PGT Exam Performance Feedback 2017/2018 Semester 1

COMP60411 Modelling data on the Web

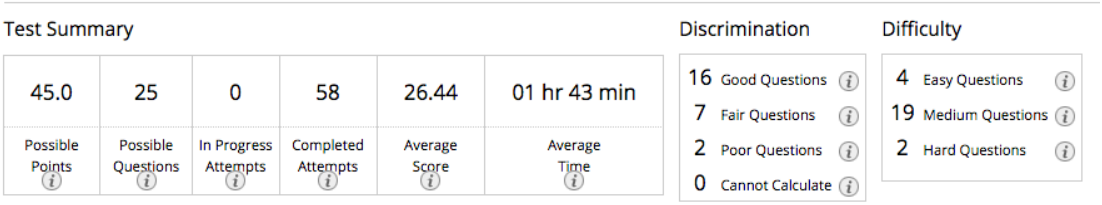
Uli Sattler
Bijan Parsia

Comments Please see the attached.

Exam Performance Feedback COMP60411 2017/2018

The exam consisted of 21 multiple choice questions worth 25 marks, and 4 essay questions worth 20 marks. The average score overall is 26.44/45 or 58%, and it is 14.77/25 or 59% for the MCQs and 11.65/20 or 58% for the essay questions.

We ran an Item Analysis on the whole exam: we had a spread of hard-easy questions, and reviewed the questions with poor discrimination (mostly related to JSON - we suspect this was a subject a little overlooked by some students). The 4 essay questions had the highest discrimination and were of medium difficulty (between 50-65%).



Question 22 was pure bookwork on the core subject of the course unit and mostly answered well, though a surprising number of students overlooked the role of well-formedness or had difficulties relating parsing/serialising with the internal/external representation.

Question 23 required students to remember how a RelaxNG schema roughly looks and works like, and to create one that captures the example. Most students got the (very simple) regular expressions for the element content correct, yet some had difficulties with optional attributes and order or value range.

Question 24 was a more advanced question and required a good understanding of the role different schemas can have. Many students failed to see that the validation strategy was for the whole application (and not just the two constraints) and that the two constraints were on different conceptual levels. In contrast, most students noticed that the second constraint was well suited for Schematron.

Question 25 was, again, a more advanced question and required students to present and discuss Postal's law; many didn't even mention it.

A number of students wrongly claimed that schemas can "repair" data.

A more common confusion was that students thought that the "multiple schema" have to be in *different* schema languages or even that we were talking about schema languages. One can have a "liberal" and a "strict" schema for the same format in XML Schema (alone) or RELAX NG (alone).

Some students gave very abbreviated answers (not much more than "use multiple loose schemas for input and one strict schema for output"). This is the basic idea of the answer (well, at least 2, a loose one for input and a strict one for output), but a proper answer requires more detail both of the system and preferably the format!

Some students wrongly assumed that Postal's law is the same thing as "robustness", but it's a principle for building robust networked systems.

Some students wrongly assumed that RELAX NG *can't* do stuff with datatypes. String and integer are built in and most implementations can handle a large fraction of XML Schema simple datatypes, and we used this in our coursework!