General comments: The marks showed a wide spread of performances - with many students getting good to excellent marks, but also a cohort with very low marks, who seemed neither to have grasped the basics nor to have bothered to work on the material, or revise properly or attend lectures.

Question 1: This is mostly lectured material/bookwork on the basic concepts of complexity analysis. All students who tried to master the material got goodish (or better) marks. Marks were lost through imprecision in answers. The final part (an exptime algorithm) was straightforward - some had been seen in tutorials (eg standard recursive fibonacci) and others described in the lectures (eg exhaustive, unbounded backtrack algorithms). Again marks tended to be lost through imprecise answers.

Question 2: The first half was lectured material/bookwork, and again students who mastered the material got good marks. Quite a few students couldn't reproduce an algorithm and its analysis with sufficient precision to get full marks. Some gave the wrong sort of algorithm - which got zero marks. The second part was to devise an algorithm from an outline. Many understood the algorithm required and good marks were awarded if a precise enough description was presented.

Question 3: No feedback provided - see GDG

Question 4: No feedback provided - see GDG
Question 1, parts a-e:

This was a compulsory question. There were no particular problems.

f) Generally well answered.

g) Too many confuse "directory server" with a file server for a hierarchical file system.

h) Too many told me why timeouts in general were a good idea! And there was some confusion about the difference between the timeout on the response to the initial "election" message and that on the result, e.g. while waiting for the "coordinator" message.

i) Inventing a common-sense meaning for "callback promise" got no marks! In AFS it has a very particular meaning.

j) Some wanted to explain the purpose of a digital signature, whereas the question wanted an explanation of the presence of the secure digest in the signature!

Question 2:

All students except 2 answered this question. There were no particular problems and the average mark appears to be reasonable. Some of the common mistakes: a) several students failed to go beyond the obvious answer that "latency is not zero"; d-i) the vast majority of students failed to realize that the size of the problem was so small that exhaustive search could be used and they applied heuristics.

Q3

a) Note - clock synchronization algorithms were not required!

b) Not well answered in general. The point most missed was that logical time gives a spurious total ordering of events, because time is represent as a single integer, whereas vector time gives a more realistic partial ordering.

c) The question describes "totally ordered multicast" - so all the servers deal with requests in the same order. Too many answers invented ways in which some of the servers would not deal with some requests, or invented timeouts and resends.

d) There are two distinct possibilities: the message lost is a request from a client to a server, or the message lost is an ack from a server. To get full marks, both should be considered!

e) The most convincing way to show why the assumption about messages not overtaking is necessary is to invent a simple sequence in which single overtaking leads to an undesired result. Few did this!

Q3: A small percentage of students attempted my question. In general the students answered the question reasonably well but there was some confusion about the methods of describing resources and how an ontological approach could help. The sections on web services and REST were generally better attempted.

Comments:

No feedback provided - see setter
Question 1.b.
The question was to find a model using the pure literal rule. Some students tried to solve the problem using other methods, for example, splitting.

Question 2.a.
There was a typical error: some formulas on branches were not completely simplified so that after simplification they still contained subformulas containing true or false. In addition, many solutions claimed to have a model for this (unsatisfiable formula). This error was easy to prevent by checking whether the claimed model is indeed a model (e.g. by computing the value of the formula in the claimed model).

Question 2.b.2.
There were several answers that the probability is 1/2: it is 1/2 after 2 steps but not after 2k steps.

Question 2.d
There were too many mistakes in applications of tableau rules.

Question 3.b
After finding that
\[ p_1 \lor p_2 \]
\[ p_0 \lor p_2 \lor p_4 \]
are the candidate clauses some students assumed that every occurrence of a literal in these clauses contributes with equal probability, for example that the probability of selecting \( p_1 \) is 1/5. The way WSAT works is that it first selects a clause (with an equal probability among all clauses) and then a literal in this clause. So the probability of selecting \( p_1 \) is 1/4.

Question 3.d
Some answers used formulas \( x' = x+1 \). The language of PLFD only has atomic formulas of the form \( x=v \), where \( x \) is a variable and \( v \) is a value, so one cannot use expressssions like \( x+1 \) in formulas.

Question 5.a
In some solutions positions were omitted

Q1) Not attempted by many students. Those that did attempt it lost marks by not showing how they had worked out the final scaling for drain current and gate delay. The effect of track delay not scaling was understood, but students found it tough to calculate the relative track to gate delay.

Q2) Most students who answered this question lost marks on the calculation of the delay through the exor logic circuit. Many failed to realise that the delay through three inverting gates would mean taking the sum for the same edges on the input and output gates, but the opposite edge for the central gate. This was despite having done a similar example in the examples class. All other parts of the question were generally answered very well.

Q3 & Q4: the averages were high and in general the questions were answered well by the students.
This course is designed as a group project exercise which counts for 50% of the final mark with the exam counting the other 50%. Last year was the first year of running this course and in the light of experience, the group size was reduced from 4 to 3 members (last year, one of the tasks was much easier than the others). However, this was not entirely successful as when groups fell apart, there was not enough critical mass for the remaining group member(s) to act as a group. There were some complaints that the project was too open-ended and that insufficiently prescribed. Nevertheless, those that took the lab seriously did well. The main problem was that some students were not disciplined enough to attend the lab sessions regularly (or at all!!!) and thereby threw away a large percentage of the marks that they could have gained for the course.

The exam results were much higher than I expected. As a group, they did not seem interested in interacting during the lectures at all and I had expected poor results. Part of the reason for the high scores are:

a) The granularity of the questions was such that a student could either get 0%, or 100% for each part of the question. In general, most students answered with an imprecision that should demanded less than 100%, but more than 0%. I tended to err on the generous side. Next year, I shall revise the exam so that more discrimination is possible.

b) The exam was very similar in content and context to last year’s exam (which I did not set). On marking the exam, I realised that many of the questions depended on the same material and did not really test the students in depth. Again, I shall remedy this next year.

I would be quite happy for the marks to be scaled down by some amount - perhaps related to the average score of the cohort of students taking the course - I would guess at between 5% and 10%
OVERVIEW

COMP20312 was taken by 117 undergraduate students mainly from the School of Computer Science but also including BCS, ABIS, with BM, Geology with IT, Physics with IT, and 3rd year Mathematics students. This year however there has been one major change regarding the exam format which has had an impact on the examination results for this year. Firstly, the examination has now been changed into two parts, one compulsory - comprising 20 multiple choice questions (MCQ) valued at 2 marks (incorrectly see later) each; the whole representing two conventional 20 mark questions (again in error); and one part comprising 2 longer style (conventional) questions of which two must be answered. Part one aims to test Blooms (Cognitive) Taxonomy on Knowledge and application while part two aims to test Blooms (Cognitive) Taxonomy on Comprehension, Application, and Synthesis over each question. The effect of these changes can now be seen in this summers 2008 examination.

Students answered 1 out of 2 questions with topic areas spread through all questions avoiding the problem of answering single topic areas and better testing the spread of a students knowledge. I am happy to say that the average standard deviation per student over the questions was only 3.0 meaning that students scoring low - scored low over all questions (allowing for the rubric error), and students scoring high, scored high over all questions. Some anomalies did exist but these were mostly due to students running out of time and not answering all parts of their last question.

Part A
Questions 1-20 (2 marks each - in error see later)
Marks: Average 15.87 / Mode 17.00 / Median 17.00 / Average % 79.36 / Standard Deviation 8.11 - Answered by 117 (all) students.
This was the highest scoring question mainly due to the fact that each question investigated 'bookwork' and application of technique. However, the question again skewed the results slightly because the conventional 3 questions previously allowed a total of 12 marks to be earned for bookwork overall - this year we allowed 20 but over a much broader spectrum of topics. Last year I said 'Skewing will not occur next year as I intend to create another 10 part MCQ testing Blooms (Cognitive) Taxonomy - Application (re-evaluating laboratory and examples clinic work) and move the testing to 2 large MCQ questions testing all topics; and then a choice of 1 (from 2) conventional questions testing Blooms (Cognitive) Taxonomy - Comprehension and Synthesis.' I did this with disastrous results for the Average (see below).

Part B
Questions 1 (20 marks)
Marks: Average 11.81 / Mode 13.00 / Median 11.00 / Average % 59.06 / Standard Deviation 3.28 - Answered by 69 students.
This was the third most answered question and tested relational schema's, functional dependencies and normalisation, and database performance. This part of the question was answered very well by most students. The main problem here was with regard to the 'original thought' (synthesis) question which always proves problematic.

Questions 2 (20 marks)
Marks: Average 11.02 / Mode 14.00 / Median 11.50 / Average % 55.12 / Standard Deviation 3.85 - Answered by 42 students.
The least answered question but those that attempted it did well. The questions tested transactions, the SQL language (application in detail), and ER diagramming from an original thought aspect. The SQL question was answered well, again the original thought question proved problematic.

PART A MARKING RUBRIC ERROR
This year I have unfortunately introduced an error into my databases examination. Last year I had 1 multiple choice question of 10 parts with two marks per part, and the choice of two from three essay style questions, making the multiple choice question and the two essay style questions 20 marks each giving a total of 60 marks; each question representing one third. This year I increased the multiple choice questions and intended there to be 20 multiple choice parts with a choice of one from two exam style questions representing a 50-50 split, however inadvertently I did not change the marking rubric for the multiple choice questions and left it at two marks per question meaning there is now a 65-35 split. This means that the
average mark for comp20312 is approximately 72% however with the 50-50 split this is reduced to 66% (close to last year). Even so I will be making changes to this exam with regard to the marking of the multiple choice questions and their value for next year. However, this year I would like the external examiners to consider allocating one mark for each correct answer of the 20 parts in the multiple choice question.

SUMMER 2007 EXAM CIRCULATION
My database exam from last year has been removed from the examination hall and scanned, this would not be a problem, however, it contains multiple choice questions and these questions have been repeated this year as is normal in cases such as this. There are also 10 extra multiple choice questions (20 total) and the essay style questions are similar to last year but have been reformulated. An e-mail was sent by facebook, certainly to eight of my students, over the weekend before the exam. The release of the exam paper for last year has had no major effect on the first 10 MCQs or the essay style questions. I do think that we need to keep an eye on the processes which has allowed this information to be removed from the examination hall, and I will of course reformulate the questions for 2009.

IN SUMMARY
I was very pleased with the students performance this year (well done) and I am further convinced that practical exposure to subjects, active learning, the full unit text (no ad-hoc lecture notes), the weekly SAQ's, and a change in the exam style (as stated above) is the correct direction for the unit. However, with regard to the examination I intend to section MCQs into different levels of difficulty and allocate marks accordingly - so book-work may get 0.5 marks each which application of techniques may get 1.0 - I will investigate the current pedagogic thinking on the subject in advance to inform my decision on this point. Hopefully, this will increase the students comprehension of, and practical exposure to, databases in support of their future careers.
a) Although there were some very good answers to this question, no one managed to get full marks. Probably the main cause of lost marks was answers which were too vague and general, and were not related specifically to the scenario given in the question. For example, some students correctly identified SPI as being a suitable approach in scenario iv, but then wrote half a page of text describing the general benefits and costs of this approach, rather than picking out the specific features of the scenario that meant that the approach was sensible in this case. In fact, in this scenario, the company had been trying to improve its processes for some time, but these ad hoc attempts had failed. The specific role of SPI in this context is to give a sensible route for process improvement, so that the improvement actions undertaken are those that are likely to succeed and provide immediate and visible benefits. A reference to this aspect of SPI was needed for full marks.

In terms of costs, I was hoping for answers that compared the costs to the benefits, and showed that they were worthwhile for the gains that the company would likely experience if the chosen technique was implemented. Few people discussed this issue explicitly and students tended instead to list the general costs of the approach. Such answers earned some marks, but only answers which specifically tied the costs to the scenario could be awarded full marks for this part of the question.

Several misconceptions were revealed by some answers:

+ Some answers suggested that a defect tracking system would need to be implemented from scratch (i.e. writing the code for the defect tracking. Many commercial and open source defect tracking systems already exist. It is unlikely that the companies described in the scenarios would have such specific requirements for defect tracking that they would need to build their own system.

+ A number of students seemed to be mixing up "checklists & templates" with "coding standards". Although some code-level checklists are available, most software engineering templates are for document deliverables rather than code deliverables. Templates will not help you to force new members of staff to adhere to your coding standards. Nor will templates help to enforce use of a consistent methodology, except in terms of the set of deliverables produced. Templates can help to avoid omissions (i.e. forgetting to produce some deliverable or part of a deliverable) but they say very little about the process by which the deliverable is created.

+ SPI is a very heavyweight technique, and it is unlikely to be a cost effective way of tackling problems that are specific to one development team. SPI is more appropriate when problems are systemic and process-related; that is, when a problem is common to multiple teams within a company.

+ A small number of people stated (erroneously) that FTR is only suitable for use on code level deliverables. In fact, the opposite is the case: FTR (and reviews in general) are the best technique we have currently for reviewing the correctness of non-code-level deliverables. It is here that the benefits of reviews are really felt, although of course they can be applied to code level deliverables too.

b) There few only a few really excellent answers to this part of the question. Although most people who answered it had the right general idea, marks were lost due to a failure to state clearly exactly what quantity would be measured, or due to a choice of measure that is useful but impractical. For example, a number of answers suggested that it would be good to measure "market share", without actually specifying what quantities would be measured to assess this. Others suggested measuring the "cost-effectiveness" of improvements to processes, "user feedback" and "rate of satisfactorily fixed bugs" but it was unclear what was actually meant by these terms. How do you know which bugs have been satisfactorily fixed? For full marks, it was necessary to specify quantities that are precise, unambiguous and clearly measurable. For example, "number of bugs reported on product A per month" is an example of a specific measure, as is "hours spent on rework tasks per month".
2. a) Answers generally OK, but a fair number fail to relate to the given scenario, and just
discuss DbC in general. You should discuss it for the supermarket scenario.
b) Answers are generally good, giving sensible reasons for your choice of language.
c) Answers are fairly good overall, but are not as good as those for a) and b), mainly because
the contracts are not clearly defined and explained, and most answers leave out invariants.

3. No feedback provided - see JTL.

4. No feedback provided - see JTL.

Q1. I was delighted by the performance of students on this question, which required them to
perform calculations involving the semantics of tenses (something they had never seen).
Obviously, I underestimated what the students could do. As a result, there were many high
marks, though a few crashes.

Q2. First part okay. Second part, really disappointing. The students did not really know how
to set out simple resolution proofs. This part of the course needs to be strengthened.

Q3. From memory, I believe that very few students attempted this question. Those that did
seemed to struggle. Such a nice question, and so few good answers.

Q4. Students did relatively well on this—but then it was non-technical. In retrospect, the last
part was hard to mark fairly. I was pleased with the answers to part c), which showed that
many of the students had thought about the significance of the technical work (on semantics)
they did in the course. And I needed cheering up after questions 2 and 3.
Richard Neville:
Over the lecture period three coursework components were undertaken (RNCW1, ARCW2, & ARCW3). The results from the piece of coursework [RN_CW1] set by RN showed the students had a moderately good understanding of a sub-set of concepts utilised to undertake the course work. The exam results enforced this view - with a good spread of results - [RN's] three components of the exam (A1-10, B1, & B2) each had an overall aggregate score of above average (e.g. in the low 60's). The exam results show that 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 good comprehension of the LOs.

The only [may be] across-the-board lack of understanding (i.e. that a number of students appeared to have) in RN's exam components was related to Exceptions - some students were not able to write code snippets utilising the 'throw' keyword - action to be taken on the issue - is to make sure this issue is given greater stress in the next lecture series.

Allan Ramsay:
Richard Neville may have supplementary comments to make on his part of the exam (though his marks seem to have a perfectly sensible spread, so I don't think there's much to be said about them).

To the marks checker: because there were two people marking this exam, the aggregation of marks on each paper is a bit messy. RN marked questions A1-A10 and B1 or B2, AR marked A11-A20 and C1 or C2. To try to keep things readable, particularly in section A, we've put a mark by each question, a total for each page in a circle at the bottom of the page, and mark for 1-10 or 11-20 in a square at the end of the group. AR managed to get into the habit of doing the marks per question in blue and the others in red, which may help. A small number of students did the section A questions in a higgledy-piggledy order, but I think we've dealt with these, and indicated them reasonably clearly. Just to make matters worse, right at the end I realised that I had been marking A17 wrongly--mostly giving 1 more mark than was deserved, but in a very small number of cases giving 1 less. I have gone through the full set making the changes: where a change was necessary it will be seen clearly on A17 itself, at the bottom of the page where A17 was answered, on the running total for 11-20 and on the cover sheet. Where there is no change for this question then the original mark was right (usually because it was 0!).

Comments on the exam: at least one student got each of my questions absolutely right, which is always reassuring. C2 is rather an all-or-nothing question: having seen the answers, I'm satisfied that it wasn't actually too hard or too easy, but it is true that not many people got marks in the middle of the range for it, so it has tended to exaggerate how well or badly people did overall. However, the four people who have clear fail marks did pretty dreadfully on the other questions they attempted as well, so I don't think that this particular question is what did the damage for them. There is an odd pattern where quite a few students got very similar marks on 1-10 and 11-20, which would suggest that these questions were at about the same level; and quite a few got much lower on 11-20 than on 1-10, but very few were the other way round. I can't interpret this pattern very easily, because some who did worse on my short questions actually did better on my long ones, so it can't just be that they stopped coming to my bits of the course. That's about all. No particular topic from my parts of the course seems to have been much worse than any other topic, the overall spread looks OK to me. I would be inclined to draw a line between the ones who got 29 and below, who clearly don't have much of a clue, and the ones who got 36 and above, who do know something.
COMP27020 Web Technology and Practice 2
Weigang Wang (MBS)

Comments:
General: 9 students sat the examination. The average mark (all unconfirmed marks) was 64.51%, with SD being 7.39. The highest mark was ~76%, the lowest ~50%. Although the average is ~4.5% above the normally acceptable range, the numbers and performance are affected by the small group size, which permitted more interaction and discussion.

Detail:
Examiner 1:
The average performance on questions of A4-A7 on key concepts of Web technology is good. The average performance on question B1 on the impact of AJAX technology and its comparison to traditional Web technology is very good. This cohort of students demonstrated better understanding on AJAX technology than that in last year. This may be a result of the first hand experience they gained from their programming coursework on using Google Web Toolkit to develop an AJAX based Web application. The average coursework performance is also good. No students selected B2, which consisted of 4 questions on some additional distributed system concepts.

Examiner 2:
For questions A2-A3: A1 was answered satisfactorily. A2 was answered by several students in relation to simpleType rather than simpleContent. Such confusion was further manifested in some answers to A3, where several students attempted to define an element rather than a type “appliance”. The notion of extension was handled well by some, presumably those who had undertaken the provided set of XMLS exercises in the laboratory sessions or practised XMLS writing in their own time. The XMLS questions thus proved good distinguishers.

In section C, 5 students answered C1 and 4 C2. Answers to C1a included solutions that eschewed the “natural” (and taught) XSLT style of using recursion via apply-templates in favour of iterative approaches using for-each and specific as opposed to general patterns. This tallies with the commonly-observed aversion to recursion in general in the student body and challenges thus teaching staff to find ways of better communicating recursive approaches to problem solving. Answers to C1b displayed a good knowledge of the advantages/disadvantages of SAX/DOM in various situations and showed that students were well aware of XML parsing strategies. Answers to C2a and C2b generally displayed very good knowledge of the mechanics of JSP and commented informatively on how designers and developers can work within a JSP environment, bringing in wider knowledge of Java bean technology and also of the MVC framework.

COMP27310 Software Engineering 2
Chris Harrison

Comments:
The students’ performance is in the same line as the last year and in fact their answers to the exam questions are slightly better than the students of the last year. In other words, that they understand and catch the teaching materials are a little better than the students of the last year.

COMP27330 Interactive Technologies & Design
Tim Morris

Comments:
Tim Morris: No problems were encountered in delivering the course. The assessment was straightforward, no significant problems occurred. The exam mean was 56% ± 13%

Andrew Howes: There were some very good answers but also a number that were extremely poor.
I am pleased that so many of the students acquired the analytic and evidence-based abilities on which the second half of the module focused. However, next year I need to put renewed effort into engaging a greater proportion of the students.