Høgskolen i Telemark

EKSAMEN

5602 DATABASER

08.12.2011

Tid: *9-14*

Målform: *Bokmål/nynorsk/engelsk*

Sidetall: *28 (inkludert denne forsiden)*

Hjelpemidler: *Ingen*

Merknader: *Ingen*

Vedlegg: *A: Eksempeldata og B: Svarark til oppgave 4*

Eksamensresultatene blir offentliggjort på Studentweb.

Fakultet for allmennvitenskapelige fag

Each question is assigned a weight, written in parenthesis, e.g. (30%).

Question 1 (30%)

The database tables below is to be used by the police to keep track of individuals and vehicles involved in accidents, where primary keys are underlined and foreign keys are marked with a star:

- Person(<u>ID</u>, FirstName, LastName, BirthDate, Gender)
- Vehicle(<u>RegNo</u>, Brand, Model, ProdYear)
- Accident(<u>ANr</u>, AccDate, Road)
- PersonInAccident(<u>ANr*, ID*</u>, Role)
- VehicleInAccident(<u>ANr*, RegNo*</u>)

Person.ID og Accident.ANr are serial numbers (auto numbers) identifying persons and accidents, respectively, whereas Vehicle.RegNo is the standard registration number (licence plate). For each accident, the date (AccDate) and road is stored. The tables PersonInAccident and VehicleInAccident show which persons and vehicles have been involved in which accidents. PersonInAccident.Role shows the role a given person had in a particular accident, for example driver, passenger or pedestrian. Appendix A shows example data for each table. You will now use SQL to solve typical tasks using this database.

1-a (5%)

Write an SQL query to show all accidents on road E134 in the year 2010. Tip: You may use the function Year to extract the year from a date.

1-b (5%)

Write an SQL query to show all accidents a given vehicle has been involved in, ordered with respect to date – from old to new. Let the registration number be a parameter.

1-c (5%)

Write an SQL query to show the number of accidents in 2010 for each brand, i.e. the query should, for each brand, show the number of accidents where vehicles of that brand has been involved.

1-d (5%)

Write an SQL query to define the PersonInAccident table. Make sure that Role must be entered.

1-e (5%)

For the accident having ANr=1002 (see appendix A) it is found that a person with name Karl Hansen born 14.07.1973 has been involved as a pedestrian. Give him ID=5. Use SQL to store this information.

1-f (5%)

Write an SQL query to show the number of persons and the number of vehicles involved in each accident. Tip: You may wish to use several queries here.

Question 2 (30%)

This question is about E/R and contains two independent parts.

2-a (5%)

Translate the data model shown below to database tables (SQL code is not necessary). Underline <u>primary keys</u> and mark foreign keys with stars (*). Entities and attributes are given "symbolic names" with a purpose – the aim is to test understanding of the rules for translating E/R diagrams to tables.
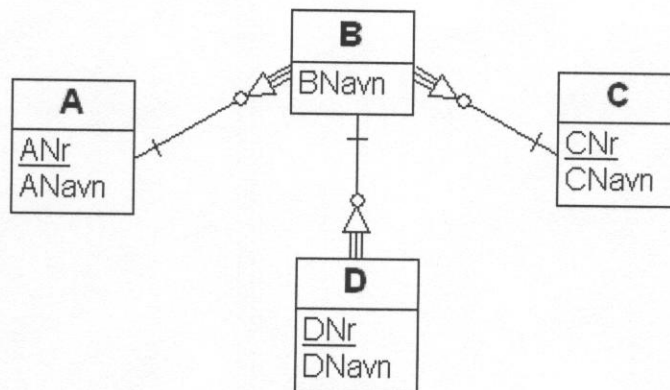
2-b (25%)

Friskus is a physical training company in Utopia, having a number of health clubs throughout the country. The company is now about to build an information system to handle customers, training offers, instructors and ticket sale.

The system should store contact information about each health club (address, zip code and telephone). Customers need to register with name and telephone number before using the facilities. These data should be stored in the system.

Friskus offers guided group training of various kinds, e.g. aerobic, spinning, sumba and yoga, but also have training rooms for individual exercise using manuals and other kinds of training devices. The type of group training offered at each health club varies.

Every health club has the same opening hours, from 8 am to 8 pm – Monday to Friday. Each class is 1 hour, and may therefore be numbered from class 1 (8 am to 9 am) through to class 12 (7 pm to 8 pm). The schedule is the same for all weeks, but may vary from health club to health club. The system should keep track of the schedule for each health club, where for example it is defined that yoga is offered in class 1 every Monday.

All health clubs are identical, having one room for individual exercise and one room for group training. All rooms are of the same size.

Friskus hires instructors to lead group trainings. The system should store the name and telephone number for each instructor, and also store what types of training each instructor may lead. When the schedule is defined, the instructors are assigned to classes, and this is also stored in the system. Example: Instructor Marit Mo leads the yoga training in class 3 each Monday at health club 27.

The customers may buy single tickets both for individual exercise and for group training. Also, it is possible to buy tickets for 10 or 30 classes, as well as month and season tickets. The system should store the price for all types of tickets. The system should also store each sale with date.

Customers can make reservations for given classes, in order to be guaranteed admission. This should also be stored in the system.

You should build a conceptual data model (E/R diagram) for this system. Write down assumptions you make in case the text is unclear or ambiguous.

Question 3 (20%)

3-a (10%)

The table Sample contains data about water samples that pupils at an elementary school have collected in lakes nearby:

- Sample(sampleno, pno, firstname, datetaken, lid, lname, ph, temp)

The example row below shows that sample 176 was taken by pupil 24 (Kari) on August 24 2011 in lake number 5 (Pyttvann), and in this sample the pH (acidity) was measured to 5.6 and the temperature to 13.2 degrees Celcius:

(176, 24, 'Kari', '24.08.2011', 5, 'Pyttvann', 5.6, 13.2)

Make any assumptions you find necessary and reasonable. Write down all functional dependencies, decide the candidate key and normal form, and then normalize the table to BCNF.

3-b (5%)

The database used in question 1 will be used by the police, and all police(wo)men in the country should be able to view the table content, but only a smaller number of employees should be allowed to make changes (insert, update, delete). Show examples of SQL code to achieve this, and explain the purpose of using roles (groups).

3-c (5%)

Two phase locking does not guarantee against deadlock. Give an example showing that a deadlock may occur even if all transactions follow two phase locking.

Question 4 (20%)

- This part contains 25 multiple choice questions.
- Some questions refer to tables in appendix A. These are marked [See appendix].
- Each question has 4 alternatives. Only 1 alternative is correct. You get 3 points for each correct answer, -1 points for each wrong answer and 0 points for unanswered questions.
- Answer the question by writing a cross (×) in the correct alternative in appendix B.

1. In which SQL command do we use the reserved word VALUES?
   a. INSERT
   b. SELECT
   c. DELETE
   d. UPDATE
2. [See appendix] What is correct?
   a. A determinant is a candidate key.
   b. A super key is a minimal candidate key.
   c. A primary key is also a candidate key.
   d. A candidate key is also a primary key.
3. [See appendix] Which column is a foreign key?
   a. Accident.ANr
   b. Person.ID
   c. Vehicle.RegNo
   d. PersonInAccident.ANr
4. [See appendix] Which WHERE expression means the same as ANr<1002 ?
   a. NOT (ANr >= 1002)
   b. NOT (ANr=1002)
   c. NOT (ANr<1002)
   d. ANr<=1002 OR ANr=1002
5. [See appendix] How many rows are deleted by the query DELETE FROM Vehicle WHERE Brand = "Ford" OR ProdYear<2010 ?
   a. 1
   b. 2
   c. 3
   d. 4
6. [See appendix] How many rows are returned from the SQL query SELECT * FROM Person, PersonInAccident, Accident?
   a. 6
   b. 15
   c. 45
   d. 120

7. [See appendix] How many rows are returned from the SQL query SELECT * FROM Person INNER JOIN PersonInAccident ON Person.ID= PersonInAccident.ID ?
   a. 4
   b. 6
   c. 8
   d. 24

8. [See appendix] How many rows are returned from the SQL query SELECT * FROM Accident LEFT OUTER JOIN VehicleInAccident ON Accident.ANo = VehicleInAccident.ANo ?
   a. 6
   b. 0
   c. 30
   d. 5

9. [See appendix] How many rows are returned from the SQL query SELECT Brand, COUNT(*) FROM Vehicle GROUP BY Brand?
   a. 4
   b. 2
   c. 0
   d. 3

10. [See appendix] When do we need to use GROUP BY?
    a. To show all accidents ordered by date.
    b. To show the number of accidents in 2010.
    c. To show the number of males involved in accidents in 2010.
    d. To show the number of accidents in each month in 2010.

11. Which expression in relation algebra returns the most number of rows, if we assume that both relation (table) A and B contains more than 2 rows?
    a. A∩B   (intersection)
    b. A×B   (cross product)
    c. A∪B   (union)
    d. A–B   (difference)

12. What do we mean when we say that there is a functional dependency from A to B?
    a. A and B contain the same number of values.
    b. A cannot contain repetitions if B contain repetitions.
    c. Two rows having equal A values must also have equal B values.
    d. Two rows having equal B values must also have equal A values.

13. [See appendix] Assume that information about the production year was added to the table VehicleInAccident. What normal form would this table have?
    a. 1NF
    b. 2NF
    c. 3NF
    d. BCNF

14. Entities correspond to tables and attributes correspond to …?
    a. Primary keys
    b. Columns
    c. Rows
    d. Foreign keys

15. How are many-to-many relationships handled during logical design?
    a. The primary key on the left hand side is copied to the right hand side.
    b. The primary key on the right hand side is copied to the left hand side.
    c. A junction table (koblingstabell) is added.
    d. The entities are collapsed into a single table.
16. [See appendix] Which view (utsnitt) is updatable?
    a. CREATE VIEW X AS SELECT DISTINCT Firstname FROM Person
    b. CREATE VIEW X AS SELECT COUNT(*) FROM Person
    c. CREATE VIEW X AS SELECT ID, Firstname FROM Person WHERE Firstname LIKE 'A*'
    d. CREATE VIEW X AS SELECT ID, COUNT(*) FROM Person GROUP BY ID
17. How many comparisons are needed in the worst case for finding a given number using binary search in a sorted table containing 400 numbers?
    a. 200.
    b. 40.
    c. 20.
    d. 10.
18. A person table containing 400 000 rows are indexed using a B tree on the column PNr. It is room for 40 table rows in each block, and the B tree has 4 levels. How many blocks do we need to read in order to find the last name (family name) for a person having a particular PNr?
    a. 200 000
    b. 5
    c. 20
    d. 5000
19. How many bits are sufficient (in theory) for storing 15 different values?
    a. 4
    b. 5
    c. 2
    d. 3
20. What does A in ACID stands for?
    a. Alter
    b. Atomicity
    c. Abstract
    d. Analytical
21. What is correct to say regarding read locks and write locks?
    a. Several transactions may get a read lock on the same row simultaneously.
    b. Several transactions may get a write lock on the same row simultaneously.
    c. A transaction must get both a read lock and a write lock before it can start.
    d. A transaction must set all its read locks before its write locks.
22. In two-phase locking, a transaction must:
    a. Set all its read locks before its write locks.
    b. Set all its locks at the end, immediately before COMMIT.
    c. Never set a lock after it has released its first lock.
    d. Either use read locks or write locks (but not both).

23. What does the command COMMIT do?
    a. Revokes a privilege on a database object from a user.
    b. Confirms a transaction.
    c. Grants a privilege on a database object to a user.
    d. Regrets a transaction.
24. In which order is the phases executed according to the waterfall model?
    a. Normalization, data modelling, table definition.
    b. Analysis, design, implementation.
    c. CREATE TABLE, INSERT, SELECT.
    d. 1NF, 2NF, 3NF, BCNF.
25. What is the difference between a star schema and a snowflake schema?
    a. The star schema is normalized.
    b. The snowflake schema is normalized.
    c. The star schema does not contain fact tables.
    d. The snowflake schema does not contain fact tables.
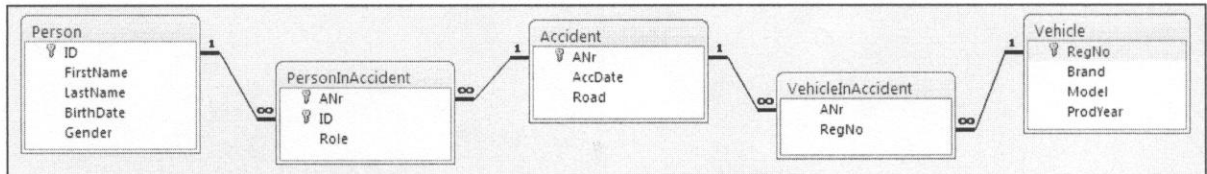
--- The end ---

Appendix A: Example data

This appendix contains example data for the database tables defined in question 1. The tables are also referred to in some parts of question 4.

Data model

The figure below shows the relationship window in Access for this database.

Person

ID	Firstname	Lastname	BirthDate	Gender
1	Per	Hansen	03.04.1982	M
2	Lise	Jensen	17.07.1988	F
3	Ola	Li	23.10.1950	M
4	Åse	Mo	04.09.1993	F

Vehicle

RegNo	Brand	Model	ProdYear
DA88997	Volkswagen	Golf	2003
LY12345	Toyota	Avensis	2007
NV33221	Ford	Focus	2010
PN41412	Ford	Focus	2008

PersonInAccident

ANr	ID	Role
1001	1	Driver
1001	2	Driver
1001	3	Passenger
1002	1	Driver
1003	2	Driver
1003	4	Pedestrian

Accident

ANr	AccDate	Road
1001	08.05.2008	E18
1002	09.11.2010	E134
1003	17.01.2010	E6
1004	08.04.2010	E134
1005	12.11.2010	E134

VehicleInAccident

ANr	RegNo
1001	LY12345
1001	PN41412
1002	NV33221
1003	DA88997
1004	LY12345

Appendix B: To be used in question 4

Write a single × on each line (question).

Question	a	b	c	d
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				