

UG Exam Performance Feedback

Third Year

2017/2018 Semester 2

COMP38120 Documents, Services and Data on the Web

Sandra Sampaio
Jock McNaught
Riza Batista-Navarro

Comments Q1: General comments

Comments are made in respect of unmoderated marks. 61 candidates answered Question 1. The mean mark was 13.05/20, the highest mark 19/20 and the lowest 5/20. Several answers were characterised by overanswering for the marks available; several were characterised by underanswering for the marks available or being too general, not entering into specifics and not addressing the question.

Specific comments

1a) Some answers provided no working, or no explanation for the order obtained. Others failed to deal correctly with set complements.

1b) This was mostly answered well, although several candidates did not enter into many specifics, did not give reasons why or how (indication of surface learning/memorising).

1c) Parts i)–iii) were answered well, the last two parts were not in many cases, with vague answers being given that then left doubt in the mind of the examiner about level of understanding of index construction.

1d) A simple solution was neglected by numerous candidates in favour of a more complex, inefficient one (the latter being given a partial mark). Suggestions to (manually) update the index of a running system to answer a query would be highly impractical.

1e) This gave rise to some excellent answers, showing very good evidence of deep learning, of reading around in the topic and of the ability to synthesise knowledge. Several answers remained very general and/or were over-brief for the marks available, or were experiential without referring to techniques underpinning semantic search, e.g., analysis of documents via named entity recognition or fact extraction, or use of techniques such as clustering.

Q2:

This question was composed of two bookwork items and two application of technique items.

Overall, the students did very well, particularly in the application of technique items, having the majority of the student scoring very high. Less consistency in the quality of answers was noticed for the bookwork items, indicating that a significant number of the students did not take the opportunity to read from the suggested reading material when revising for the exam.

Q3: In general, majority of the students answered the bookwork questions well, for example: benefits of the RDF model in the Linked Data context (Question 3a), advantages of URI aliases (Question 3b), definition of RDFS classes and properties (Question 3d.i).

The most common mistakes made by students arise from the questions requiring application of concepts.

1. Except for one or two students, no one made use of the `rdfs:domain` construct to specify that the duration property only applies to the video class, in Question 3d.iii.

2. In their analysis of the SPARQL query in Question 3e, most students:

a. Did not specify what the return values consist of (i.e., name of the product and the date when the review was created)

b. Did not specify that only reviews pertaining to products containing "phone" in their names will be matched

c. Did not explain that casting the string to a date type is necessary so that the comparison will be based on the date values rather than string values.

3. A few students did not perform merging by using only a unique `sus2` node in drawing the graph corresponding to the given RDF/XML document in Question 3c.
