

## Tentamen 2003-05-20 DATABASTEKNIK - 1DL116

Datum .....Tisdagen den 20 Maj, 2003  
Tid .....12:00-17:00  
Jourhavande lärare ...Kjell Orsborn, tel. 471 11 54 eller 070 425 06 91  
Hjälpmedel .....Inga

### 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. DB terminologi: 4p

Förklara kortfattade följande begrepp i databassammanhang (det går bra med definitioner där sådana finns).

- (a) referensnyckel (eng. foreign key)
- (b) transaktion
- (c) dödlig låsning (eng. dead lock)
- (d) BCNF - Boyce-Codd Normal Form

2. Databassystem - konventionell filprocessering: 4p

Nämna och motivera fyra fördelaktiga egenskaper hos ett databassystem som normalt skiljer sig från egenskaperna hos ett programsystem som baseras på konventionell filprocessering.

3. Datamodeller: 4p

- (a) Hur representerar man begreppen entitetstyp (eng. entity type) och attribut (eng. attribute) från en E-R modell i relationsmodell? (1p)
- (b) Hur kan man representera binära relationstyper i relationsmodellen (notera att det finns flera binära relationstyper)? (3p)

4. Fysisk databasdesign - indexering: 4p

Förklara för vilka ändamål (typer av databasfrågor) som följande index kan, och inte kan, effektivisera exekveringen:

- (a) hashindex
- (b)  $B^+$ -träd

5. Aktiva databaser: 4p

- (a) Vad består komponenterna i ECA-modellen av i en relationsdatabas? (3 p)
- (b) Varför är det ofta svårt att implementera komplexa integritetsvillkor m.h.a. triggers? (1 p)

6. Objekt-orienterade databaser: 4p

- (a) Vilka tre sorters användardefinierade databasutvidgningsmekanismer finns i ett objekt-relacionellt databashanteringssystem?

- (b) Vilka av ovanstående utvidgningsmekanismer saknas eller är svaga i enkla objekt-orienterade databashanterare (dvs. i s.k. object stores)?

7. Databasapplikationsgränssnitt: 4p

- (a) Vad är JDBC? (1 p)  
(b) Beskriv JDBC:s arkitektur med text och bild. (2 p)  
(c) Vad är skillnaden mellan JDBC och ODBC? (1 p)

8. Frågeoptimering: 4p

En stor firma har en relationsdatabastabell över hur bra deras försäljare är:

`SALES(PNR, SALES, NAME, ...)`

Tabellen innehåller försäljning (SALES) i kr för varje försäljare med nyckel personnummer (PNR). Det finns klustrat primärindex på PNR och oklustrat sekundärindex (B-träd) på SALES. Det finns 10000 försäljare i tabellen. Det får plats 10 tabellrader och 100 indexnoder i varje diskblock.

Företagsledningen behöver ofta veta de 10 bästa försäljarna och du ombeds att designa ett applikationsprogram som snabbt finner dessa stjärnor. För sådana frågor tillhandahåller SQL nuförtiden en utvidgad syntax:

```
select ssn, sales
stop after 10 rows
from sales
order by sales descending
```

Klausulen 'stop after 10 rows' betyder att bara de 10 första raderna i frågan kommer att returneras.

Vilken är den optimala exekveringsplanen uttryckt som utvidgad relationsalgebra innehållande stopafter(x,n) operator? Visa varför den är optimal. Vilken betydelse har 'stop after 10 rows' klausulen för effektiviteten av exekveringsplanen?

Lycka till!

/ Kjell och Tore