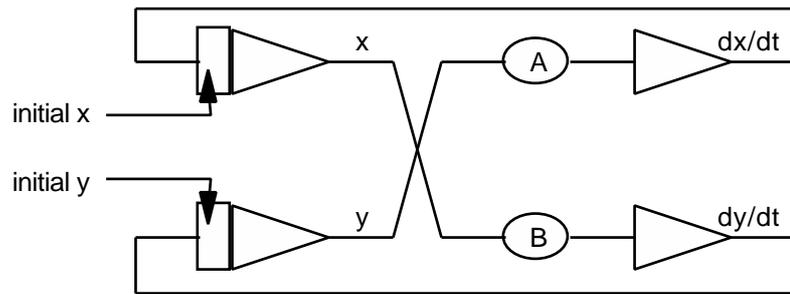
given initial values for x and y .

An analogue computer is composed of adders, integrators, and potentiometers, with some means of connecting them together in arbitrary ways. It must also have means of setting initial values of integrator outputs at the beginning of an experiment.

[2 marks]

The output is inverted because the operation of the adder depends on the sum of output and input being very close to zero. It would be easy to add an inverter stage, but that would approximately double the number of components in the system and potentially introduce additional errors.

[3 marks]



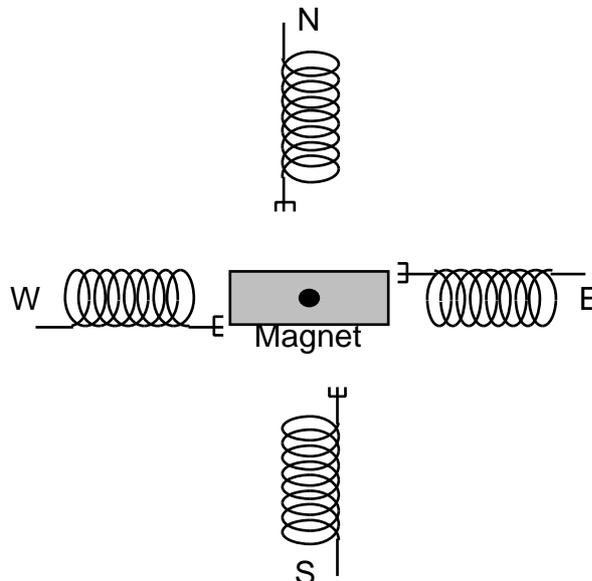
[5 marks]

Nobody did this one. Eaay to mark - but a waste of time setting it.

QUESTION 16.

Describe the operation of a simple stepper motor, explain why stepper motors are useful, and show how a stepper motor's input signals must be controlled to give continuing motion in one direction.

A simple stepper motor might be constructed as in the diagram :



It is operated by energising the coils so that the magnetic field at the rotor magnet is shifted to drive the magnet round in the desired direction.

[4 marks]

Stepper motors are useful because they provide precise positioning without the need for position sensors and feedback control.

[3 marks]

For the motor in the diagram, this sequence of input signals will give continuing motion in the clockwise direction :

N	E	S	W
+	0	-	0
0	+	0	-
-	0	+	0
0	-	0	+

[3 marks]

Also non-impressive, for the most part. If you knew how stepper motors worked, you didn't think it worth telling me about it. As I asked for it, that was poor tactics.