



Information page for written examinations at Linköping University

Examination date	2014-01-16
Room (1) If the exam is given in different rooms you have to attach an information paper for each room and <u>mark intended place</u>	T1
Time	14-18
Course code	TDDD43
Exam code	TEN1
Course name Exam name	Datamodeller och databaser, avancerad kurs Skriftlig tentamen
Department	IDA
Number of questions in the examination	8
Teacher responsible/contact person during the exam time	Patrick Lambrix / Fang Wei-Kleiner
Contact number during the exam time	2605
Visit to the examination room approx.	15:15; 16:45
Name and contact details to the course administrator (name + phone nr + mail)	Madeleine Häger Dahlqvist, madeleine.hager.dahlqvist@liu.se, 2360
Equipment permitted	
Other important information	
Which type of paper should be used, cross-ruled or lined	
Number of exams in the bag	

Institutionen för datavetenskap
Linköpings universitet

TENTAMEN

TDDD43 Advanced Data Models and Databases

January 16, 2014, 14-18

Grades: For a pass grade you need to obtain 50% of the total points.

Instructions: In addition to the instructions on the cover page:

- Write clearly.
- Start the answers to a question on a new page.
- If you make assumptions that are not given in a question, then clearly describe these assumptions. (Of course, these assumptions cannot change the exercise.)
- Give relevant answers to the questions. Points can be deducted for answers that are not answers to the question.
- Answer in English.

LYCKA TILL!

1. XML querying (3 + 2 = 5p)

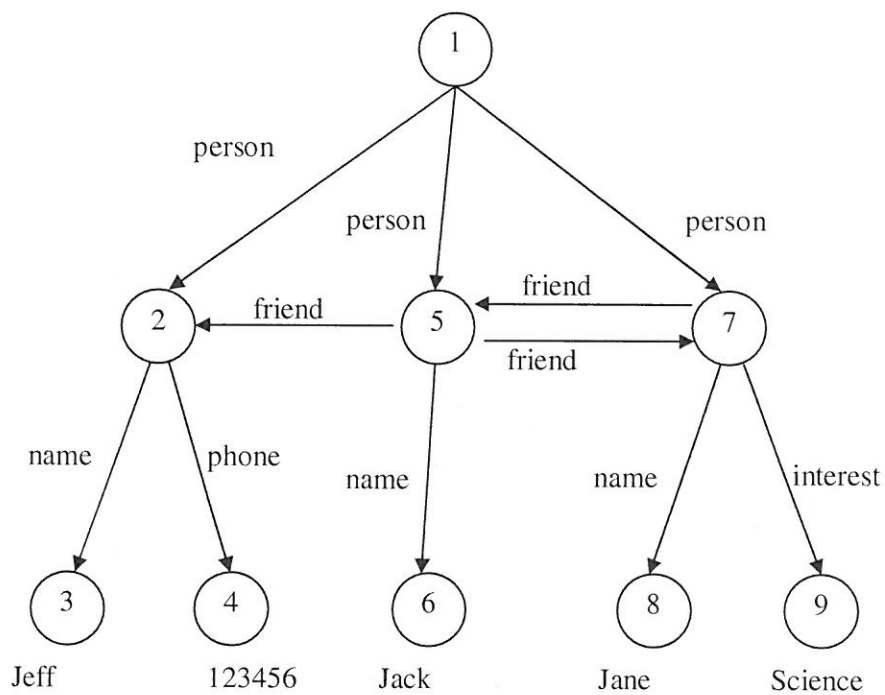
Study the following XML file.

- a) What is the result of executing the following XPath expressions on the XML file?
 - i) `//country[members]/population`
 - ii) `/geodata//name`
 - iii) `//countries[country/@id=2]/country/members`
- b) Express “*Find the population of the country whose capital name is Stockholm*” as an XQuery query.

```
<?xml version="1.0" encoding="UTF-8"?>
<geodata>
  <countries>
    <country id = "1" name="Sweden" capital="c1">
      <population>9,540,065</population>
    </country>
    <country id = "2" name="Denmark" capital="c3">
      <members>
        <member>UN</member>
        <member>EU</member>
      </members>
    </country>
    <country id = "3" name="Germany" capital="c2">
      <members>
        <member>UN</member>
        <member>EU</member>
      </members>
      <population>81,726,000</population>
    </country>
  </countries>
  <cities>
    <city id="c1">
      <name>Stockholm</name>
    </city>
    <city id="c2">
      <name>Berlin</name>
    </city>
    <city id="c3">
      <name>Copenhagen </name>
    </city>
  </cities>
</geodata>
```

2. Data Guides (3p)

Draw a strong data guide for the data model below.



3. NoSQL databases (2p)

Consider the figure on the last page. Let P1, P2 and P3 be three processes and each of them maintain a vector clock with the initial value of (0,0,0).

Fill in the values of the vector clock of each process for the events illustrated with black points. Hand in the last page with your answer.

4. Information Retrieval (2p)

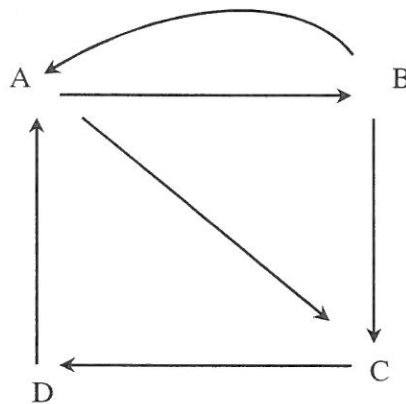
Explain tf and idf in the vector model.

5. Page rank (3p)

There are four Web pages A, B, C, D linked with each other as illustrated in the following graph.

(1) Without considering the teleporting, which page has the highest page rank? Which the lowest? Justify your answer.

(2) By considering teleporting and assuming the constant β as 0.85, how will the page rank of the pages be changed?



6. Description logics (2 + 2 + 1 = 5p)

- (i) Explain the notions of subsumption and satisfiability.
Can subsumption be reduced to (un)satisfiability? Explain your answer.
- (ii) Does *FORALL PlaysAt.Team* subsume *EXISTS PlaysAt.Team*?
Prove your answer using a tableau algorithm.
- (iii) Given the following statement:
Teacher(Patrick, databases)
What is the result of the following queries using open, respectively closed world semantics:
 - a. Teacher(Patrick, databases)?
 - b. Teacher(Patrick, data mining)?

7. Integration (3p)

Describe the local as view approach for integration of data sources. Discuss knowledge representation, query formulation and query processing, and give examples of each.

8. Ontologies (2 + 2 = 4p)

- (i) Give and explain 4 principles of the OBO foundry.
- (ii) Given the concept 'cell death' in ontology O1 and the concept 'apoptosis' in ontology O2. What kinds of matchers would give this pair of concepts a high similarity value and in which cases? Discuss the situations. You may make assumptions about the content, structure etc. of the ontologies.

