The marks were generally in line with results from previous years and the performance on the questions were usually at the same level for a given student, indicating that the questions were equal in difficulty. The main problems were to do with students running out of time. This was often because they gave a large amount of writing in answer to some of the questions rather than answering the questions more directly and succinctly. It is important to understand that information of relevance to the actual question is what is marked, this is an important skill in examination technique. Most of the marks lost were due to this aspect. At MSc level, question answers should be more than just a repetition of the lecture notes and the questions were designed to probe the students' deeper understanding of the topic. In this respect the exam did test this and the performance of the cohort of students was satisfactory. All but one student were spread in the range 77% to 45%. In the one exception, the student's marks were consistently low on all questions attempted, indicating a lack of understanding of the material of the whole of the course. The reasons why this was so are therefore individual rather than general.

The quality of answers relating to the Unicore software and concepts was generally the highest of all answers, this indicates that the work in the labs and coursework enhanced the students' understanding of the concepts of Grid computing.

1. The most popular question; answered by almost all the candidates. Generally the answers were good. The key points to note were as follows:
   In part ii) the sequential initialisation in Implementation A leads to many remote memory accesses that were not present in Implementation B.
   In part iii) both loops are parallelisable, but both contain a reduction variable that must be addressed. Parallelisation of the outer loop introduces load imbalance issues.

2. This was the second most popular question and was again relatively well answered.
   Graph should be carefully presented with good (superlinear) performance in the middle, but with poorer performance at the bottom end where cache limitations impact performance, and with a flattening out, and bending over, at the top end where communication begins to dominate. Final part should specifically identify those features of the speed-up graph that cannot be explained by Amdahl's Law.

3. a) Variation on bookwork - note that there are p-squared processors in this case.
   b) Need to recognise that the call to ISSQUARE affects the workload and influences the most suitable scheduling technique. Also note that there is a sequential outer loop which makes dynamic schedules less suitable.

4. Least well done question - attempted by only three candidates. Note that many of the marks are given for the correct algebraic derivations of the cyclic reduction algorithms.

Comments:
I consider that the students did quite well in all questions, despite the fact that the paper was considered by both the moderator and the external examiner as challenging. There were a couple of avoidable mistakes, but I am satisfied with the overall performance. There were a very few students (7), and they had a choice to do 2 out of 4 questions. Thus, the sample is too small to notice any common mistakes in this context.
Q1:
(a) A mixture of interpretations. Some discussing the operating system environment and some the mobile wireless environment. Both were accepted. There was a mix of good well focused answers and much weaker answers.

(b) Some people clearly missed the point of this question which was to use but not write about the error correction laboratory for the course. Answer which were about error correction got almost no marks. The question was about the differences between the laboratory solution. In fact about the software engineering and coding differences between writing for a desktop/laptop computer and a typical modern mobile telephone. A wide interpretation of this was accepted and there were some very good answers and some that were very weak.

(c) Reasonably well done. Some people wrote too much and did not focus specifically on the question. Others got the order of bit-error correction and bit-error detection wrong. Also the questions asks ‘how’ not ‘why’ - some people explained why the measures were needed.

(d) Most people did not state that irreparably damaged packets are usually discarded when an ARQ is activated. Most people understood the point about adding a multiple of the generator polynomial, but some missed it. The email part - very few explained that as irreparably damaged packets are not delivered to the transport layer (they are discarded), and that TCP is available for email, erroneous character are rarely seen in emails.

Q2:
This was a 1 point per significant point made essay question. There was a strong suggestion on the question paper to do a plan. Most did not do one. But one student got 2 points extra from the plan and another got 1 mark. The 2 students with proper, fairly detailed plans got high marks. The others who attempted this question got marks below 50% as they did not cover the topic widely or deeply enough and failed to make lots of point scoring points. This question was so broad that it should have been quite easy to pick up marks. But you have to plan and come up with as many different statements as possible taking into account the maximum required. Ideally for 20 marks make a few more than 20 of what you consider to be point scoring statements.

Q3:
All marks for this question were lower than I had hoped for.

a) Most students who attempted this made some valid comments but few mentioned the end-to-end semantics of TCP and the effect this has on sensor networks. Few mentioned timing issues in TCP which, if delivery is delayed can quickly lead to lots of extra traffic being generated. There were some reasonable answers.

b) Some answers did not suggest features of a transport protocol instead describing what TCP and UDP do but without analytical comment. Many suggested using RTP/RTCP which would be overkill for the given system whereas, for example, suggestions to time-stamp and use sequence numbers would have been more appropriate to allow re-ordering, detection of lost data and knowledge of when samples were collected (though this is more of an application rather than a transport issue). There were a mixture of reasonable and weak answers.

c) Answers to this were generally weaker than for 3a & b. Some suggested using Mobile IP which may be helpful but is not a total answer. Few separated the long periods of static (e.g. planted in a field) from the mobile periods. These two disparate situations should have suggested very different e.g. table and on-demand driven solutions. Some students did not notice this was about the network layer.

Q4:
This question was well done and the majority ‘made mincemeat’ of the example, though many did not really explain its significance. Some people missed the significance of the RC4 cipher stream generator and its initialisation.
Q5:
(a) Many good answers, though often a little disorganised. Some superb answers also.

(b) There are many issues here; many well understood, though not all.

© Good old Shannon Hartley Law, how could anybody miss out here? But quite a number did, by not converting from dB to SNR (i.e. not using the magic formula). Regarding the comparison between single carrier and OFDM - well, I learnt something here. In principle the max bit-rate remains the same, but some guys calculated the effect of the cyclic extension and and the effect of not using all 64 carriers in IEEE802.11. Well done these guys.

Its good to see CS students grasping these Phy layer issues, understanding the relationship (and difference) between bit-rate and bandwidth and appreciating the advantages and disadvantages of multi-carrier OFDM which will be a vital aspect of next generation mobile comms.
Q1.1 is a bookwork question, and was mostly answered well (average 2.97/4 marks) - but some students got confused wrt validity and well-formedness through some copy-and-paste errors, questions 1.1d and 1.1e coincide -- this was explained during the exam, and students were told that answering any of these 2 or both was fine. As a consequence, the total mark for Q1 is 29 rather than 30.

Q1.2 was discussed at length in class, and was answered well (average 3.3/4). Some students had difficulties *explaining* what the structure or content of a document are or giving suitable examples.

Q1.3 was also discussed in class. Since Q1.3b is only worth 1 mark, I couldn't subtract marks for "not" explaining what 'competing' means (it was not asked for explicitly). Since 1.3c and d cover some more complex matters, the results are less good than for 1.3a and b.

Q1.4 requires understanding of XML schema and, since students worked with it during class, most students performed well.

Q1.5 requires good understanding of XQuery: all students showed the most basic understanding, but different students failed on different more advanced aspects (especially the default value for the attribute in the internal DTD and the exact meaning of XPath expression like $art/name)

Q1.6 Students did surprisingly badly on this question (1.27/4), despite the fact that uniqueness constraints are a central concept in CS and data management, and the fact that we have discussed this in class. In particular, some didn't even mention xs:key or xs:unique, but instead mentioned other, general Schema mechanisms to restrict the value of attributes...

Q1.7 required some thinking since it was only briefly mentioned in class. This is reflected in the mark (average 1.3/3) and the misconceptions shown by various students (e.g., conflating 'empty schema', in the context of schema containment, with one that does not contain any constraints, i.e. would be universal).

Q2.1 was discussed in class at length. Some people wrote that Schemas can be used for "describing" and "prescribing". While true, this distinction does not articulate distinct *reasons* for having a schema. E.g., prescriptive human documentation and descriptive human documentation are both human documentation. Automated error checking is automated error checking whether the errors are treated as warnings or as fatal.

In general, it helps to be explicit. For example, if you say one purpose is to validate a document, you should explain what validation is or why you might want to do it (note that there are lots of different reasons ranging from data integrity checking to adding default values.

Q2.2 was pure recall. As such, elaborations which either overstated or mistated aspects of the law lost a point, even if the answer included an exact statement of the law.

Q2.3 was discussed heavily in class in the context of namespace patterns in general and specific XML Schema support for various patterns.

People have confused the contained namespace pattern with Clark notation for universal names, i.e., [http://ex.org/]{foo}. The contained namespace pattern refers to a way of using namespaces for extensibility and composition of formats. It is supported in XML Schema by the any element with a specific namespace and processContents=strict (for example).


Q2.4 was looking for the simple idea that a datatype is a set of values, as was stated often in class, with emphasis.

Q2.5 was discusses extensively in class. Many students inverted the relationship (e.g., asserting that nominal type systems equate types based on structure) while others conflated
it with having user defined types.

Q2.6 The key distinction is that with Matching, V is already of type T (if successful) and thus returns true or false whereas in validation, V may be "not" of type T but convertible into T (and we often get a new V as a result.)

Q2.7 There were some really nice, sophisticated answers which discussed the difference between RELAX NG's built in types and type libraries.

Q2.8 Minor syntax issues (missing parens) did not cost points. Even fairly syntactically broken DTDs were accepted. Missing information about attributes cost a point (e.g., #REQUIRED).

Many people mispelt "argument" as "arguement". That took a point. This was not a case of misremembering DTD syntax!

All dtds were used (after syntax repair) to try validate the example.

Some people got confused about the nesting of elements (arguers is a child of argument, not topic. The indentation is a bit misleading there, but topic is clearly an empty element).

Overall, people did very well on this question (average of 5 out of 6).

Q2.10 Several people missed the fact that Schematron is geared to providing human-understandable errors. That is, in fact, it's entire raison d'etre.

Q2.11 The expression has a union type. Similar examples were discussed in class.

**Comments:**

Please Aphrodite Galata for feedback.
Q1.1 Answered well overall

Q1.2 Patchy answers here, with a number of candidates confusing rigidity and identity

Q1.3 Answered well overall.

Q2 A combination of bookwork and application of tableaux reasoning. Overall answered well.

Q3 Answered well overall

Q4 There were reasonable answers to this question, although few answers considered a wider picture or exhibited students thinking about wider issues.

Q5.1 This question was not tackled very well. Few answers actually referred to particular aspects of the OWL API.

Q6.3 There was some confusion about the benefits of tractable profiles: The key is robustness. Tractable profiles have a more predictable performance model.

Q7.3 The key point for the third one is that the ontology itself might contain wrong modelling, thus even if the procedure is fine you might get garbage out. (And, of course, the implementation might be buggy.)

Q8.1 There was a lot of variance on this question. The most obvious benefits discussed in class is that description based development allows for e.g., automation verification of the representation (as well as reducing e.g., the number of "links" that must be manually maintained and globally inspected) and that post-coordination requires such development like features at runtime.

Q8.2 People had a lot of trouble with this (it's known hard). If an answer reasonably overlapped with a real justification, I gave credit.

Q9.1 Distinguished variables range over only "names" where as undistinguished variables range over "arbitrary elements" of the domain.
Q9.2 The key difference is that concept expressions can only detect tree like models where as conjunctive queries can detect graphs.
Q9.3 Some people just had the describe present the same results as an ask (without further justification for this). Generally, the describe should provide information that you couldn't easily get any other way.

The other questions were fine.
I am very satisfied with the paper. The results were generally good to very good, with answers fairly equally spread amongst the optional questions.

5 students failed the paper out of 49 (i.e. got a mark of less than 50%). Most of those who failed got a very low mark, indicating no serious preparation. I did not identify any systematic error (i.e. made by a significant number). The course unit is popular (49 students) and it continues under a new code, but with the same title.

---

**Comments:**

1. Overall, if I remember correctly, this year's exam performance is better than last year. Most students have obtained over 30 (out of 50) and a small number of them have achieved more than 40 (out of 50).

2. Of 84 students, only 14 chose B2 or B3 question; the rest all chose B1, which is the main subject of the course - not surprisingly!

3. Two common mistakes: (1) nearly half of the students were confused with the definitions of patterns in general and patterns in software in particular, resulting in a reduction of marks in question A. (2) a large number of students who did B1 confused Self-Service pattern with Information Aggregation pattern.

4. Lack of understanding: One main reason for most students not to choose B2 is that the question concerned software design patterns and object-oriented techniques. It appears to me that most of the students in my class don't have enough background knowledge in software design so the question was quite challenging to them.

5. I have used negative marking this time as I wish to show the students where they got it wrong and how many marks have been deducted and where appropriate, a comment on why they got it wrong. This way of marking will help the students if they ask for viewing the exam scripts.

6. On reflection, I am satisfied that the students have done well and their results are within my expectation.