

أسئلة الامتحان النهائي لمادة : ادارة قواعد بيانات
رمز المادة: CS424 التاريخ: 2020/03/04
الزمن: ساعتان

القسم: الحاسب الالى
مطلبة الفصل: السابع



للفصل الدراسي : خريف 2019

رقم القيد

اسم الطالب

Q1 Choose the right answers: (10 points)

1. What is an XML parser used for?
 - a. to validate an XML-document
 - b. to compile XML
 - c. to provide an interface between an application and an XML file
 - d. to check the XML content for spelling errors
2. Which of the following is not the characteristics of data warehouse ?
 - a. Data warehouse is outside the operational system
 - b. Integrates data from more than one operational system
 - c. It has short span of life
 - d. Systematic organization of data
3. Which of the following is true concerning an ODBMS?
 - a. They have the ability to store complex data types on the Web.
 - b. They are overtaking RDBMS for all applications.
 - c. They are most useful for traditional, two-dimensional database table applications.
 - d. All of the above.
4. Choose which data mining task is the most suitable for the following scenario: Given the records of books that a group of people read, find relationship of the genre pattern
 - a. Classification
 - b. Association rules
 - c. Clustering
 - d. Prediction
5. Which statement is true regarding OQL
 - a. OQL is the object-oriented query standard.
 - b. It uses ODL as its schema definition language.
 - c. Types in OQL are like ODL's.
 - d. All of the above
6. Web server data include
 - a. IP address
 - b. Page reference
 - c. Status
 - d. All of the above



7. An XML file has the following structure. What XPath syntax will return the color of the shirt?

```
<menstore>
  <category>Shirts</category>
  <color>Red</color>
  <size>M</size>
  <price>$25.99</price>
  <clothing>
</menstore>
```

- path='/menstore/clothing/color';
- path='/menstore/ color';
- path='/mensstore/clothing/colors';
- path='/clothing/color';

8. All sites of the database system have identical setup same database system software

- Heterogeneous
- Federated
- Homogeneous
- None of the above

9. In two-tier client/server architecture, the running of application programs and the user interface programs is in control of

- modulation side
- client side
- server side
- host side

10. Which of the following is a disadvantage of replication

- Reduced network traffic
- If the database fails at one site, a copy can be located at another site.
- Each site must have the same storage capacity.
- Each transaction may proceed without coordination across the network.

Q2 What are the limitations of the following? (7 points)

- Three tier architecture?
- Database as a server

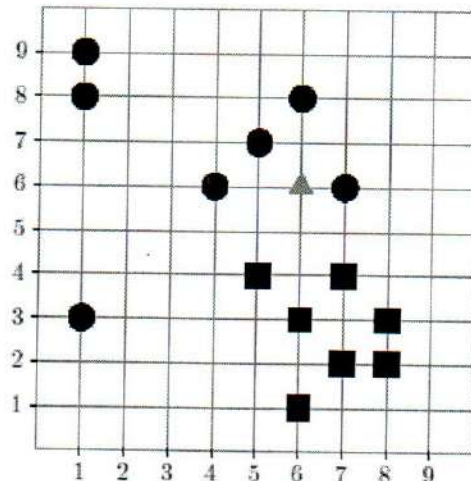
Consider the following ODL schema:

```
interface department
( extent Departments )
{ attribute String      name;
  attribute Short      dno;
  attribute String      address;
};
interface person
( extent Persons )
{ attribute String name;
  attribute Short   age;
};
interface employee: person
( extent Employees )
{ attribute String address;
  attribute set<person> children;
  attribute department dept;
  attribute employee   manager;
};
```

- Give an OQL query which returns names of employees and their addresses and their children names.

Q3 The 2D feature vectors in the figure below belong to two different classes (circles and rectangles). Classify the object at (6, 6) in the image represented using a triangle, using k nearest neighbor classification. Perform KNN classification for the following values of k and compare the results. (7 points)

- K=4





- b. What is web log? Explain web structure mining and web usage mining in detail ?
- c. Based on the below query, identify if the DW uses Star Schema versus Snow-Flake Schema , then identify Dimension, Facts.

```
SELECT
P.Brand,S.Country AS Countries,
SUM(F.Units_Sold)
FROM Fact_Sales F
INNER JOIN Dim_Date D ON (F.Date_Id = D.Id)
INNER JOIN Dim_Store S ON (F.Store_Id = S.Id)
INNER JOIN Dim_Product P ON (F.Product_Id = P.Id)
WHERE
D.YEAR = 2020 AND P.Product_Category ='Computer'
GROUP BY
P.Brand, S.Country
```

Q4 Consider the XML document listings in Figure below (8points)

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<loans>
  <borrower>
    <name>
      <first>Michael</first>
      <last>Hansen</last>
    </name>
    <address>Freisha Way, 23456 Copenhagen, Denmark</address>
    <loan>
      <amount>1000</amount>
      <payout-date>2016-01-01</payout-date>
      <repayment amount="100" date="2017-01-01"/>
      <repayment amount="100" date="2018-01-01"/>
    </loan>
    <loan>
      <amount>500</amount>
      <payout-date>2017-01-01</payout-date>
      <repayment amount="100" date="2015-01-01"/>
    </loan>
  </borrower>
  <borrower>
    <name>
      <first>Mansha</first>
      <last>Daleep</last>
    </name>
    <address>Costoria Street 8232, 98764 Mumbai, India</address>
  </borrower>
</loans>
```

- a. Write an XPath expression that returns all the names of borrowers, who have at least one loan.
- b. Determine the output produced from the above XQuery code when it is run against the following XML document



```
<borrowers>
{
for $x in doc("borrowers.xml")//borrower
return <borrower>
{$x/name}
<debt> {fn:sum($x/loan/amount) - fn:sum($x/loan/repayment/@amount)} </debt>
</borrower>
}
</borrowers>
```

- c. What is the difference between Native XML database and XML-enabled database and which one is better?

Q5 Distributed Database (8points)

- a. Given the following relation EMP and the predicates p1: SAL > 23000, p2: SAL < 23000 . Perform a horizontal fragmentation of the table based on the given predicates ? Is this a correct fragmentation.

ID	Name	Salary
01	Ahmed	12000
02	Ali	67050
03	Mohamed	51980
04	Areej	23000
05	Fatema	28760
06	Mahmood	23000
07	Ashraf	29999

- b. If the answer to (a) is no, explain why, and give the predicates that would correctly fragment the table.
- c. Give the names of each system give the names of each system from where Distribution, Heterogeneity and Autonomy
1. A0 , D0 , H1
 2. A2 , D1 , H*
 3. A1 , D2 , H*