

Written exam in

TDDD43 Databases and Data Models Advanced course

Location: U4, U6

Date: 2011-04-29

Time: 8-12

Equipment: None

Requirements:

16 points is required for passing the exam

Approx. 22 points is required for grade 4

Approx. 28 points is required for grade 5

The results are announced within two weeks

Teacher in charge: Lena Strömbäck, 0709-39 6776, 013-28 23 24,
Patrick Lambrix, 013-28 26 05

Use one sheet for each question.

Write name and personal number on each sheet.

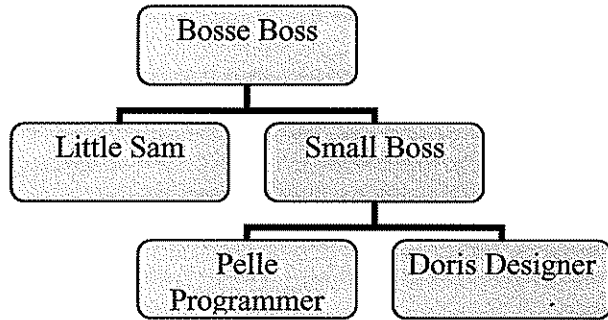
Only write on one side of each sheet.

Give relevant answers to the question.

Points can be deducted for answers that are not answers to the question.

1. XML modeling and querying (8p)

The XML schema at the end of this exam show a model for the personal in a company.



- a) Use this schema to model the personal in the small company described in the diagram to the right. (If you need to add more information make something up.) (2p)
- b) Explain the meaning of key/keyref in the schema. Give an example file that do not validate because of the key/keyref definition. (2p)
- c) State one drawback with the given schema. (1p)
- d) Express “What is Little Sam’s e-mail address” as an XPath query. (1p)
- e) Express “What is the name of Little Sams Boss?” as an XQuery query. (2p)

2. RDF (2p)

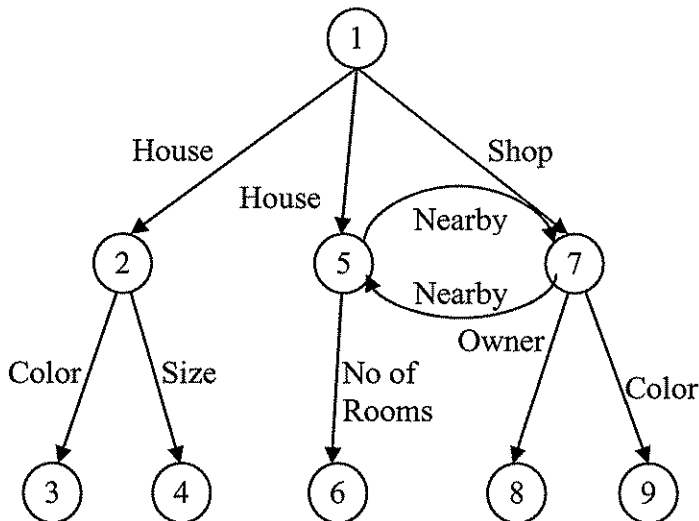
Explain the main characteristics behind the data model used in RDF. Give a simple example of how the manager relation from exercise 1 can be modeled in RDF. You do not need to use the RDF syntax.

3. XML storage (3p)

- a) What is the main idea behind Native XML storage? (1p)
- b) What is the main idea behind shredding XML into a RDBMS? (1p)
- c) What is hybrid XML storage? (1p)

4. Data Guides (3p)

Below a simple sample data model is given. Draw the strong data guide for this model. Use the OEM model.



5. OODBMS modeling (4p)

In a pharmaceutical company different kinds of reports are stored including strategic documents, research documents and public material.

Managers in the company can read all kinds of documents and write strategic documents and public material. Researchers can read and write research documents and read public material. PR employees can read research documents and read and write public material. Other employees can only read public material.

- a) Draw the subject, authorization object and authorization type hierarchies. (2p)
- b) Draw an authorization model using implicit/explicit, weak/strong and positive/negative authorizations for each of the following: managers, researchers, PR employees. Assume that when a positive authorization is not stated in the text that the authorization is not given. (2p)

6. Semantic Web (3p)

Discuss the following problems for the syntactic web and show how the semantic web could alleviate these problems:

- locating relevant information
- retrieving relevant information
- integrating information

7. Databases and Description logics (3p)

Databases use the closed world semantics while description logics use the open world semantics. Describe what open and closed world semantics mean in terms of interpretations, completeness of the information and query evaluation.

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 (3p)

- Describe and explain the different steps in data source integration.
- How are these different steps dealt with in the mediation/view integration approach that uses local as view?

9. Ontology alignment (0.5p + 2.5p = 3p)

- An ontology alignment system contains a number of matchers. What is the functionality of a matcher?
- Give the different kinds of strategies for matchers. For each kind of matcher, give the following:
 - o Which kind of knowledge do the matchers use.
 - o Give an example of such a matcher.

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

  <xs:element name="personnel">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="person" minOccurs="1" maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:complexType>
    <xs:key name="empid">
      <xs:selector xpath="person"/>
      <xs:field xpath="@id"/>
    </xs:key>
    <xs:keyref name="keyref1" refer="empid">
      <xs:selector xpath="person"/>
      <xs:field xpath="link/@manager"/>
    </xs:keyref>
  </xs:element>

  <xs:element name="person">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="name"/>
        <xs:element ref="email" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element ref="link" minOccurs="0" maxOccurs="1"/>
      </xs:sequence>
      <xs:attribute name="id" type="xs:string" use="required">
      </xs:attribute>
      <xs:attribute name="salary" type="xs:integer">
      </xs:attribute>
    </xs:complexType>
  </xs:element>

  <xs:element name="name">
    <xs:complexType>
      <xs:all>
        <xs:element ref="family"/>
        <xs:element ref="given"/>
      </xs:all>
    </xs:complexType>
  </xs:element>
```

Continues on next page.

```
<xs:element name="family" type="xs:string">
  </xs:element>

  <xs:element name="given" type="xs:string">
    </xs:element>

  <xs:element name="email" type="xs:string">
    </xs:element>

  <xs:element name="link">
    <xs:complexType>
      <xs:attribute name="manager" type="xs:IDREF">
        </xs:attribute>
      <xs:attribute name="subordinates" type="xs:IDREFS">
        </xs:attribute>
      </xs:complexType>
    </xs:element>

</xs:schema>
```

