



Försättsblad till skriftlig tentamen vid Linköpings Universitet

Datum för tentamen	2014-08-12
Sal (1) Om tentan går i flera salar ska du bifoga ett försättsblad till varje sal och <u>ringa in</u> vilken sal som avses	TER3
Tid	8-12
Kurskod	732G16
Provkod	TEN1
Kursnamn/benämning Provnamn/benämning	Databaser: design och programmering Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	9
Jour/Kursansvarig Ange vem som besöker salen	Erik Prytz
Telefon under skrivtiden	070-2762748
Besöker salen ca kl.	ca kl. 09
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Anna Grabska Eklund, ankn. 2362, anna.grabska.eklund@liu.se
Tillåtna hjälpmedel	inga
Övrigt	
Vilken typ av papper ska användas, rutigt eller linjerat	Valfritt
Antal exemplar i påsen	

1 (8)

Linköpings Universitet
Institutionen för Datavetenskap
Erik Prytz

TENTAMEN

732G16

Datum: **2014-08-12**
Jourhavande lärare: **Erik Prytz**
Telefon: **070-2762748**
Besöker tentamenslokalen c:a kl. 09:00.

Poängfördelning

Fråga	Poäng
1	4
2	1
3	3
4	2
5	4
6	4
7	2
8	3
9	4

Totalt antal poäng: 27
Poäng för G: 13
Poäng för VG: 20

Instruktioner: Skriv läsligt. Visa ditt arbete när så krävs av uppgiften. Läs igenom tentan innan du börjar, så att du kan fråga jourhavande lärare om det finns oklarheter. Använd ett papper per fråga.

Lycka till!

1. (4p) **Begrepp:**

- a. **Flervalsfråga, välj endast ett alternativ:** ER-modellen är ett exempel på vilken typ av datamodell?
 A: Hierarkisk modell
 B: Konceptuell modell
 C: Implementationsmodell
 D: Fysisk datamodell
- b. **Flervalsfråga, välj endast ett alternativ:** Under kursen har vi diskuterat relationsmodellen. I detta sammanhang, vad är en relation?
 A: En koppling mellan olika tabeller i databasen
 B: En beskrivning av de tabeller som kan finnas i databasen
 C: En tabell bestående av en mängd tupler i databasen
 D: En ordnad lista av attributvärden i databasen
- c. **Begrepp:** Enligt den så kallade tre-schema-arkitekturen kan man betrakta sin databas på tre olika nivåer. Namnge och förklara kort vad minst två av dessa nivåer innebär.

2. (1p) **Relation algebra:** Skriv följande SQL-query som ett relationsalgebraiskt uttryck:

```
SELECT PRODUCT FROM INVENTORY WHERE COST<100 OR
COLOR=BLACK;
```

3. (3p) **Fysisk design:** En mängd data i en databas ska sparas i en fil. Filen består av 40 000 poster och varje post är 50 byte stor. Denna fil ska nu sparas på en hårddisk med en blockstorlek på 1024 byte.

- a. Beräkna blockningsfaktor och sedan hur många block som krävs för att lagra filen. Visa ditt arbete.
- b. Till denna fil ska ett glest index skapas för primärnyckeln. Primärnyckeln tar 8 byte och adressen till blocken tar 12 byte. Beräkna indexfilens poststorlek, blockfaktor, samt hur många block som krävs för att lagra indexfilen. Visa ditt arbete.
- c. En annan databasadministratör ifrågasätter om ett index verkligen är bästa lösningen, och föreslår att filen ska sparas med en hashstruktur istället. Det skulle då krävas maximalt 2 spillblock, enligt administratören. Beräkna och visa på skillnaderna i tidsåtgång för att söka reda på en viss post med filen enligt föreslagen hashstruktur jämfört med filen med index. Antag att en blockåtkomst tar 5 millisekunder (0.005 sekunder). Visa ditt arbete.

Du kan använda följande uträkningar till hjälp:

$\log_2(2000) \approx 11$	$\log_2(40) \approx 6$	$\log_2(32) = 5$
$\log_2(1000) \approx 10$	$\log_2(40000) \approx 16$	$\log_2(16) = 4$
$\log_2(256) = 8$	$\log_2(120) \approx 7$	$\log_2(64) = 6$

4. (2p) **Transaktioner:** För att förhindra problem med parallella körningar av databasen kan man använda sig av olika låsfunktioner. Vad innebär binära lås, och hur skiljer det sig från läs- och skrivlås? Till detta kan man även använda tvåfaslåsning, vad innebär det?
5. (4p) **Design/ER:** Biblioteket Bookbuster vill ha en ny databas för att hantera sin verksamhet. De har anlitat dig för att ta fram ett designförslag på hur den nya databasen ska se ut. I detta skede vill de alltså ha ett ER-diagram som visar en lämplig lösning utifrån deras kravspecifikation.

-Det första Bookbusters vill kunna lagra är förstås de böcker som de har i lager. De vill spara titeln, antalet sidor, hur många kopior de har, isbn, publiceringsår, samt värdet på boken. De vill även lagra författare med namn och författar-id (ett unikt nummer för alla publicerade författare). Eftersom samma författare kan ha skrivit flera olika böcker och vissa böcker har mer än en författare vill de kunna spara den informationen separat från böckernas information. Alla böcker har förstås minst en författare.

-Det är även viktigt att spara information om förlagen som publicerar böcker. Man måste kunna hitta namn samt förlags-id (ett unikt nummer för varje förlag). Alla böcker är alltid publicerade av ett enda förlag. Förlagen har även ett antal lokalkontor i olika länder, så Bookbusters vill även ha information om telefonnummer och namn på kontaktpersonen för varje lokalkontor, samt vilket land lokalkontoret ligger i. Förlagen har aldrig mer än ett kontor per land, men kan ha kontor i väldigt många länder.

-Till sist vill Bookbusters även spara data om låntagarna. Det är viktigt att spara namn, adress, samt bibliotekskortsnummer för alla låntagare. En låntagare kan låna hur många böcker som helst, och alla lån har ett återlämningsdatum. Det är viktigt att spara information om vilka böcker som en låntagare har lånat samt när de böckerna ska lämnas tillbaka. En låntagare kan också reservera böcker. Enligt Bookbusters får en låntagare bara reservera en bok åt gången, och om en bok redan är reserverad får ingen annan reservera den boken.

Rita ett ER-diagram baserat på denna kravspecifikation. Markera nycklar, deltagande, samt kardinalitet där det behövs. Du behöver **inte** översätta ER-diagrammet till relationstabeller. Om information saknas eller kravspecifikationen är för vag i något avseende (ex. rörande unika värden för nycklar, deltagande, eller kardinalitet) när du bestämmer din design måste du specificera de antaganden du gör.

5 (8)

7. (2p) **Normalisering:** En klantig databasdesigner har implementerat en dåligt designad relationsdatabas åt polisen som ska spara information om vapenlicenser och vapenägan. I dagsläget ser databasen ut så här:

VAPENÄGANDE(personnummer, licensnummer, personnamn, märke, ändamål, kaliber, vapentyp, inköpsdatum, hagelvapentillstånd, kulvapentillstånd, adress, tillverkningsår, jägarexamen, vapennummer, modell)

Du vet att följande fullt funktionella beroenden (FFB) finns:

{personnummer} → {namn, adress, jägarexamen}
{jägarexamen} → {hagelvapentillstånd, kulvapentillstånd}
{vapennummer} → {vapentyp, kaliber, märke, tillverkningsår, modell}
{licensnummer} → {vapentyp, kaliber, märke, vapennummer, personnummer, ändamål, tillverkningsår}
{personnummer, vapennummer} → {inköpsdatum}
{modell} → {märke}

Normalisera tabellen VAPENÄGANDE så att den uppfyller Boyce-Codd Normalform (BCNF). Ange eventuella antaganden du gör.

6 (8)

8. (3p) **Integritet:** Man brukar tala om två typer av villkor som man måste ta hänsyn till för databasens integritet. Vad kallas dessa två villkorstyper, och vad innebär de? Man kan även ha olika inställningar för att hantera situationer då en ändring i databasen medför att ett villkor inte längre är uppfyllt. En sådan inställning kan vara "on delete cascade". Vad innebär detta?
9. (4p) **Design+SQL:** En databas över filmer ser ut som följer:

Skådespelare(SkådisID, Förnamn, Efternamn, Födelsedatum)
Regissör(RegID, Förnamn, Efternamn, Födelsedatum)
SkådisIFilm(Skådis, Film, Lön)
RegiIFilm(Regi, Film, Lön)
Film(FilmID, Titel, År, Längd, Studio, Inspelningskostnad, Inkomst)
Studio(Namn, Land, TotalOmsättning)

Primärnycklarna är understrukna, och det finns även följande främmande nycklar:

SkådisIFilm(Skådis) *references* Skådespelare(SkådisID)
SkådisIFilm(Film) *references* Film(FilmID)
RegiIFilm(Regi) *references* Regissör(RegID)
RegiIFilm (Film) *references* Film(FilmID)
Film(Studio) *references* Studio(Namn)

Skriv ned de SQL-frågor som behövs för att besvara de tre nedanstående frågorna. Om du tycker det saknas information så att du behöver göra antaganden ange då samtliga sådana antaganden du gör.

- Vilka filmer har skådespelaren Peter Småblåsig, som har SkådisID 1001, deltagit i? Svaret ska innehålla endast filmens titel och år, tack.
- Ta fram den genomsnittliga vinsten (inkomst minus inspelningskostnad) för de filmer som har en regissör vars lön översteg 1 miljon.
- Lista alla år då en film har producerats där regissören har haft samma efternamn som (minst) en skådespelare i filmen.