# Försättsblad till skriftlig tentamen vid Linköpings universitet

(fylls i av ansvarig)

| | |
|---|---|
| **Datum för tentamen** | Januari 10, 2013 |
| **Sal** | TER2 |
| **Tid** | 14.00-18.00 |
| **Kurskod** | TDDD12 |
| **Provkod** | TEN1 |
| **Kursnamn/benämning** | Databasteknik |
| **Institution** | *IDA* |
| **Antal uppgifter som ingår i tentamen** | 7 |
| **Antal sidor på tentamen (inkl. försättsbladet)** | 5 |
| **Jour/Kursansvarig** | Jose M. Peña (jose.m.pena@liu.se) |
| **Telefon under skrivtid** | 0708229596 |
| **Besöker salen ca kl.** | 16.30 |
| **Kursadministratör (namn + tfnnr + mailadress)** | Madeleine Häger Dahlqvist, 282360, madeleine.hager.dahlqvist@liu.se |
| **Tillåtna hjälpmedel** | ordbok |
| **Övrigt (exempel när resultat kan ses på webben, betygsgränser, visning, övriga salar tentan går i m.m.)** | |
| **Vilken typ av papper ska användas, rutigt eller linjerat** | |
| **Antal exemplar i påsen** | |

Institutionen för datavetenskap
Linköpings universitet

# EXAM
# TDDD12 Databasteknik
# TDDD46 Databasteknik

# January 10, 2013, 14.00-18.00

## Help

Dictionary.

## Grades

You can get max 30 points. To pass the exam, grade 3, you need 7.5 points in both the practical and theoretical part of the exam. For grade 4 and 5, you need 21 and 27 points, respectively.

## Questions

Jose M. Peña will visit the room at 16.30.

## Instructions

Write clearly. Give relevant and motivated answers only to the questions asked. State the assumptions you make besides those in the questions. None of these additional assumptions should change the spirit of the exercises. Please, answer in English.

Good luck!

**Practical part (15 points)**

**Question 1. Data modeling with EER diagram (5 p):**

We want to create a database to store information about a group of people and the presents they gave and received in their last party. Specifically, before the party, every person listed the presents they wanted to received as well as the presents they did not want to receive. We want to store these lists in our database. During the party, every person received some presents from some other people. We want to store who gave the present and whom received it (you can assume that the exchange of presents is between single persons, i.e. no between groups). If a person received a present she did not like, she could give it to another person. We also want to store this information in our database.

Draw an EER diagram for the description above. Clearly write down your choices and assumptions in case you find that something in the information above is not clear.

**Question 2. SQL (1 + 2 + 2 = 5 p):**

Consider the following database schema

Country(Name, Code, Capital, Area, Population)
Organization(Name, Abbreviation, Established)
IsMember(Organization, Country, Joined)

The attribute *Organization* in the table IsMember is a foreign key reference to *Abbreviation* in the table Organization.
The attribute *Country* in table IsMember is a foreign key reference to *Code* in the table Country.

Examples of the tuples from the above relational schema are as follows:

Country(Sweden, SWE, Stockholm, 449964, 9514000)
Organization(European Union, EU, 1952)
IsMember(EU, SWE, 1995-01-01)

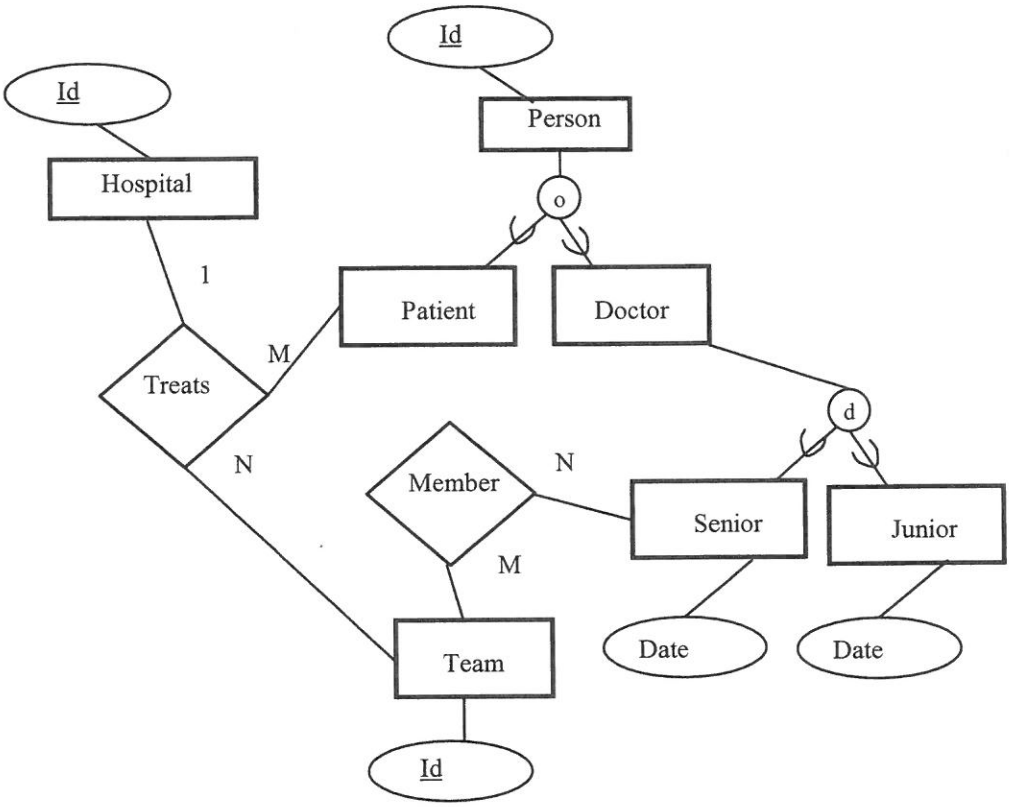
1. List the country names of all the European Union (Abbreviation: 'EU') members.

2. Compute the sum of the areas from all the 'EU' countries.

3. List the name of every country that is the member of at least five organizations.

**Question 3. EER diagram and relational schema (5 p):**

Translate the EER diagram below to a relational schema (use the algorithm seen in the course).



**Theoretical part (15 points)**

**Question 4. Normalization (3 p):**

Normalize (1NF→2NF→3NF→BCNF) the relation R(A, B, C, D) with functional dependencies {A→C, B→D, C→A, D→B}. Explain your solution step by step. Bear in mind that a relation can have several candidate keys.

**Question 5. Data structures (2 + 3 = 5 p):**

We have a file with 1000000 records. Each record is 40 bytes long. The records have two key attributes X and Y. The file is ordered on X. The database uses a block size of B=40000 bytes and unspanning allocation. Each index record is 4 bytes long.

1. Calculate the average number of block access needed to find a record with a given value for X when using (a) the primary access method and (b) a single level index.
2. Calculate the average number of block access needed to find a record with a given value for Y when using (a) the primary access method, (b) a single level index and (c) static multi-level index.

Recall that $\log_2 2^x = x$. That is, $\log_2 1 = 0$, $\log_2 2 = 1$, $\log_2 4 = 2$, $\log_2 8 = 3$, $\log_2 16 = 4$, $\log_2 32 = 5$, $\log_2 64 = 6$, $\log_2 128 = 7$, $\log_2 256 = 8$, $\log_2 512 = 9$, $\log_2 1024 = 10$, $\log_2 2048 = 11$, $\log_2 4096 = 12$, $\log_2 8192 = 13$, $\log_2 16384 = 14$ etc.

**Question 6. Transactions and concurrency control (1 + 1 + 1 = 3 p):**

3. Give a schedule that is not serial but that is serializable. Justify your answer.

4. Apply the two-phase locking protocol to all the transactions in the schedule you gave in exercise 1.

5. Show how the transactions in the exercise 2 (i.e. after you applied the two-phase locking protocol) interleave when executed.

**Question 7. Database recovery (2 + 1 + 1 = 4 p):**

1. Apply the deferred update recovery method and the immediate update recovery method (version 1, i.e. NO-REDO/UNDO) seen in the course to the system log below. Show all operations (in the correct order) that are performed during the recovery.

   Part of system log:
   Start-transaction T2
   Write-item T2, B, 3, 4
   Start-transaction T3
   Write-item T3, A, 7, 8
   Checkpoint
   Write-item T3, A, 8, 1
   Commit T2
   Checkpoint

Write-item T3, A, 1, 5
Start-transaction T4
Write-item T4, B, 4, 5
Write-item T4, B, 5, 10
Commit T3
Checkpoint
Start-transaction T1
Write-item T1, C, 8, 9
Commit T4
→system crash

2. Why do we not store read operations in the log file ?
3. When do we write the log file to disk in the deferred update and immediate update recovery methods ?