

Tentamen 2004-04-05

DATABASTEKNIK - 1DL116, 1MB025

DatumTisdagen den 5 April, 2005
Tid14:00-19:00
Jourhavande lärare ...Kjell Orsborn, tel. 471 11 54 eller 070 425 06 91
Hjälpmedelminiräknare

Anvisningar:

- Läs igenom hela skrivningen och notera eventuella oklarheter innan du börjar lösa uppgifterna. Förutom anvisningarna på skrivningsomslaget så gäller följande:
 - Skriv tydligt och klart. Lösningar som inte går att läsa kan naturligtvis inte ge några poäng och oklara formuleringar kan dessutom misstolkas.
 - Antaganden utöver de som står i uppgiften måste anges. Gjorda antaganden får förstås inte förändra den givna uppgiften.
 - Skriv endast på en sida av papperet och använd ett nytt papper för varje uppgift för att underlätta rättning och minska risken för missförstånd.
- För godkänt krävs det cirka 50% av maxpoäng.

1. **Database terminology:**

2 pts

Explain the following concepts in a database context.

- (a) candidate key
- (b) secondary index

Answer:

b) Sekundärindex är en ordnad fil av dataposter med 2 fält där första fältet är av samma typ som som indexeringsfältet, dvs vilket fält som helst i datafilen. Andra fältet är en blockpekare. Indexeringsfältet kan vara ett icke-nyckelfält eller ett sekundärnyckelfält och datafilen ej sorterad efter indexeringsfältet. Sekundärindex kan vara glesa eller täta. Index ger en avsevärd effektivisering vid sökning av dataposter. Vid uppdatering av datafilen måste också tillhörande index uppdateras vilket medför en viss ökad kostnad för dessa operationer.

2. **Data models:**

4 pts

How are the concepts entity type and attribute in the ER (entity-relationship) model represented in the following implementation data models:

- (a) the relational data model

Answer: E-R modellens begrepp entitetstyp representeras som en tabell (relation) i relationsmodellen och attribut representeras som en namngiven kolumn (attribut) i en tabell (relation).

- (b) the object-oriented data model

Answer: object types/classes and object attribute

3. **SQL:**

2 pts

Suppose that we in a database have two relations (tables) with the following schemas:

```
PRODUCT(PID,PNAME)
COMPONENT(CID,CNAME,WEIGHT,COST,PID)
```

, where xID's denote keys.

Formulate an SQL query that retrieves the product id, name and the number of components for each product (i.e. how many components that are part of each product).

```
SELECT P.PID,P.PNAME, COUNT(*) AS NO_OF_COMPONENTS
FROM PRODUCT P, COMPONENT C
WHERE P.PID = C.PID
GROUP BY PID,PNAME
```

4. **Transactions:**

4 pts

Describe the properties that one would like transactions to fulfill in a database context (hint: ACID).

5. Object databases:

4 pts

- (a) Which three user-defined extensibility mechanisms of the database exists in an object-relational database management system? (3 pts)
- (b) Which of the above extensibility mechanisms is/are missing or is/are weak in simple object-oriented database management systems (i.e. in so called “object stores”)? (1 pt)

Answer:

- a) Användardefinierade datatyper, Användardefinierade index, Användardefinierad frågeoptimering
- b) Det saknas stöd för användardefinierad frågeoptimering. Det saknas stöd för användardefinierade index

6. Query processing:

4 pts

- (a) What is an unclustered index? (1 pt)
- (b) Which are the three most common join methods? (3 pts)

Answer:

- a) Ett oklustrat index är ett index vars nycklar har annan sorteringsordning än raderna i tabellen.
- b) Nested loop join, sort-merge join, hash join

7. XML-databases:

4 pts

- (a) What is the difference between DTD and XMLSchema? (2 pts)
- (b) Which are the two most important query languages for XML? (2 pts)

Answer:

- a) Både DTD och XMLSchema beskriver innehåll i XML-filer. XMLSchema är betydligt mer semantiskt rikt genom att man har understöd för många olika datatyper, arv och relationer mellan entiteter. DTDn har bara datatypen sträng och är mer likt en grammatik där man beskriver layout av tillåtna dokument. DTDn uttrycks i annat språk än XML, medan XMLSchema uttrycks i XML.
- b) XQuery och Xpath

8. Multimedia databases:

4 pts

- (a) Which advantages have “object-relational” databases to store multi-media data? (2 pts)
- (b) What is a BLOB and what is it used for? (1 pt)
- (c) Why is RAID good to store multi-media data in databases? (1 pt)
 - a) Man kan lagra komplexa objekt mot vars struktur man kan ställa effektiva optimerade databasfrågor. Man kan indexera innehållet i objekten. Komplexa objekt kan t.ex. vara filmer som kan vara representerade som objekt bestående av scener bestående av bilder innehållande aktörer.
 - b) BLOB = Binary Large Object är en datatyp för att lagra stora bitsträngar i relationsdatabaser utan att veta något om själva innehållet i

BLOBen (t.ex. foton). Man kan således inte ställa frågor som filtrerar baserat på innehållet i en BLOB.

c) RAID = Redundant Array of Inexpensive Discs tillåter att filer ligger lagrade i bitar spridda över flera diskar så att man snabbt kan läsa dem parallellt. Överföringstiden av t.ex. bilder blir då snabbare än om allt ligger på en disk.

9. Database APIs:

4 pts

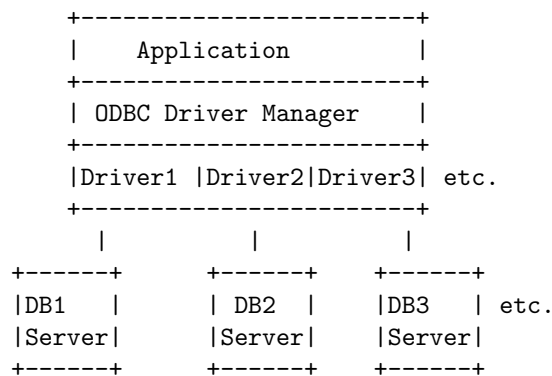
- (a) What is JDBC? (1 pt)
- (b) Explain the difference between JDBC and ODBC. (1 pt)
- (c) Describe the architecture of ODBC. Illustrate with picture. (2 pts)

Answer:

a) JDBC = Java DataBase Connectivity är ett standardgränssnitt mellan programmeringsspråket Java och en eller flera samtidigt tillgängliga SQL-baserade relationella databaser.

b) ODBC = Open DataBase Connectivity är ett programmeringsspråksberoende gränssnitt till SQL-baserade relationella databaser.

c)



DB1, DB2, DB3 different kinds of DBMSs, e.g. DB2, Oracle, MySQL, Mimer

Good luck!

/ Kjell och Tore