Ug Exam Performance Feedback
Second Year - Semester 1

Richard Neville:
Over the lecture period [semester 1] three laboratory components was undertaken. The results from my labs [lab1RN and Lab3RN] showed the students had a moderately good understanding of a subset of concepts presented during the lectures [prior to these laboratories].
The exam results enforced this view – with a good spread of results. The exam results show the module's learning outcomes were sufficiently exercised. Both the courses and the lecture learning outcomes (LOs) were tested and the students [as a whole] showed [relatively] good comprehension of the LOs. The exam and laboratories validated the pedagogical approach.

Chris Kirkham:
1g) Many answers didn't show any appreciation of the fact that there are some operations (in the instruction set of the computer) which do not work in user mode, but do work in kernel mode.

1h) The keys words in the question were "implements" and "input". The former does not mean express, i.e. using a less-than symbol, but asks for an explanation that stdin in the process created is set-up to read from the file. Too many answers spoke instead about output redirection, or even pipes!

1i) The registry in Windows is NOT where the values of registers are stored!

ij) Yes - the microkernel is a small kernel; the man in the street might guess that. To get full marks for explaining the microkernel approach I wanted to see an indication of what facilities a microkernel would provide, while being small! No - it doesn't compress its data!!

2a) "Describe the actions ...." Too many answers told me lots of things about context switches, but didn't answer the question. Of those who did tell me about the actions, too many were vague about where information would be stored (in the PCB - not on the stack!).

2c) "outline the proof" does not mean illustrate with an invented example! And "lowest average turnaround time" should be read as compared to any other order of running - so there was no need to drag in round-robin, etc.

2d) turnaround time = processing time (fixed) + waiting time.

2e)(ii) Having given an outline proof that SJF produced the lowest average turnaround time, it was rather surprising that some answers believed the modified scheduler reduced it!

3a)(ii) Kernel-level threads do NOT run with kernel privilege!

3b) Too many failed to indicate that the program created 20 threads, each of which would output two lines. So what (i) wanted was a description of these 40 lines of output - what was possible about the order of them, and what not.
Question 1
This question was about real-time multi-media using different protocols. It was shocking, even after a full term of talking about TCP and UDP, how many students do not know the difference between them or even the basics of how they work. These protocols are the backbone of the Internet!

I expected the material on 2nd life to be more challenging but it was clear that knowledge of this topic was at a similar depth to knowledge shown for other parts.

I expected to have a high average for this question. Almost everybody attempted it.

a) Answers gave too much general information about TCP with either no direct connection to real-time multi-media or with the connection left unclear. The cost of creating a connection is normally very small and only matters if an instant start to playback is required. The real problems using TCP for anything real-time, e.g. XX and YY, were not mentioned in most answers.

b) I was amazed how many students still think UDP has a Window concept (TCP does!) or that there is a connection when using UDP! Some answers indicated that only UDP can control/vary packet length whereas almost all transport protocols can do this. Others think UDP with RTP has retransmission and acknowledgement - they do not! Still others, think UDP is TCP and perhaps visa-versa but their TCP answers did not make this clear, if it was clear then mixing up mnemonics is excusable in an exam to some extent.

c) This was answered reasonably by some students who had read the notes but was very weak from others.

Question 2
This was a well sign posted topic (I thought). The few students that attempted this question struggled. In previous years this material has produced very good marks. About 1/4 of students attempted this question.

a) This is in the lecture notes. 50% of marks were for pictures and 50% for explanation. Answers were very weak. A question like this has appeared frequently in past papers.

b) This is in the lecture notes and the book. Clearly almost nobody had read the notes. Most seem to understand the use of the start/end sentinel. Some rounded to octets which was wrong as HDLC is bit oriented. Some repeated the sentinel instead of using bit stuffing to avoid the sentinel appearing in the data. A couple of candidates thought this was CRC encoding?

c) This was carefully explained in the lecture and was asked in a very similar manner to the lecture explanation. Answers had lots of weird assumptions mainly based on the behaviour of wired networks rather than wireless ones.

Question 3
a) The question asked for the answer to be illustrated with examples, many students lost marks by failing to give any examples. The question was focussed on application data, many answers were at a network rather than application level. Several answers were also about topics not related to the question; for example, proprietary protocols, data buffering, data entry, or plain text verses encrypted transmission. Some just repeated question, e.g. “for implicit data typing, the value of data and the type of data are implicit”.

b) Most answers missed the point that a real network is composed of many machines with different hardware and software each which their own native formats for values, and need to make these machines operate together.

c) The question asked for the answer to be illustrated with a diagram that showed where conversions were performed; many students failed to give a diagram and most of those that did, failed to indicate where conversions need to be performed. As the application describes runs on different types of hardware, TCP and stop-and-go are not the strategies that allow the data to be correctly received. Given the variable size of the array, the explicitly size of an instance transferred must be part of the strategy; a point many students failed to identify in their answer. The question was looking for the use of simple bookwork, canonical or receiver-makes-right strategy, applied to a specific case; some students lost marks by giving generic answers that did not consider the specific case. Other, non-sensible suggestions included using multiple packets (all of the message could be contained in one packet; sending the string, integer array size and integer values in multiple packets is creating multiple message types) and encoding the integers as strings.

d) Much of the strategy that should have been described in part c) was about support flexibility; this part of the question removes any need for flexibility. However, many answers still about flexibility. The question specifically asked for the effect of the strategy described in
the previous part in relation to bandwidth and CPU cycles, but many answers failed to mention these.

e) The question was about extensibility, which is extending or altering the functionality provided. Many of the answers concentrated on scalability, which is increasing the number of machines in a network.

Question 4

a) The question clearly stated “in terms of transport level protocol”, many answers lost marks by including information about routers (too low level) or key checking (too higher level).

b) Question asked for demonstration of how timeouts, sequence numbers of retransmissions made the transfer of a particular set of data reliable. It is not ask for an explanation of these concepts which several answers gave. Also, the question used real values for times, which should have been used on producing answers. The description indicated that five bytes were sent every four seconds, many student incorrectly interpreted this as fives bytes per four seconds. Other common errors that lost marks included putting sequence numbers on packets not bytes, acknowledging bytes 11-15 before bytes 6-10 retransmitted and successfully received, at time 8 A still sends bytes 11-15 despite fact that no ack for bytes 6-10 has been received, and poor addition (if something is sent at time X and the propagation delay 1, it arrives at X +1, not X+2 or any of the other random values given).

c) Question was about purpose of window to improve utilisation and use in flow-control, many answers did not cover these points or described what there were not why they worked. A number of answers assumed that a transmission window was a period of time, when it is a number of. A number of answers also lost marked by not stating that it is the receiver that varies size of window dependant on the level on its input buffer. Not sending acks, causes retransmission, so poor use of network resources.

d) The question required to the use of CIDR to allocate the address, some students lost marks by using subnetting. Each network gives 254 address, so networks C, E and G needed more than one class C network address allocated, many students only allocated a single class C network. When allocate block of class C network addresses to a physical network block needs to be of size $2^n$ and start on an appropriate boundary this point was missed by most students. It is also necessary to allocate networks, not parts of network, which several students did. Several students seemed to think that there were less addresses than machines, as there were 64*256 addresses, this was not true.
Common mistakes:

Question 1.

a) (i) Problem with processing conjunction of three clauses, consequently corresponding subformulas and their positions are incorrect.
   (ii) Initial formula is considered to have position 1.
   (iii) Problem with polarity 0.
   (iv) Some answers assert that the initial formula is of polarity -1.

b) (i) Mixing the empty clause and the empty set of clauses.
   (ii) In some answers, an incomplete interpretation is given: the value of r is missing.
   (iii) Splitting on a variable when unit propagation can be applied.

c) (i) Propositional variables are not introduced in some cases. The domain axiom is written using equalities and disequalities i.e. P = a and P != a.
   (ii) Many forgot about the domain axioms.
   (iii) Others made mistakes in the domain axioms and in the transformed initial formula as well.

d) (i) The answer yes for 1 is wrong
   (ii) One of the states missing in the paths for 2, 3.

Question 2.

a) (i) Formula transformed in DNF instead of CNF.
   (ii) Problems with the distributivity rule.

b) (i) Probability is counted incorrectly: based on number of appearances of variables in all clauses together, without considering the probabilities of choosing each of the clauses first.
   (ii) In some cases the false clauses are found incorrectly.

c) Incorrect answers.

d) Problems with the definition of a Path. In many cases a transition (or number of transitions) is given for the answer.

Question 3.

a) (i) Mistakes in counting number of satisfied clauses.
   (ii) In some papers the flipping variable is chosen among variables satisfying the smallest number of clauses after flipping.

b) (i) The order on variables is different from the one given in the problem.
   (ii) Some answers miss final merging of nodes in the OBDD.

c) Incorrect paths given.

d) Some students do not use any connectives when writing complex LTL formulas.

Question 4.

a) (i) Problem with branching in tableaux.
   (ii) Problem in applying tableau expansion rules in general.

b) Some answers miss final merging of nodes in the OBDD.

c) A wrong transition graphs given.

Question 5.

a) (i) After splitting, some students evaluate the formula as in the case of DPLL with a propositional formula. I.e., the evaluation at each node is done without considering the quantifiers.
(ii) Again, some students mix the empty clause and the empty set of clauses.

b) (i) Some students do not use any connectives when writing complex LTL formulas.
(ii) The wrong formulas are given.

Q1.  
   a) Many students were too vague about the essential components of a load/store architecture (Reg Bank and ALU).
   b) This part asking for a description of register-to-register, load/store and branch instructions was well done by those who understood what the architecture was doing. These gave a description of accessing operands from the Reg Bank, followed by an ALU operation, and detailed the writeback performed.
   c) Most students knew that a 'hold' register was needed in a pipelined system but then failed to say exactly why it was needed.
   d) A 2-stage versus 3-stage pipeline was not well done - comparisons made did not compare like with like. A more considered answer would have resulted from drawing a timing diagram.

Q2.  
   a) The different behavioural models were not clearly expressed in some cases.
   b) This part asked for a Verilog *module*. Thus a full module was expected including all declarations and with correct syntax that executed continuously. The inner functionality was usually OK but many of the declarations and the ability to continuously monitor the inputs for changes was missing.
   c) The features of the STUMP's state machine were reasonably explained except for 'reg' (which maintains variables constant between assignments but is not necessarily implemented with flip flops).

Q3.  
   a) The definitions of controllability and observability were well done.
   b) The need for reset was not so well expressed. Parallel system initialisation to a known starting point were expected.
   c) The description of the scan path was reasonably well done and most drew a relevant diagram.
   d) The testing of a scan pathed STUMP was not well done. A detailed description of the scan path as applied to performing and checking the different phases of a register-to-register instruction was required. Instead, answers were vague or tended to generalise about scan paths.
   e) The comparison of the scan path testing approach to that used in the lab was reasonably well done.

Q4.  
   a) The description of a two-phase clock was reasonable although the diagram presented often showed both clocks on simultaneously.
   b) Most were able to sketch the logic to generate a two-phase clock but the explanations usually failed to describe what happened when the input clock changed and hence were unable to explain how the non-overlap time occurred.
   c) The increase in non-overlap time by using a chain of invertors in series with the NOR gates of the generator logic was not clearly described by most.
   d) and e) The description of clock skew and its minimisation were reasonably well understood and described.
Ke Chen

Regarding short questions (f) - (j) in Sect. A, more than 50% students achieve the satisfactory performance while a number of students do not seem to gain the essential understanding for book knowledge that had been highlighted and repeated during lectures and tested in given exercises. For example, a large number of students could not distinguish between “margin” and “decision boundary” in SVM. Even though a mini-project was done and the same question was asked in exercises, over 25% students still could not tell the underlying difference between Bayes and Naïve Bayes classifiers.

For two comprehensive questions, i.e., questions 4 and 5, in Sect. C, most of students chose question 5 that tests how to choose a distance metric for medical data and how to further apply it in the agglomerative algorithm for hierarchical clustering. A quite number of students failed to identify the asymmetric binary distance to be a proper choice by randomly selecting one he/she was familiar with. In addition, several students did not fully understand the cluster distance and the cluster merge concepts in the agglomerative algorithm. On the other hand, most of students who attempted question 4 seem to understand the relationship between Perceptron and SVM well although only few students managed to justify their solution completely with the properties of SVM. It is also worth mentioning that several students did not attempt either of questions in Section C and, unfortunately, about 15% students received zero mark from Section C.

In summary, the overall result reflects what students actually achieved from this course unit. Based on this examination, I believe that, to a great extent, a lack of revision (failing to do exercises given) and a low lecture attendance rate (normally below 50%) seem responsible for the poor performance and the lack of understanding while other factors need to be investigated further.

Gavin Brown

Part I. The short answer questions were answered well in most cases, with no particular pattern of mistakes.

Part II

Q1. In the question about the perceptron learning rule, the main point which students found difficult to answer well was mixing up the update rule with the activation rule. The correct learning rule uses (target-output) whereas in many cases students wrote (output-input), apparently mixing up the terminologies. The latter part, the 1(c) case study, was answered well. The main point missed by most students was that removal of outliers can be beneficial for the perceptron, hence removing the millionaires from the dataset was not a problem.

Q2. Very few students (~10%) chose this question. The main issue seemed to be with describing the intuitive working of the ID3 algorithm, rather than knowing the technical details of "information gain" - this is helpful to guide teaching strategy for next year.