

Final Exam I - TDBD15 – Advanced Data
Models and Systems
Friday May 28th, 2004
9.15 - 17.00
MA236 - MA246

NUMBER: _____

(Write this number on every sheet in the exam body.
-10pts if your name appears in the exam body)

NAME: _____

PERSON NUMBER: _____

EMAIL: _____@cs.umu.se

**PROVIDE ANSWERS ON THE GIVEN SHEETS.
NO EXTRA PAGES ALLOWED!**

1. Short Answer

- a. (30) Decompose the following relation into 4NF:

People(Name, Book, CD, Address)

Where Name \rightarrow Book, Name \rightarrow CD, Name \rightarrow Address.

Is your decomposition in 5NF? _____

Why/why not?

- b. (30) Give a ROLAP snowflake schema for the data-cube:

Sales(Store, City, Item, Week)

where each cube cell contains the number of items sold, where cities (e.g. Stockholm) are in countries (e.g. Sweden) which are in regions (e.g. Scandinavia) which are in continents (e.g. Europe), items are item names (e.g. Hamburger) grouped by type (e.g. Meat) grouped by category (e.g. Food) and weeks are grouped into months.

- c. (30) Given 25 rectangular objects on a Euclidean plane, what is the height of the worst case 2-3 R-tree? _____.

In the worst case 2-3 tree, what is the worst case number of nodes touched in a “where am I?” query? _____.

2. EER

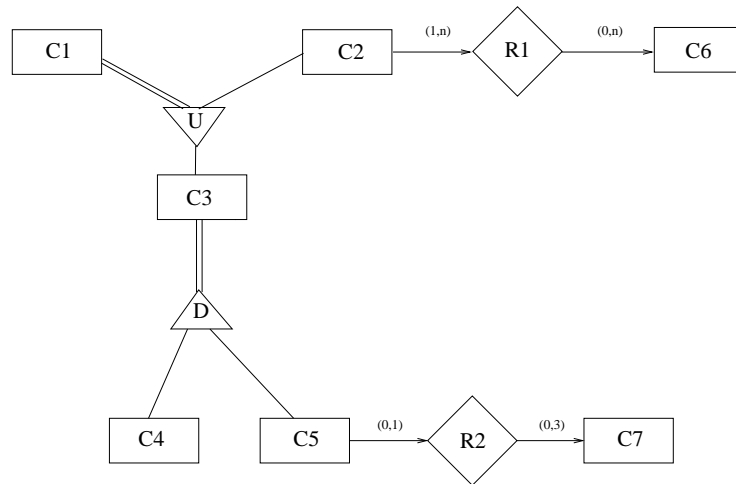


Figure 1: EER Diagram

(30) For the given EER diagram, state which of the following are models¹:

- $C_1(i_1) \wedge C_3(i_1)$. _____.
- $C_1(i_1) \wedge C_3(i_1) \wedge C_5(i_1)$. _____.
- $C_2(i_1) \wedge C_3(i_1) \wedge C_4(i_1) \wedge C_5(i_1)$. _____.
- $C_1(i_1) \wedge C_3(i_1) \wedge C_5(i_1) \wedge C_7(i_2) \wedge C_7(i_3) \wedge R_2(i_1, i_2) \wedge R_2(i_1, i_3)$. _____.
- $C_1(i_1) \wedge C_6(i_1) \wedge C_5(i_1) \wedge C_7(i_2) \wedge R_2(i_1, i_2)$. _____.

Now assume that the following string typed attributes are attached to the entities and relationships:

C3 : a	C6: 1d6 (PK)	C6: d
C3 : id3 (PK)	C4: b	C7: e
C1 : id1 (PK)	C5: c	
C2 : id2 (PK)	C7: id7 (PK)	

¹Anything not stated is false. For a relationship, the subject argument is followed by the object argument.

EXAM NUMBER: (no name!)

(30) Using the mapping algorithm on page 18, give a relational schema that corresponds to EER diagram with attributes on the previous page.

(20) Discuss the limitations of the resulting schema if you are limited to just the primary and foreign key constraints of SQL2.

3. Temporal Databases

Consider that we have a cash register that records the exact moment at which orders are placed (and perhaps when they are retracted). There also exists a record of the cash transfers that signal when an order completes. Here we see a person orders a Burger, Fries and a Cola, but then changed their mind from Cola to Sprite.

```
Sale(T#, R#, Item, Price, TST, TET)
001, R1, Burger, 1.20, 10:02.03, uc)
001, R1, Fries, .70, 10:02.13, uc)
001, R1, Cola, 1.10, 10:02.23, 10:02.45)
001, R1, Sprite, 1.10, 10:03.01, uc)
```

```
CashExchange(T#, R#, TST, TET)
001, R1, 10:03.13 uc
```

```
CashInRegister(R#, Time)
R1, 10.00 67.34
R1, 22.15 384.32
```

In SQL

- a. (30) Give the total number of burgers sold on register R1 between 10.00 and 13.00 (you may assume that all transactions must complete within 15 minutes, that this query is being issued at 17.00, and that the store opens at 10.00)

EXAM NUMBER: (no name!)

b. (30) Give all transactions that were revised more than once.

c. (20) Give all the registers that don't balance over the course of the day.

4. Datalog Queries

```
Person(id, name, nationality, gender).
Parent(parentId, childId).
AncientTribeMember(id, nationality).
```

```
Person(001, clovis, french, M)
Person(002, marie, french, F)
Person(003, gerhart, german, M)
Person(004, ann, french, F)
Person(005, paul, french, M)
Person(006, susan, french, F)
Person(007, dan, french, M)
Parent(001,005)
Parent(002,005)
Parent(003,006)
Parent(004,006)
Parent(005,007)
Parent(006,007)
AncientTribeMember(001, french).
AncientTribeMember(002, french).
AncientTribeMember(004, french).
```

- a. (30) Give the program that derives the ids for people with a French ancestor (French ancient tribe members have French ancestors).

EXAM NUMBER: (no name!)

b. (30) Give the program that derives the ids of persons who have only French ancestors (use negation).

c. (30) Repeat b without using negation.

EXAM NUMBER: (no name!)

5. XML

- a. (30) Define a DTD corresponding to the schema in problem 4.

EXAM NUMBER: (no name!)

b. (30) Create a document of this DTD type that captures the database state depicted in problem 4. (You need only do this for persons 001,002 and 005)

c. (20) Give XQuery/XPath query that obtains all French males.

6. Association Rule Mining

You are given the following transactions

1. {a, b, c, d}
2. {a, b, c}
3. {a, b, d}
4. {b, c, d}
5. {a, e}
6. {a, b, c}
7. {a, c, d}
8. {b, c, e}

(80) Now, using the Apriori algorithm (page 85) find the associations rules with support above $1/4$ (that is $\beta > 1/4$) and confidence over $2/3$ (that is $\alpha > 2/3$). Indicate which item sets have counts calculated over the database.