Final exam in 1DL301 Database Design I

Department of Information Technology, Uppsala University
August 21, 2018, 08.00-13.00

This is a multiple-choice exam. Each question might have several correct choices, but you must select only one choice. If you select more than one, your answer will be marked as incorrect (even if you have selected only correct choices).

The questions are divided into four sections:

| Section | Questions |
| :--- | ---: |
| Data modeling | 10 |
| SQL | 10 |
| FDs, CKs, NFs and normalization | 10 |
| Other | 6 |

To pass the exam you need:

- to answer correctly to at least $50 \%$ of questions in each section, and
- at least 22 correct answers in total.

Failing to meet any of these criteria means failing the exam (i. e. grade U).
If you fulfill both criteria, your grade will be determined by the number of correct answers:

| Correct answers | Grade |
| :--- | :--- |
| $22-26$ | 3 |
| $27-31$ | 4 |
| $32-36$ | 5 |

Your answers must be given on the answer sheet which will be handed in. Don't forget to fill out your exam code. To mark your answer fill the answer box entirely using a dark colored pen (black or blue). The optical character recognition system will not recognize ticks, crosses, circles or any other additional notes. Remember, do not to fill more than one answer box per question.

If you make a mistake on the answer sheet, request a new one and make sure you hand in the correct sheet (if you hand in several answer sheets, your exam will not be graded).

You can keep the question sheets. We recommend that you first note your answers on these and fill out the answer sheet just before handing it in.

The answer key will be published after 2pm on http://user.it.uu.se/~janku655/db1/.

[^0]
## 1 Information about you

Question i (Not assessed.) What is your study program?
A IT
B STS
C X
D Other CS/math/physics
E Other non-CS/math/physics

Question ii (Not assessed.) In which period were you registered in the course for the first time?
A Fall 2017, period 2
B Fall 2017, period 1
C Earlier period

Question iii (Not assessed.) How many lectures have you attended?
A None or very few
B Around 25\%
C Around 50\%
D Around 75\%
E Almost all or all

## 2 Data modeling

Question 1 What ER element is depicted in the following figure?

$$
X
$$

A Weak attribute
B Identifying relationship
C Key attribute
D Weak entity type
E Important entity type

Question 2 Which of the ER diagrams is equivalent to the depicted ER model?

A


C


Question 3 In an online bookstore a customer can buy one or several books in a single order. It is of course possible to buy several copies of the same book. The prices of books might change so we need to store the prices of bought books in the order as well. Which of the following ER diagrams shows a correct way how to store orders? (Only the relevant portion of each diagram is shown, Qty stands for quantity. Ignore the value added tax.)
A

C

B



Question 4 Which of the following statements about the depicted ER diagram is correct?


A Each $E_{1}$ entity must be related to exactly one $E_{2}$ entity. An $E_{2}$ entity can be related to zero, one or more $E_{1}$ entities.

B An $E_{1}$ entity can be related to zero or one $E_{2}$ entities. Each $E_{2}$ entity must be related to one or more $E_{1}$ entities.

C An $E_{1}$ entity can be related to zero or one $E_{2}$ entities. An $E_{2}$ entity can be related to zero, one or more $E_{1}$ entities.

D The ER model is not semantically correct.
E Each $E_{1}$ entity must be related to exactly one $E_{2}$ entity but it can be related to it multiple times. An $E_{2}$ entity can be related to an $E_{1}$ entity but only once.

Question 5 The ER diagram from the previous question is equivalent to which of the following diagrams?
A

C

B

D


Question 6 Based on the depicted ER diagram and the semantics of ER models, which of the following statements is correct?


A Each employee works on at least one project.
B The salary of an employee cannot be a negative number.
C All employees have a unique full name (the combination of the first and last name).
D Each project has at least one employee working on it.
E The model is wrong, relationships cannot have attributes.

Question 7 Convert the ER model from the previous question to a relational model.
A Employee(Id, FirstName, LastName, Salary, WorksOnId) with WorksOnId ${ }^{\mathrm{FK}} \rightarrow$ WorksOn(Id) Project(Id, Title, Budget, WorksOnId) with WorksOnId ${ }^{\mathrm{FK}} \rightarrow$ WorksOn(Id) WorksOn(Id, HoursSpent)

B The ER model cannot be converted to a relational model because both Employee and Project have an attribute with the same name.

C Employee(Id, FirstName, LastName, Salary)
Project(Id, Title, Budget)
WorksOn(Eid, Pid, HoursSpent) with Eid ${ }^{\mathrm{FK}} \rightarrow$ Employee(Id) and Pid ${ }^{\mathrm{FK}} \rightarrow \operatorname{Project}(\mathrm{Id})$
D Employee(Id, FirstName, LastName, Salary)
Project(Id, Title, Budget)
WorksOn(Eid, Pid, HoursSpent) with Eid ${ }^{\mathrm{FK}} \rightarrow$ Employee(Id) and Pid ${ }^{\mathrm{FK}} \rightarrow \operatorname{Project}(\mathrm{Id})$
E Employee(Id, FirstName, LastName, Salary) with Id ${ }^{\mathrm{FK}} \rightarrow$ WorksOn(Eid)
Project(Id, Title, Budget) with Id ${ }^{\mathrm{FK}} \rightarrow$ WorksOn(Pid)
WorksOn(Eid, Pid, HoursSpent)

Question 8 Convert the following ER model to a relational model.


A $E_{1}\left(\underline{A_{11}}, A_{12}\right)$
$E_{2}\left(\underline{A_{11}}, A_{21}, A_{22}\right)$ with $A_{11}{ }^{\mathrm{FK}} \rightarrow E_{1}\left(A_{11}\right)$
B $E_{1}\left(A_{11}, A_{12}\right)$
$E_{2}\left(A_{21}, A_{22}\right)$ with $A_{21}{ }^{\mathrm{FK}} \rightarrow E_{1}\left(A_{11}\right)$
C $E_{1}\left(A_{11}, A_{12}\right)$
$E_{2}\left(\underline{A_{11}}, A_{21}, A_{22}\right)$ with $A_{11}{ }^{\mathrm{FK}} \rightarrow E_{1}\left(A_{11}\right)$ and $A_{21}$ being unique
D $E_{1}\left(A_{11}, A_{12}\right)$ with $A_{11}{ }^{\mathrm{FK}} \rightarrow E_{2}\left(A_{11}\right)$
$E_{2}\left(\underline{A_{11}}, A_{21}, A_{22}\right)$
E $E_{1}\left(A_{11}, A_{21}, A_{12}\right)$ with $A_{21}{ }^{\mathrm{FK}} \rightarrow E_{2}\left(A_{21}\right)$
$E_{2}\left(\underline{A_{21}}, A_{22}\right)$

Question 9 Consider the following ER model for a database of games in the next season of SHL (a Swedish professional ice hockey league). Convert the model to a relational model.


A Team(Id, City, Coach)
Game(HomeId, AwayId, Date, Score) with HomeId ${ }^{\text {FK }} \rightarrow$ Team(Id, City), AwayId ${ }^{\text {FK }} \rightarrow$ Team(Id, City)
B Team(Id, City, Coach, AwayId) with AwayId ${ }^{\mathrm{FK}} \rightarrow$ Team(Id)
Game(HomeId, Date, Score) with HomeId ${ }^{\mathrm{FK}} \rightarrow$ Team(Id)
C Team(Id, City, Coach) Game(Date, Score)

D Team(Id, City, Coach) Game(HomeId, AwayId, Date, Score) with HomeId ${ }^{\mathrm{FK}} \rightarrow$ Team(Id), AwayId ${ }^{\mathrm{FK}} \rightarrow$ Team(Id)

E Team(Id, City, Coach) Game(HomeId, AwayId, Date, Score) with HomeId ${ }^{\mathrm{FK}} \rightarrow$ Team(Id), AwayId ${ }^{\mathrm{FK}} \rightarrow$ Team(Id)

Question 10 A company has two types of projects: X and Y. Each project is either of type X or Y, but not both at the same time. Since many (but not all) of the attributes we wish to store for both project types are the same (such as the id, title, department, project manager, start date, etc.), we want to use a superclass to store these attributes. Which of the following ER models corresponds to this specification? (The attributes are not shown.)


## 3 SQL

For questions $11-15$ consider the following database consisting of two tables T1 and T2:

T1

| C1 | C2 |
| :---: | :---: |
| A | A |
| A | B |
| B | C |
| D | D |

T2

| C 1 | C 2 |
| :---: | :---: |
| A | D |
| A | A |
| B | A |
| C | $N U L L$ |

Question 11 How many rows are in the result of the following SQL query?

```
SELECT *
FROM T1, T2
```

A 15
(B) 12
(C) 8
(D) 16

E 4

Question 12 What is the result of the following SQL query? (Not showing the header of the result.)

```
SELECT DISTINCT T1.C1, T1.C2
FROM T1 JOIN T2 ON T1.C2=T2.C1
WHERE T2.C2 IS NOT NULL;
```

A

| A | A |
| :---: | :---: |
| A | B |
| B | C |

B

| A | A |
| :---: | :---: |
| A | B |

D Empty table

| EA A <br>  A <br> A A <br>  A B |  |  |
| :---: | :---: | :---: |
|  | B | C |


| A |  | A |
| :---: | :---: | :---: |
| C | A | A |
| A |  |  |

Question 13 What is the result of the following SQL query? (Not showing the header of the result.)

```
SELECT T1.C1, COUNT(T2.C2)
FROM T1, T2
WHERE T1.C2 = T2.C1
GROUP BY T1.C1
```

| A | A | 3 |
| :--- | :--- | :--- |
|  |  |  |

D SQL is invalid

| B | A | 3 |
| :--- | :--- | :--- |
|  | B | 0 |


| E | A | 3 |
| :--- | :--- | :--- |
|  | B | 1 |
|  |  |  |


| A 2 <br>  A | 1 |
| :---: | :---: | :---: |
| B | 1 |

Question 14 What is the result of the following SQL query? (Not showing the header of the result.)

```
SELECT T1.C1
FROM T1, T2
WHERE T1.C2 = T2.C1 AND COUNT(T2.C2) > 0
GROUP BY T1.C1
```

A | A |
| :--- |

D Empty table
B SQL is invalid

| E | $\begin{array}{l}\text { A } \\ \\ \\ n\end{array}$ |
| :---: | :---: |

C A

Question 15 What is the result of the following SQL query? (Not showing the header of the result.)

```
SELECT T1.C1, T1.C2
FROM T1
LEFT JOIN T2 ON T1.C2 = T2.C1
WHERE T2.C2 = 'D' OR T2.C2 IS NULL
```

| A | A | A |
| :--- | :--- | :--- |
|  | D | D |


| D |  |  |  | A | A |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | B | C |  |  |  |
| D |  |  |  |  |  |

B Empty table

E

| A | A |
| :--- | :--- |
| B | C |

F SQL is invalid

For questions $16-20$ consider the following database model:


Question 16 What is the correct SQL statement to create the employee_project table (which stores the information about which employee is working on which project and how many hours he/she has spent on it so far), including the primary and foreign keys?

A CREATE TABLE employee_project (
employee_id int NOT NULL,
project_id int NOT NULL,
hours_spent int)
B CREATE TABLE employee_project (
employee_id int NOT NULL,
project_id int NOT NULL,
hours_spent int,
PRIMARY KEY (employee_id, project_id),
FOREIGN KEY (employee_id) REFERENCES employee(id),
FOREIGN KEY (project_id) REFERENCES project (id))
C CREATE TABLE employee_project (
employee_id int NOT NULL PRIMARY KEY,
project_id int NOT NULL,
hours_spent int,
FOREIGN KEY (employee_id) REFERENCES employee(id),
FOREIGN KEY (project_id) REFERENCES project(id))
D CREATE TABLE employee_project (
employee_id int NOT NULL PRIMARY KEY,
project_id int NOT NULL PRIMARY KEY,
hours_spent int,
FOREIGN KEY employee (id) REFERENCED BY employee_id,
FOREIGN KEY project(id) REFERENCED BY project_id)
E CREATE TABLE employee_project (
employee_id int NOT NULL,
project_id int NOT NULL,
hours_spent int,
PRIMARY KEY (employee_id, project_id),
FOREIGN KEY employee(id) REFERENCED BY employee_id AND project(id) REFERENCED BY project_id)

Question 17 Finish the SQL query to return the ID and the title of each department together with the number of its employees:

```
SELECT d.id, d.title, count(e.id)
FROM department d, employee e
(1)
```

A None of the other answers
B (1) ON e.department_id=d.id GROUP BY department
C (1) WHERE e.department_id=d.id GROUP BY d.id, d.title
(D) (1) ON d.employee_id=e.id GROUP BY d.id

E (1) WHERE e.department_id=d.id GROUP BY e.id

Question 18 Which of the following SQL queries selects the ID and the surname of the best paid employees in each department?

A SELECT id, last_name, hour_salary FROM employee GROUP BY department_id HAVING hour_salary=MAX (hour_salary)

B SELECT e.id, e.last_name, e.hour_salary FROM employee e WHERE e.hour_salary = (SELECT MAX(hour_salary) FROM employee WHERE department_id=e.department_id)

C SELECT e.id, e.last_name, e.hour_salary FROM employee e WHERE NOT EXISTS (SELECT * FROM employee WHERE hour_salary >= e.hour_salary)

D SELECT e.id, e.last_name, e.hour_salary FROM employee e WHERE EXISTS (SELECT * FROM employee WHERE hour_salary > e.hour_salary AND department_id $=$ e.department_id)

E SELECT e.id, e.last_name, MAX(e.hour_salary) FROM employee e GROUP BY e.department_id

Question 19 Finish the SQL query to return the ID and the title of each project and its cost so far:

```
SELECT p.id, p.title, (1)
FROM project p, employee_project ep, employee e
WHERE (2)
GROUP BY p.id, p.title
```

A (1) $\operatorname{SUM}(e p . h o u r s) * S U M\left(e . h o u r \_s a l a r y\right)$
(2) ep.employee_id=e.id

B None of the other answers
C (1) $\operatorname{SUM}(e p . h o u r s) * A V G\left(e . h o u r \_s a l a r y\right)$
(2) ep.employee_id=e.id AND ep.project_id=p.id

D (1) SUM (ep.hours) *e.hour_salary
(2) employee_id=id AND project_id=id

E (1) SUM (ep.hours_spent*e.hour_salary)
(2) employee_id=e.id AND project_id=p.id

Question 20 Finish the query to select the ID of employees working on both project 14 and project 15:

```
SELECT employee_id
FROM employee_project
WHERE (1)
```

A project_id IN (14, 15)
B project_id = 14 OR project_id=15
(C) project_id = 14 AND project_id=15

D project_id = 14 INTERSECT project_id=15
E None of the other answers

## 4 FDs, CKs, NFs and normalization

For questions $21-24$, consider the following relation in 1NF:

```
TestResult(Date, SSN, FirstName, LastName, TestID, TestName, Value, Unit)
```

with the following non-trivial FFDs:

- \{Date, SSN, TestID $\} \rightarrow\{$ Value $\}$
- $\{$ SSN $\} \rightarrow\{$ FirstName, LastName $\}$
- \{TestID $\} \rightarrow$ \{TestName, Unit $\}$

Question 21 Which of the following options lists all candidate keys?
A $\{\mathrm{SSN}\}$ and $\{$ TestID $\}$
B \{Date, SSN, TestID $\}$
C] \{Date, SSN, TestID $\},\{$ SSN $\}$ and $\{$ TestID $\}$
D $\{$ Date $\},\{\mathrm{SSN}\}$ and $\{$ TestID $\}$
E There are no candidate keys

Question 22 Which of the following attributes is not a prime attribute?
(A) Date

B All attributes are prime
C] SSN
D Unit
E TestID

Question 23 In which NF is the TestResult relation?
A In 2NF but not in 3NF
B In 3NF but not in 2 NF
C In BCNF
D In 1NF but not in 2 NF
E In 3NF but not in BCNF

Question 24 Which of the following options shows the result of the normalization of the original table with all tables in BCNF (not showing the foreign key constraints)?

A TestResult(Date, Value)
Person(SSN, FirstName, LastName)
Test(TestID, TestName, Unit)
B TestResult(Date, SSN, TestID, Value)
Person(SSN, FirstName, LastName)
Test(TestID, TestName, Unit)
C TestResult(Date, SSN, TestID, TestName, Value, Unit)
Person(SSN, FirstName, LastName)
D TestResult(Date, SSN, TestID)
Person(SSN, FirstName, LastName)
Test(TestID, TestName, Value, Unit)
E TestResult(Date, SSN, TestName, Value)
Person(SSN, FirstName, LastName)
Test(TestName, TestID, Unit)

For questions $25-29$, consider a relation in $1 \mathrm{NF} R(A, B, C, D, E)$ with the following dependencies:

- $\{A, B\} \rightarrow\{C, D, E\}$
- $\{C\} \rightarrow\{A, B, D, E\}$
- $\{E\} \rightarrow\{D\}$

Question 25 What is $\{E\}^{+}$(i. e. the set of all attributes which can be determined by $\{E\}$ )?
A Empty set
B $\{D\}$
C $\{D, E\}$
D $\{A, B, C, D, E\}$
E $\{E\}$
F None of the other answers

Question 26 Which of the following options lists all candidate keys?
A $\{A, B, C, E\}$
B All sets including both $A$ and $B$
C There are no candidate keys
D $\{A, B\}$ and $\{C\}$
E $\{A, B\},\{C\}$ and $\{E\}$
F None of the other answers

Question 27 If we choose $C$ to be the primary key, which of the following options list all prime attributes?
A A, B, C
B C
C A, B, D, E
D All attributes are prime
E A, B
F There are no prime attributes

Question 28 In which NF is $R$ ?
A In 1NF but not in 2 NF
B In 3NF but not in BCNF
C In 3NF but not in 2NF
D In BCNF
E In BCNF but not in 3NF
F In 2NF but not in 3NF

Question 29 Which of the following normalized databases contains all the information contained in the original table, with all relations in BCNF (not showing the primary and foreign key constraints)?

A $R_{1}(A, B, E), R_{2}(C, E), R_{3}(E, D)$
B $R_{1}(A, B, C, D, E), R_{2}(C, A, B, D, E), R_{3}(E, D)$
(C) $R_{1}(A, B, C, E), R_{2}(E, D)$

D $R(A, B, C, D, E)$
E $R_{1}(A, B, D), R_{2}(C, D), R_{3}(E, D)$
F None of the other answers

Question 30 The following table shows the current state of a relation. Is it true that $\left\{C_{1}\right\} \rightarrow\left\{C_{3}\right\}$ ?

| $C_{1}$ | $C_{2}$ | $C_{3}$ | $C_{4}$ |
| :---: | :---: | :---: | :---: |
| 1 | 1 | XX | 157 |
| 2 | 2 | XX | 178 |
| 3 | 1 | XX | 192 |
| 4 | 3 | XY | 183 |
| 5 | 4 | XY | 166 |

A Yes, since if $C_{1}$ is 1,2 or 3 , then $C_{3}$ is XX and if $C_{1}$ is 4 or $5, C_{3}$ is XY.
B A dependency cannot be confirmed (to hold) by just checking one state of the relation.
C No, values of $C_{3}$ are not unique.
D No, since for some values of $C_{3}$ there are several different values of $C_{1}$ (e. g. if $C_{3}$ is XX, $C_{1}$ can be 1, 2 or 3 ).
E Yes, since all rows are distinct.

## 5 Other

Question 31 In the relational model, if $K$ is a candidate key of relation $R$ and $X$ is one of the remaining attributes (i. e. $X \notin K$ ), then:

A $K \cap\{X\}$ is the primary key
B $K \cap\{X\}$ is a candidate key
C $K \cup\{X\}$ is a candidate key
D $K \cup\{X\}$ is a superkey
E None of the other answers

Question 32 If $P_{1}$ is NULL and $P_{2}$ is FALSE, which of the following statements is correct?
A $P_{1}$ AND $P_{2}$ is FALSE, $P_{1}$ OR $P_{2}$ is NULL
B $P_{1}$ AND $P_{2}$ is FALSE, $P_{1}$ OR $P_{2}$ is FALSE
C $P_{1}$ AND $P_{2}$ is TRUE, $P_{1}$ OR $P_{2}$ is FALSE
D $P_{1}$ AND $P_{2}$ is NULL, $P_{1}$ OR $P_{2}$ is FALSE
E $P_{1}$ AND $P_{2}$ is NULL, $P_{1}$ OR $P_{2}$ is NULL

Question 33 The current state of the Account table is

| Account | Balance |
| :---: | :---: |
| X | 2500 |
| Y | 1000 |

An application connects to the database to execute the following set of SQL statements at the isolation level serializable:

BEGIN
UPDATE Account SET Balance=Balance-1000 WHERE Account=' $\mathrm{X}^{\prime}$
UPDATE Account SET Balance=Balance+1000 WHERE Account='Y'
COMMIT
After running the first UPDATE the database server crashes. After the server restarts, what are the balances of accounts X and Y ?

A X: 1500: Y: 2000
B None of the other answers
(C) X: 2500, Y: 1000

D X: 1500 or 2500 depending on if the server managed to save the updated balance back before the crash, Y: 1000

E X: 1500, Y: 1000

Question 34 Bob creates the table $X$. Then, the following sequence of statements is executed, in the given order (the name of the user executing the statement is indicated at the beginning of each statement):

Bob: GRANT SELECT ON X TO Jim WITH GRANT OPTION
Bob: GRANT SELECT, UPDATE ON X TO Ann WITH GRANT OPTION
Jim: GRANT SELECT ON X TO Tim
Ann: GRANT SELECT ON X TO Tim
Jim: REVOKE SELECT ON X FROM Tim
Which privileges (concerning the table X ) does Tim have?
A SELECT with the grant option
B SELECT, UPDATE, both with the grant option
C SELECT without the grant option
D SELECT, UPDATE, both without the grant option
E None

Question 35 Consider a relation $R\left(A_{1}, A_{2}, \ldots, A_{n}\right)$, and a subset of its attributes, $K \subseteq\left\{A_{1}, A_{2}, \ldots, A_{n}\right\}$. If for any legal instance $r$ and any two distinct tuples $t_{1}, t_{2} \in r, t_{1} \neq t_{2}$ it holds that $t_{1}[K] \neq t_{2}[K]$ (where $t[K]$ denotes the projection of $t$ onto the attributes in $K$, i. e. the values in $t$ but only for the attributes in $K$ ), then $K$ is:

A Candidate key
B Primary key
C Superkey
D Foreign key
E Determinant

Question 36 Consider a relation $R(\underline{A}, B, C, D)$ containing $10^{7}$ records. $A$ is the primary key and $C$ contains $10^{5}$ distinct values. The following prepared statement is executed very frequently:

## SELECT D FROM R WHERE C=?

Which of the following indexes can be useful in executing the query?
A One index on A and one on D
B One index on $C$
C One index on (C, D)
D One index on (D, C)
E One index on D

Final exam in 1DL301 Database Design I - Answer sheet
Department of Information Technology, Uppsala University
August 21, 2018, 08.00 - 13.00

| Your exam code |
| :--- |
| $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$ |

Encode the number in your exam code in the table on the right side. For example, if your exam code is AB2097, fill in the box no. 2 in the first column, the box no. 0 in the second, the box no. 9 in the third, and the box no. 7 in the last column. Fill in the boxes entirely!

Use a dark colored pen (blue or black). To mark your answer fill in the box entirely (■)! The OCR software will not recognize ticks, crosses, circles, etc. Do not make any additional notes on this sheet!

If you make an error, ask for a new answer sheet!


Remember, choose only one answer for each question!




[^0]:    Allowed aids One A4 sheet with handwritten notes (both sides can be used) which must be handed in with your exam (remember to write your exam code in a corner). An English explanatory dictionary and/or a translation dictionary between English and your mother tongue.

