COS326 Database Systems **Semester Test** **Examiners:** Dr. P.E.N. Lutu Dr Serena Coetzee **Date:** 11 October 2012	**Time:** 2 hours **Marks:** 80 **No of Pages:** 9	UNIVERSITEIT VAN PRETORIA UNIVERSITY OF PRETORIA YUNIBESITHI YA PRETORIA Faculty of Engineering, Built Environment and Information Technology ***Department of Computer Science***

Surname:..... Name:.....

Student number:..... Signature:.....

ATTEMPT ALL QUESTIONS

No marks will be awarded for one- word answers.

SECTION ONE: 30 marks

Question 1 (5 marks)

Fill in the blanks:

a. Given n sets of values, D_1, D_2 ,..., D_n , the _____ is the set of all ordered pairs (d_1, d_2,. .., d_n), where each d_i is a member of D_i (that is: $d_i \in D_i$, $i = 1,.., n$)

b. A _____ is a set of tuples. A _____ is a set of allowable values for one or more attributes of a *relation*.

c. The _____ of a relation is the number of attributes it contains while the _____ of a relation is the number of tuples it contains.

d. _____ is a named relation defined by a set of attribute and domain name pairs while a _____ is a set of relation schemas, each with a distinct name.

e. A _____ is a collection of normalized relations with distinct relation names.

Question 2 (5 marks)

State whether each of the following statements is TRUE or FALSE.

a. Relational algebra provides the mathematical theory for relational database queries while the structured query language (SQL) specifies the implementation of relational database queries. _____

b. Relational algebra defines a basic set of operations on relations and specifies how data is obtained from relations. _____

c. Expression 1 and expression 2 are equivalent: *expression 1*: $(R \bowtie_F S)$

expression 2: $\sigma_F (R \times S)$ _____

d. The degree of a theta join between two relations R and S is the product of the degrees of the operand relations R and S. _____

e. The relational algebra expression: $\rho_{empNo, lName}$ **(Employee)** specifies that the attributes *empNo* and *lName* should be selected from the employee relation in order to form a new relation. _____

Question 3 (3 marks)

State whether each of the following statements is TRUE or FALSE.

a. One approach to achieving persistence in an object-oriented programming language is to use the Internet as the underlying storage engine. _____

b. Pointer swizzling refers to the action of converting object identifiers to main memory pointers and back again. _____

c. An object-oriented data model is a persistent and sharable collection of objects. _____

Question 4 (5 marks)

State whether each of the following statements about the object-relational database is TRUE or FALSE.

a. Resolves many known weaknesses of RDBMSes. _____

b. Preserves significant body of knowledge and experience gone into developing relational databases and applications. _____

c. The object-relational database is a very complex data structure compared to the relational database. _____

d. Proponents of the relational approach believe that when the RDBMS is replaced by the ORDBMS, the simplicity and purity of the relational model are lost. _____

e. The object-relational database system has resulted in a much simpler version of **SQL.** _____

Question 5 (5 marks)

State whether each of the following statements is TRUE or FALSE.

a. Object-relational DBMSes are suitable for storing geospatial data because one can define the classes (attributes and methods) that are required to store and manipulate geospatial data. _____

b. A geographic feature is represented by geometry and attributes. _____

c. Geospatial data stored as raster data can be used for network analysis, including shortest path routing. _____

d. Storing spatial data in a spatial DBMS holds the advantage that data can be queried in SQL. _____

e. Three approaches to storing spatial data are: proprietary format, geo-relational model and geodatabase model. _____

Question 6 (5 marks)

State whether each of the following statements is TRUE or FALSE.

a. Semi-structured data is data that conforms to a semi-relational data model. _____
b. Semi-structured data is well-formed XML data that conforms to an XML schema. _____
c. Semi-structured data may be irregular or incomplete. _____
d. An XML schema is used to enforce data validation of an XML document. _____
e. Schematron is an extension of the XML schema. _____

Question 7 (2 marks)

Choose the most correct statement about the order of execution of triggers.

a. If an AFTER trigger exists on the table, the value that is being modified in a table can be changed before it is finally stored in the database.
b. During the execution of a BEFORE-trigger, the new value can be modified by the trigger. This is not possible with an AFTER-trigger.
c. During the execution of a BEFORE-trigger, the new value cannot be modified by the trigger. This is, however, possible with an AFTER-trigger.
d. A BEFORE trigger is fired after the execution of an INSERT or UPDATE statement and when the new value(s) has/have already been stored in the database table.

My choice is: _____

SECTION 2: 50 marks

Question 8. **(10 marks)**

a. Convert the following relational algebra (RA) expressions to SQL statements:

i. $\pi_{city}(Branch) - \pi_{city}(PropertyForRent)$ [2]

_____

_____

_____

_____

ii. $\pi_{city}(Branch) \cap \pi_{city}(PropertyForRent)$ [2]

b. For the following RA natural join expression:

$(\pi_{clientNo, fName, lName}(\mathbf{Client})) \bowtie (\pi_{clientNo, propertyNo, comment}(\mathbf{Viewing}))$

i. What is/are the join attribute(s) for the expression? [1]

ii. Write the SQL statement corresponding to the RA expression. [2]

c. For the following RA expression:

$\pi_{branchNo, street, city}(Branch) \bowtie \pi_{propertyNo}(Property)$

i. What type of join is specified in the expression? [1]

ii. Write the SQL statement corresponding to the RA expression. [2]

Question 9. Object-oriented DBMS **(10 marks)**

a. Following are two of the weaknesses of RDBMSes:

i. Poor representation of "real world" entities

ii. Semantic Overloading due to the fact that a single construct (the relation) is used to represent data and relationships between data.

Explain how OODBMSes (e.g. db4Objects) overcome these weaknesses. [4]

b. What is a persistent programming language and how does it differ from a database programming language? [2]

c. Below is a list of some of the mandatory features on an object-oriented DBMS according to the OODBMS Manifesto of 1989. Complete the table in relation to db4objects. (4)

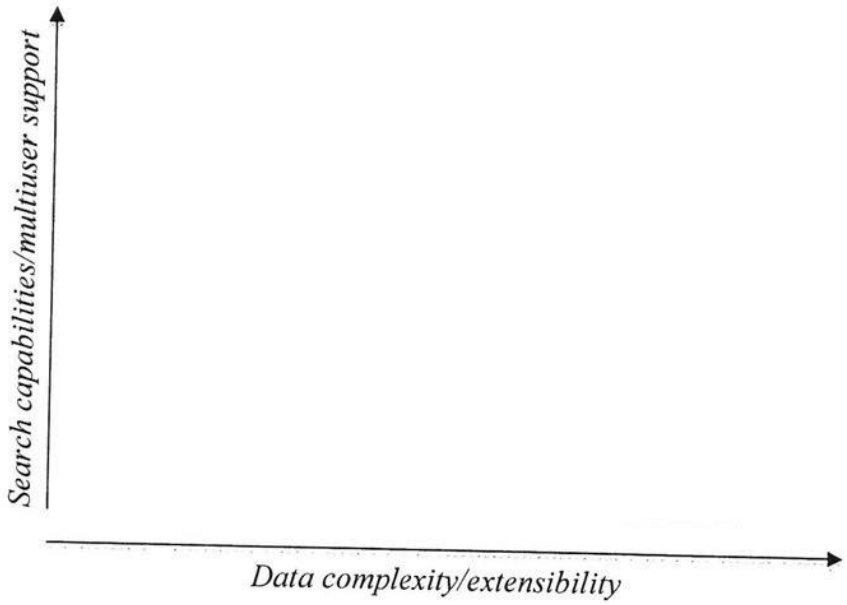
Mandatory feature	Does db4Objects support this feature (yes /no)?	If yes, motivate your answer in one sentence
Complex objects must be supported.		
Object identity must be supported.		
The set of data types must be extensible.		
Data persistence must be supported.		

Question 10. **[10 marks]**

a. Define the term 'impedance mismatch' in the database context. [2]

b. How do object-relational DBMS overcome this mismatch? [2]

c. Complete Stonebraker's four-quandrant view of the database world below. [2]



d. SQL was recently extended to include object-oriented features. Identify (name) the object-oriented feature that is illustrated in each of the SQL statements below. [4]

i. `SELECT * FROM student s where s.isFirstYear;`

ii.
```
CREATE TYPE lecturer UNDER employeeType AS (
    academicDepartment VARCHAR(32),
    publicationCount INTEGER,
    INSTANTIABLE
    NOT FINAL)
```

iii.
```
INSERT INTO student VALUES('2906837', ROW('COS 326', 'PASS WITH
DISTINCTION'));
```

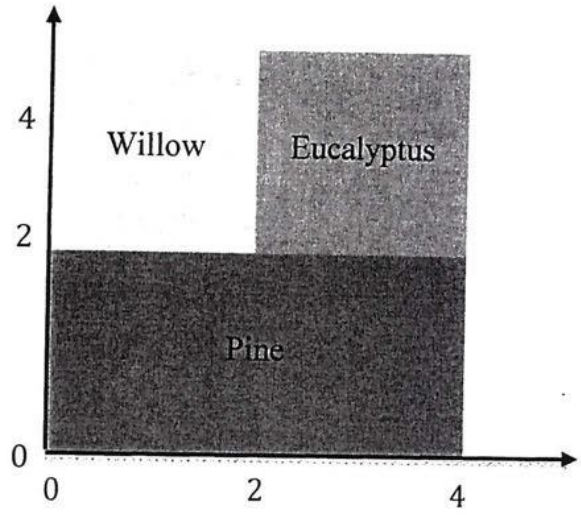
iv.
```
CREATE PROCEDURE printVAT (IN amount DECIMAL)
    DECLARE
    /* variable declarations*/
    BEGIN
    /* print VAT */
    RETURN
    END;
```

Question 11 **(10 marks)**

a. Explain why the following are two important properties of geospatial data:

i. Geospatial data is referenced to a geographic space [2]
ii. Geospatial data can be represented at a variety of scales [2]

b. Use the representation of forest cover below to explain the difference between object-based and field-based geospatial data. [3]



c. Explain why object-relational databases are suitable for storing and maintaining geospatial data. [2]

d. Explain why indexes used for alpnanumeric data in relational databases are not suitable for geospatial data. [1]

Qu **(10 marks)**

a. Discuss the difference between a tree-based and an event-based API for XML. [4]

b. Study the XML data for property viewings in the `viewings.xml` file below, and then answer i.- iii.

i. Identify and discuss a feature of semi-structured data that is illustrated in this data [2]

ii. Write an XQuery expression to find the property (property number) that was viewed for which the viewing ID is 23. [2]

iii. Write a FLWOR expression to list the property numbers of viewings that ended in a positive "WishToRent"? [2]

```
<?xml version="1.0" encoding="UTF-8"?>
<Viewings>
    <Viewing ID="22">
        <Client>Wilma Flintstone</Client>
        <PropertyNo>PA14</PropertyNo>
        <ViewDate>2007-06-20T12:00:00</ViewDate>
        <WishToRent>0</WishToRent>
    </Viewing>
    <Viewing ID="23">
        <Client>Joe Bloggs</Client>
        <PropertyNo>PG21</PropertyNo>
        <ViewDate>9:00 on 22 November 2007</ViewDate>
        <Comment>Not bad at all.</Comment>
        <WishToRent>1</WishToRent>
    </Viewing>
    <Viewing ID="24">
        <Client>Joe Bloggs</Client>
        <PropertyNo>PG36</PropertyNo>
        <ViewDate>2007-11-23T10:30:00</ViewDate>
        <Comment>Kitchen too small.</Comment>
        <WishToRent>0</WishToRent>
    </Viewing>
</Viewings>
```