Information page for written examinations at Linköping University

Examination date	2011-12-16
Room (1) If the exam is given in different rooms you have to attach an information paper for each room and mark intended place	G36
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
Contact number during the exam time	2605
Visit to the examination room approx.	15.15, 16.30
Name and contact details to the course administrator (name + phone nr + mail)	Madeleine Häger Dahlqvist, 2360, madeleine.hager.dahlqvist@liu.se
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

December 16, 2011, 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 modeling and querying (2 + 1 + 1 = 4p)

Study the following XML schedule (*continues on next page*!):

a) Give a sample XML file that validates for this query.
b) Assume that the company using this schema administrates hundreds of orders from the same customer. State one possible disadvantage with the given schema.
c) Give an XPath query for finding all items with a price higher than 20.

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

<xs:simpleType name="stringtype">
 <xs:restriction base="xs:string"/>
</xs:simpleType>

<xs:simpleType name="inttype">
 <xs:restriction base="xs:positiveInteger"/>
</xs:simpleType>

<xs:simpleType name="dectype">
 <xs:restriction base="xs:decimal"/>
</xs:simpleType>

<xs:simpleType name="orderidtype">
 <xs:restriction base="xs:string">
  <xs:pattern value="[0-9]{6}"/>
 </xs:restriction>
</xs:simpleType>

<xs:complexType name="shiptotype">
 <xs:sequence>
  <xs:element name="name" type="stringtype"/>
  <xs:element name="address" type="stringtype"/>
  <xs:element name="city" type="stringtype"/>
  <xs:element name="country" type="stringtype"/>
 </xs:sequence>
</xs:complexType>

<xs:complexType name="itemtype">
 <xs:sequence>
  <xs:element name="title" type="stringtype"/>
  <xs:element name="note" type="stringtype" minOccurs="0"/>
  <xs:element name="quantity" type="inttype"/>
  <xs:element name="price" type="dectype"/>
 </xs:sequence>
</xs:complexType>
```

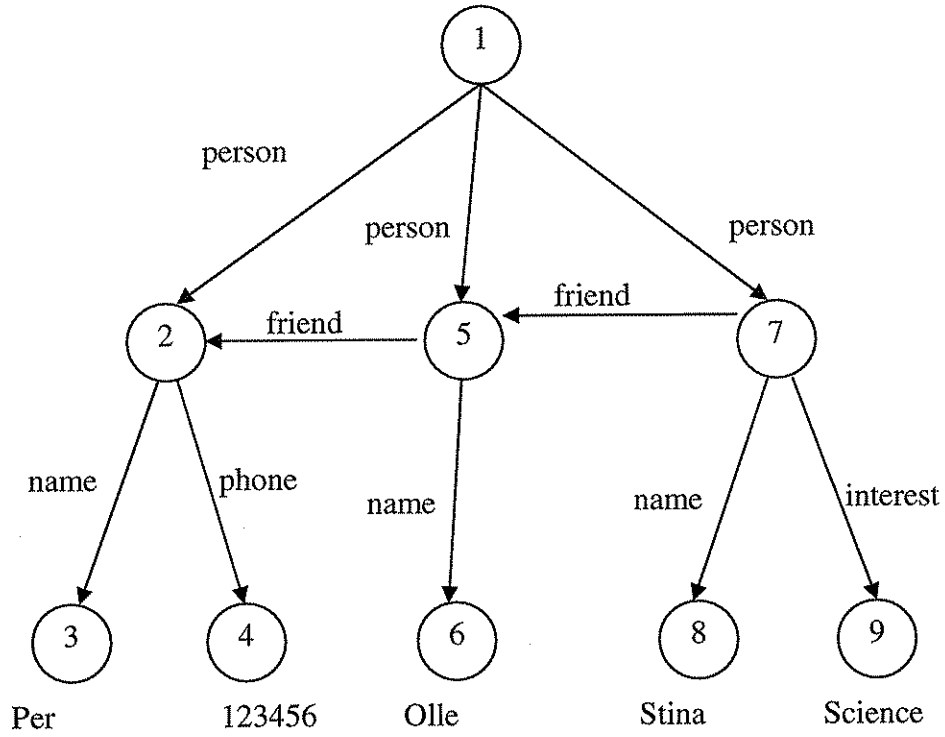
```
<xs:complexType name="shipordertype">
 <xs:sequence>
  <xs:element name="orderperson" type="stringtype"/>
  <xs:element name="shipto" type="shiptotype"/>
  <xs:element name="item" maxOccurs="unbounded" type="itemtype"/>
 </xs:sequence>
 <xs:attribute name="orderid" type="orderidtype" use="required"/>
</xs:complexType>

<xs:element name="shiporder" type="shipordertype"/>

</xs:schema>
```

2. XML storage (1 + 1 + 1 = 3p)

a) What are the main ideas behind Native XML storage?
b) What are the main characteristics of shredding XML into a RDBMS?
c) What is meant by hybrid storage techniques?

3. Data Guides (4p)

Draw a strong data guide for the data model below.



4. 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.

5. OO databases (1 + 3 = 4p)

In a university information system the documents about the employees, students, and courses are stored. Employee documents are either personnel (such as application material and the salaries) or public (such as research and publications). Student documents consist of the student achievements and the transcripts. Course documents consist of the course description, course materials and the subscription lists.

University administrative staff members can read and write all the information of employees and students. Teaching staff members can read and write the course documents and the student transcripts, and in addition, read the public employee documents.
Students can only read the course description and the public employee information.

a) Draw the subject, authorization object and authorization type hierarchies.
b) Draw an authorization model using implicit/explicit, weak/strong and positive/negative authorizations for each of the following: administrative staff members, teaching staff members and students. Assume that when a positive authorization is not stated in the text that the authorization is not given.

6. Ontologies (1 + 2 = 3p)

- Give 2 application areas where ontologies are used. Explain how they are used.
- Give and explain 4 principles of the OBO foundry.

7. Description logics and reasoning (1 + 2 + 1 = 4p)

- Define the semantics of
  - $\forall$ *hasMember.SoccerPlayer*
  - $\exists$ *hasMember.SoccerPlayer*

- Does $\forall$ *hasMember.SoccerPlayer* subsume $\exists$ *hasMember.SoccerPlayer*?
  Prove your answer using a tableau algorithm.

- Given the following statement:
  Teacher(Patrick, databases)
  What is the result of the following queries using open, respectively closed world semantics:
  - Teacher(Patrick,databases)?
  - Teacher(Patrick, data mining)?

8. Integration (4p)

Data source 1 has information about used cars and contains model, year, price, seller information. Data source 2 contains information about luxury cars for which the price is at least 500 000 euro and contains model, year, price, seller information. Data source 3 contains reviews for cars based on the model and the year. Assume the global schema defines the following relations: Car(model, price, year) and Car-review(model, year, review)

- Give local as view mappings for all data sources with respect to the given global schema.
- Describe the query: Return reviews for cars from 2010 with price lower than 10 000 euro.
- Apply then the bucket algorithm to generate the queries to the local data sources.

Hand in this page for question 4.