Final exam in 1DL301 Database Design I

Department of Information Technology, Uppsala University
January 11, 2019, 14.00 – 19.00

This is a multiple-choice exam with two types of questions:

- If a **question is marked with ♣** you must **select all correct choices**. If you don't select all correct choices or you include any incorrect choice, your answer will be marked as incorrect.

- For all **other questions** you must **select only one choice** even if there are several correct choices. 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 at least 50% of the questions in each section, and**
- **at least 22 correct answers in total.**

Failing to meet either 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.

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. Consider noting your answers on these first and filling out the answer sheet just before handing it in.

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 one of the corners). An English explanatory dictionary and/or a translation dictionary between English and your mother tongue.

1   Information about you

Question i   (Not assessed.) What is your study program?

- ■ Bioinformatics
- ■ IT
- ■ STS
- ■ X
- ■ Other

Question ii   (Not assessed.) In which period were you registered on the course for the first time?
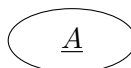
- ■ Fall 2018, period 2 (this period)
- ■ Fall 2018, period 1 (previous period)
- ■ Earlier period

Question iii   (Not assessed.) How many lectures have you attended?

- ■ None or very few
- ■ Around 25%
- ■ Around 50%
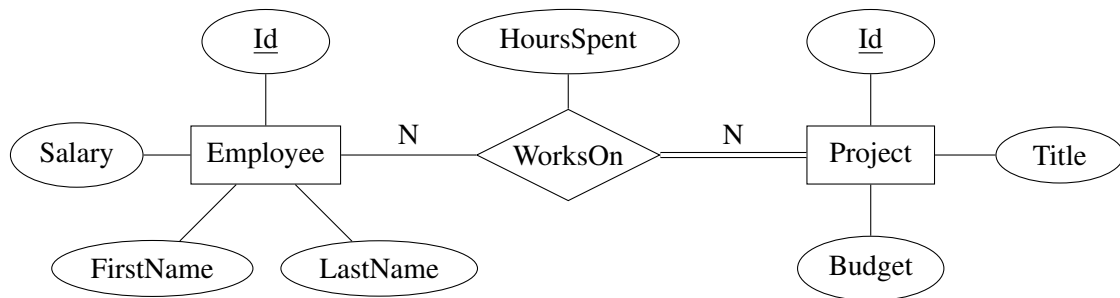- ■ Around 75%
- ■ Almost all or all

2   Data modeling

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

- A Weak attribute
- B Derived attribute
- ■ Key attribute
- D Composite attribute
- E Multi-valued attribute

Question 2 ♣   Based on the ER model depicted below and the semantics of ER models, which of the following statements are correct? (Select **all** correct choices.)


A   All projects have a unique title.

B   The model is wrong, relationships cannot have attributes.

C   Some projects do not have any employees working on them.

D   Each employee works on at least one project.

■   An employee does not have to work on any project, but can also work on several.

■   All employees have a unique id.

■   Each project has at least one employee working on it.

Question 3   We wish to record the weight of patients during their stay in a hospital. Each patient is weighed once a week. Which of the following ER models can be used? (Only the relevant portion of each model is shown.)
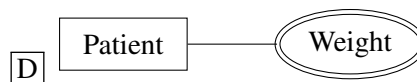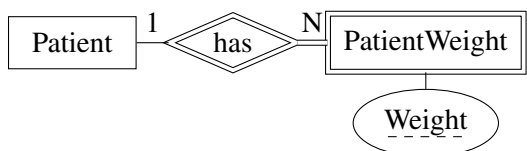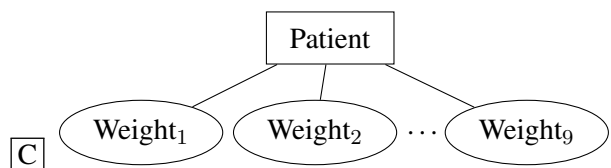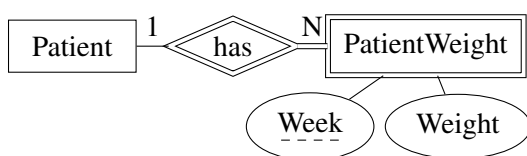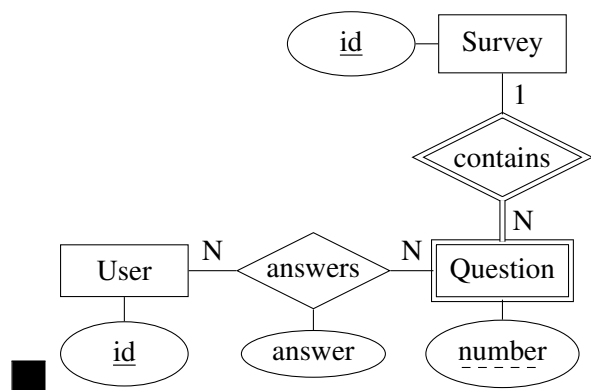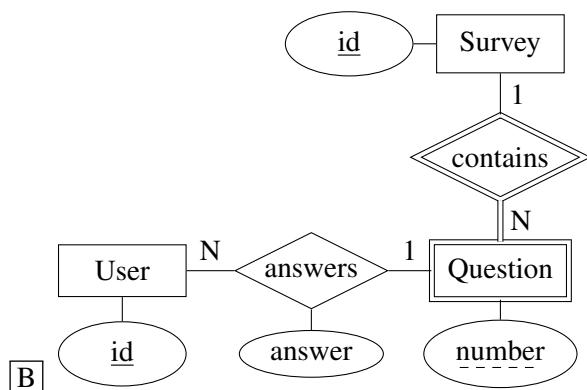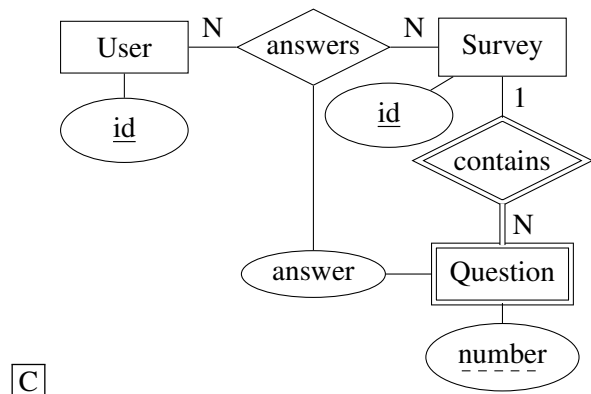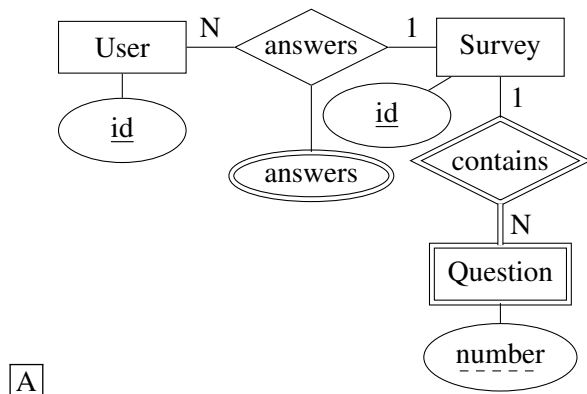
Question 4   Consider the following ER model:

$$E_1 \quad 1 \quad R \quad N \quad E_2$$
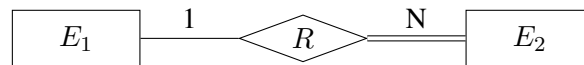
Which of the following diagrams using the min-max notation depicts the same model?

A E_1 (1,1) R (0,N) E_2

■ E_1 (0,N) R (1,1) E_2

B E_1 (1,N) R (0,1) E_2

D E_1 (0,1) R (1,N) E_2

Question 5   You have been asked to implement a very simple survey application. A survey consists of one or more scale questions: respondents choose the answer for each question on a scale from 1 to 5. Each registered user can fill out each survey only once. Surveys are not anonymous, i.e., we need to know which answers were given by which users. Which of the following ER diagrams shows a correct way to store the answers? (Only the relevant portion of each diagram is shown.)

A   User N answers 1 Survey, id, answers, id, contains 1/N Question, number

C   User N answers N Survey, id, id, contains 1/N Question, answer, number

B   id Survey, contains 1/N, User N answers 1 Question, id, answer, number

■   id Survey, contains 1/N, User N answers N Question, id, answer, number
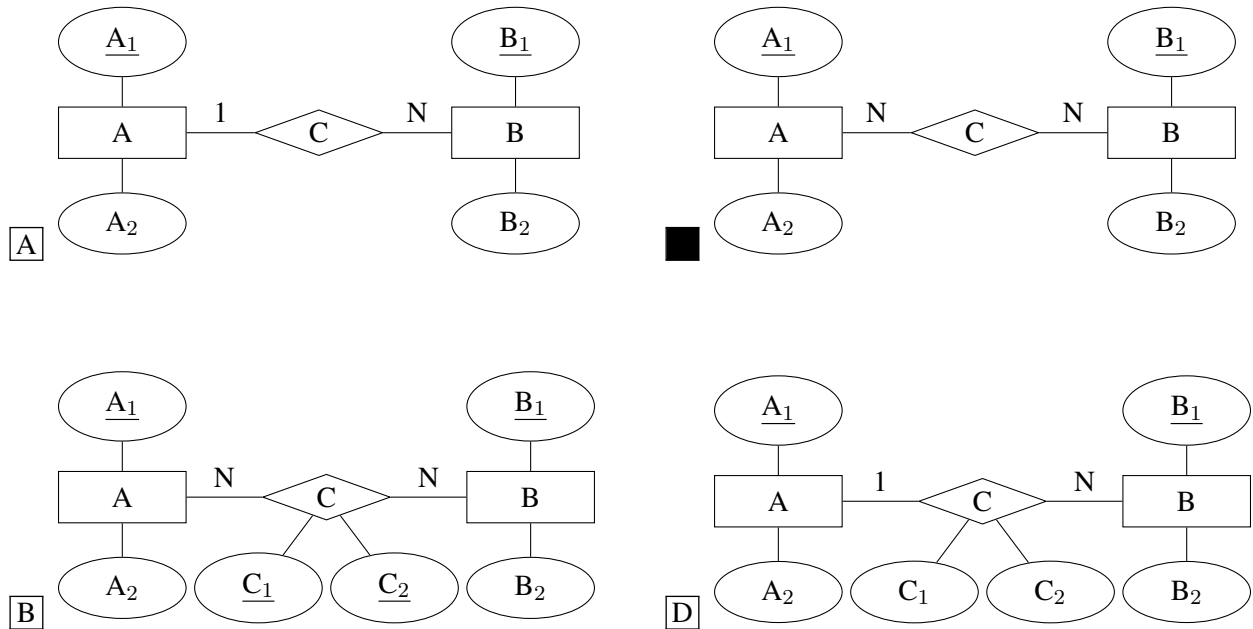
Question 6   Which of the ER models corresponds to the following relational model?

$A(\underline{A_1}, A_2)$
$B(\underline{B_1}, B_2)$
$C(\underline{C_1, C_2})$ with $C_1 \xrightarrow{FK} A(A_1)$ and $C_2 \xrightarrow{FK} B(B_1)$

Question 7   Convert the following ER model to a relational model (foreign key constraints are not shown).
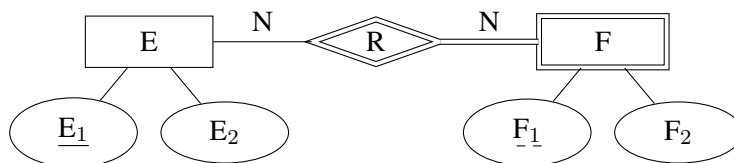
A $E(\underline{E_1}, E_2), F(\underline{F_1}, F_2), R(\underline{E_1, E_2})$

B $E(\underline{E_1}, E_2), F(\underline{E_1, F_1}, F_2)$

C $E(\underline{E_1}, E_2), F(\underline{E_1, F_1}, F_2), R(\underline{E_1, E_2})$

D $F(\underline{E_1, F_1}, E_2, F_2)$

■   None of the other answers

Question 8 Which of the listed ER diagrams **cannot** be used when converting the depicted EER diagram to an ER diagram?

Note: (o) means *optional*.

A

C

D
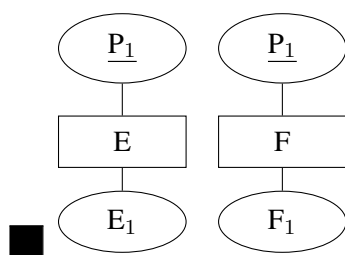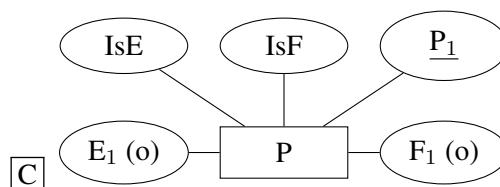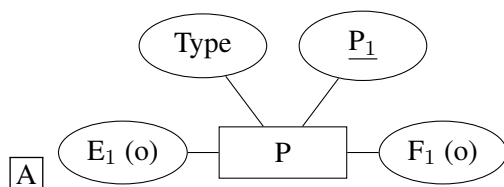
Question 9 Independently of your answer to the previous question, convert the ER diagram with the weak entities (i.e., choice D in Question 8) to a relational model.

A $P(\underline{P_1}, E_1, F_1)$

■ $P(\underline{P_1})$
   $E(\underline{P_1}, E_1)$ with $P_1 \xrightarrow{FK} P(P_1)$
   $F(\underline{P_1}, F_1)$ with $P_1 \xrightarrow{FK} P(P_1)$

C $P(\underline{P_1})$
   $E(P_1, \underline{E_1})$ with $P_1 \xrightarrow{FK} P(P_1)$
   $F(P_1, \underline{F_1})$ with $P_1 \xrightarrow{FK} P(P_1)$

D $P(\underline{P_1})$
   $E(\underline{E_1})$
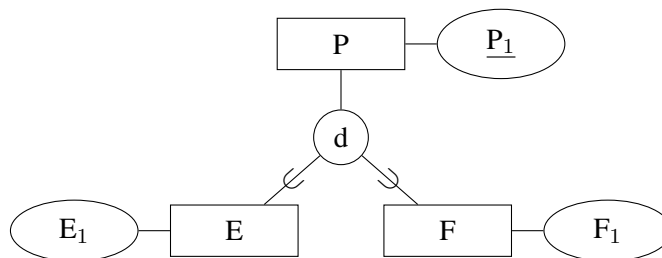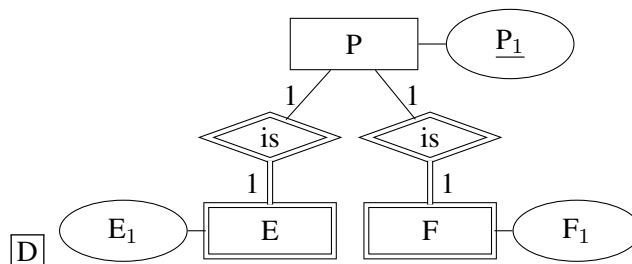 $F(\underline{F_1})$

E $P(\underline{P_1})$
   $E(\underline{P_1}, \underline{E_1})$ with $P_1 \xrightarrow{FK} P(P_1)$
   $F(\underline{P_1}, \underline{F_1})$ with $P_1 \xrightarrow{FK} P(P_1)$

Question 10 ♣   One-to-many relationships can be translated in several ways. Select **all** correct relational models for the depicted ER model. (Hint: There are two correct choices.)

$\boxed{A}$ $E(\underline{E_1}, E_2, F_1)$ with $F_1 \xrightarrow{FK} F(F_1)$
    $F(\underline{F_1}, F_2, E_1)$

■ $E(\underline{E_1}, E_2)$
    $F(\underline{F_1}, F_2, E_1)$ with $E_1 \xrightarrow{FK} E(E_1)$

■ $E(\underline{E_1}, E_2)$
    $F(\underline{F_1}, F_2)$
    $R(E_1, \underline{F_1})$ with $E_1 \xrightarrow{FK} E(E_1)$ and $F_1 \xrightarrow{FK} F(F_1)$

$\boxed{D}$ $EF(\underline{E_1}, E_2, F_1, F_2)$

$\boxed{E}$ $E(\underline{E_1}, E_2)$
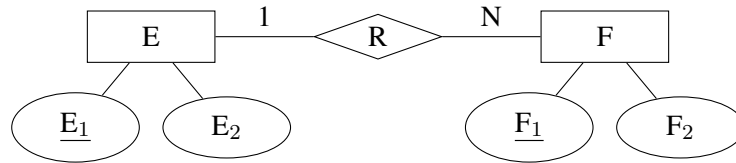 $F(\underline{F_1}, F_2)$
    $R(\underline{E_1}, F_1)$ with $E_1 \xrightarrow{FK} E(E_1)$ and $F_1 \xrightarrow{FK} F(F_1)$

3 SQL

For questions 11 – 17 consider the following database consisting of two tables T1 and T2:

T1

C1	C2	C3	C4
1	100	1	5
2	110	1	NULL
3	120	1	2
4	100	2	2
5	115	4	NULL

T2

C1	C2
1	A
2	B
3	C
4	D
5	E

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

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

$\boxed{A}$ 10

$\boxed{B}$ 15

$\boxed{C}$ None

■ 25

$\boxed{E}$ 8

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

```
SELECT DISTINCT T1.C3, T2.C2
FROM T1, T2
WHERE T1.C3=T2.C1 AND T1.C4 IS NOT NULL
```

A Empty result.

B The SQL is invalid, a JOIN clause must be used
to join T1 and T2.

C

1	A
1	A
2	B

D

1	A
2	B
4	D

■

1	A
2	B

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

```
SELECT C3, COUNT(*), AVG(C2)
FROM T1
GROUP BY C3
```

A

1	5	109
2	5	109
4	5	109

B

1	3	110
2	1	100
3	0	NULL
4	1	115
5	0	NULL

■

1	3	110
2	1	100
4	1	115

D

1	5	110
2	5	100
4	5	115

E

1	2	110
2	1	100
4	0	115

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

```
SELECT C1, C2
FROM T1 AS M
WHERE NOT EXISTS (SELECT C1 FROM T1 WHERE C2 < M.C2)
```

A

2	110
3	120
5	115

B The SQL is invalid.

C Empty result.

D

3	120

■

1	100
4	100

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

```
SELECT T2.C2, COUNT(T1.C1)
FROM T1
RIGHT JOIN T2 ON T1.C3=T2.C1
GROUP BY T2.C2
```

■

A	3
B	1
C	0
D	1
E	0

B

A	3
B	1
D	1

C

A	2
B	1
D	NULL

D

A	2
B	1
D	0

E

A	3
B	1
C	NULL
D	1
E	NULL

Question 16   In SQLite you will get the following error when you try to execute the query from the previous question:

```
Error:  RIGHT and FULL OUTER JOINs are not currently supported
```

What query can you run instead to get the same result?

■
```
SELECT T2.C2, COUNT(T1.C1)
FROM T2
LEFT JOIN T1 ON T1.C3=T2.C1
GROUP BY T2.C2
```

B
```
SELECT T2.C2, COUNT(T1.C1)
FROM T1, T2
WHERE T1.C3=T2.C1 OR COUNT(T1.C1) IS NULL
GROUP BY T2.C2
```

C It is not possible to use a single query to get the same result.

D
```
SELECT T2.C2, COUNT(T1.C1)
FROM T2
INNER JOIN T1 ON T1.C3=T2.C1
WHERE COUNT(T1.C1) >= 0
GROUP BY T2.C2
```

E
```
SELECT T2.C2, COUNT(T1.C1)
FROM T1
INNER JOIN T2 ON T1.C3=T2.C1 OR T2.C1 IS NULL
GROUP BY T2.C2
```

Question 17   What is the result of the following SQL query? (Not showing the header of the result.)

```
SELECT A.C3, B.C3
FROM T1 AS A
LEFT JOIN T1 AS B ON A.C4=B.C1
```

A The SQL is invalid since a table cannot be joined with itself!

B Empty result.

D
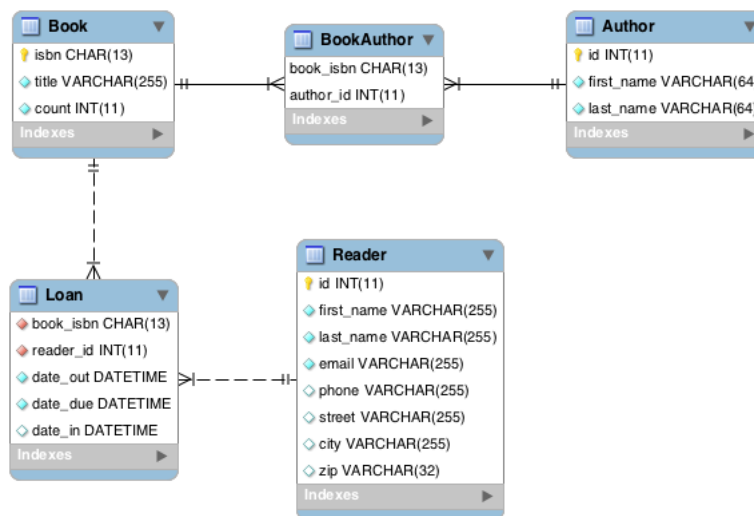
1	4
1	1
2	1

■

1	4
1	NULL
1	1
2	1
4	NULL

E

NULL	1
1	1
2	1
NULL	1
NULL	2
1	4

For questions 18 – 20 consider the following simplified data model for books, readers and loans in a library:

Notes:

- The count attribute in the Book table is the number of copies of the book owned by the library.

- When a book is checked out (i.e., a reader borrows a book), a new row is inserted in the Loan table with date_out set to the current date and time, the date_due set to 1 month from the current date and time, and date_in set to NULL. The reader can extend the loan period, in that case the date_due is updated to 1 month from the current date and time. When the book is checked in (i.e., returned to the library), the date_in gets updated to the current date and time (the row is never removed from the Loan table).

Question 18 ♣ Finish the query to select the ISBN and title of books with more than one author. Include the number of authors as well.

```
SELECT Book.isbn, Book.title, ① AS author_count
FROM Book
JOIN ②
③
```

Select one choice for ①, one choice for ② and one choice for ③.

| A | ① COUNT(DISTINCT Book.isbn) |

| ■ | ① COUNT(*) |

| C | ① SUM(Author) |

| ■ | ② BookAuthor ON BookAuthor.book_isbn=Book.isbn |

| E | ② Author ON BookAuthor.book_isbn=Book.isbn |
| | AND BookAuthor.author_id=Author.id |

| F | ③ WHERE author_count > 1 |
| | GROUP BY Book.isbn, Book.title |

| ■ | ③ GROUP BY Book.isbn, Book.title |
| | HAVING author_count > 1 |

| H | ③ WHERE author_count IS NOT 1 |
| | GROUP BY Book.isbn, Book.title |

Question 19 ♣ Finish the SQL statement to create the BookAuthor table:

```
CREATE TABLE BookAuthor (
   book_isbn char(13) NOT NULL,
   author_id int(11) NOT NULL,
   ①
)
```

In the following list select **all** relevant primary and foreign key constraints for the BookAuthor table that to-gether (separated by commas) will replace ①.

■	FOREIGN KEY (author_id) REFERENCES author(id)
■	PRIMARY KEY (book_isbn, author_id)
C	FOREIGN KEY book(isbn) REFERENCES book_isbn
D	PRIMARY KEY (author_id)
■	FOREIGN KEY (book_isbn) REFERENCES book(isbn)
F	PRIMARY KEY (book_isbn)
G	FOREIGN KEY author(id) REFERENCES author_id

Question 20 ♣   Finish the query to select the loans that are overdue (i.e., the loans not returned by the due date).

```
SELECT Book.isbn, Reader.*
FROM Loan, Book, Reader
WHERE ①
```

Select **all** parts of ① which, when joined with the AND operator, create the correct WHERE clause.
Note: The `NOW()` function returns the current date and time.

- ■ `date_in IS NULL`
- B `Book.count > 0`
- C `date_in > NOW()`
- D `date_in = NULL`
- ■ `Loan.reader_id = Reader.id`
- F `date_due > NOW()`
- ■ `Loan.book_isbn = Book.isbn`
- H `COUNT(Loan.reader_id) > 0`
- ■ `date_due < NOW()`
- J `date_due IS NULL`

4   FDs, CKs, NFs and normalization

For questions 21 – 25 consider a relation R(A, B, C, D, E) in 1NF with the following dependencies:

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

Question 21 ♣   Select **all** elements of $\{A, B\}^+$ (i. e. the set of all attributes which can be determined by the set $\{A, B\}$)?

- ■ A
- ■ B
- ■ C
- ■ D
- ■ E

Question 22 ♣   Select **all** candidate keys of R.

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

Question 23 ♣   Select **all** **non-prime** attributes.

[A] A

[B] B

■ C

■ D

■ E

Question 24   In which normal form (NF) is R?

[A] 2NF but not 3NF

[B] BCNF but not 3NF

■ 1NF but not 2NF

[D] BCNF

[E] 3NF but not BCNF

Question 25   Which of the following options shows the result of the normalization of R with all tables in BCNF (not showing the primary and foreign key constraints)?

[A] None of the other choices

[B] R_1(A, B, C, D, E)

■ R_1(A, B, C), R_2(B, E), R_3(E, D)

[D] R_1(A, B, C), R_2(B, D, E), R_3(E, D)

[E] R_1(A, B, C, E), R_2(E, D)

For questions 26 – 28 consider the following relation in 1NF:

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

with the following dependencies:

- {Date, SSN, TestID} → {Value}
- {SSN} → {FirstName, LastName}
- {TestID} → {TestName, Unit}

Question 26 ♣   Select **all** choices that are **superkeys** of the TestResult relation.

[A] {TestID, TestName, Unit}

■ {Date, SSN, FirstName, LastName, TestID, TestName, Value, Unit}

■ {Date, SSN, TestID}

[D] {SSN, FirstName, LastName}

■ {Date, SSN, TestID, Unit}

Question 27 In which NF is the TestResult relation?

A. 2NF but not in 3NF

B. 3NF but not in 2NF

C. 3NF but not in BCNF

D. BCNF

■ 1NF but not in 2NF

Question 28 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, SSN, TestName, Value)
Person(SSN, FirstName, LastName)
Test(TestName, TestID, Unit)

B. TestResult(Date, Value)
Person(SSN, FirstName, LastName)
Test(TestID, TestName, Unit)

C. TestResult(Date, SSN, TestID, TestName, Value, Unit)
Person(SSN, FirstName, LastName)

■ TestResult(Date, SSN, TestID, Value)
Person(SSN, FirstName, LastName)
Test(TestID, TestName, Unit)

E. TestResult(Date, SSN, TestID)
Person(SSN, FirstName, LastName)
Test(TestID, TestName, Value, Unit)

Question 29 In which NF is the BookAuthor table that was introduced before Question 18?

A. 1NF but not 2NF

B. 2NF but not 3NF

C. 3NF but not BCNF

■ BCNF

E. The table has no CK so it cannot be in any NF

Question 30   The following table shows the current state of a relation:

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

Is the following dependency true?

$$\{C_3\} \rightarrow \{C_2\}$$

- [A] No, the values of C_3 are not unique.
- [B] Yes, since all rows are distinct.
- [C] Yes, since if C_2 is 1 or 2, then C_3 is XX and if C_2 is 3 or 4, C_3 is XY.
- [D] A dependency cannot be confirmed or rejected by just checking one state of the relation.
- [■] No, since for some values of C_3 there are several different values of C_2.

5   Other

Question 31 ♣   In the following list, select **all** ACID properties.

- [A] Independence
- [■] Durability
- [C] Dependability
- [D] Concurrency
- [■] Isolation
- [■] Consistency
- [■] Atomicity
- [H] Aggregation

Question 32 ♣   What privileges are needed to be able to execute the following query?

```
DELETE FROM project
WHERE id NOT IN (SELECT project_id FROM employee_project)
```

Select **all** relevant privileges.
Hint: There are 3 correct choices.

- [A] DELETE ON employee_project
- [■] DELETE ON project
- [C] SELECT(project_id) ON project
- [D] DELETE(id) ON project
- [■] SELECT(id) ON project
- [■] SELECT(project_id) ON employee_project

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='X'
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?

- ■ X: 2500, Y: 1000
- B X: 1500: Y: 2000
- C X: 1500 or 2500 depending on if the server managed to save the updated balance back before the crash, Y: 1000
- D None of the other answers
- E X: 1500, Y: 1000

Question 34   Consider the following relation R:

C1	C2
1	110
2	140
3	120
4	130

We execute the following statements:

```
CREATE TABLE X AS SELECT AVG(C2) FROM R;
CREATE VIEW Y AS SELECT AVG(C2) FROM R;
UPDATE R SET C2=C2*2;
```

What are the results of the following queries (not showing headers)?
① SELECT * FROM X
② SELECT * FROM Y

- ■ ① 125   ② 250
- B CREATE VIEW is not a valid SQL statement.
- C ① 250   ② 250
- D ① 250   ② 125
- E ① 125   ② 125

Question 35   In the relational model, if an attribute K is a candidate key of a relation R and X is an attribute of R different from K then:

- A X cannot be the primary key of R.
- B K is also the primary key of R.
- C None of the other options.
- D {K, X} is also a candidate key.
- E X cannot be a candidate key.
- ■ {K, X} is always a superkey (even if X is not a candidate key).

Question 36   Consider a relation R(A, B, C, D, E) containing 10^7 records. A is the primary key, and B and C each contains 10^5 distinct values. The following prepared SQL is executed very often:

```
SELECT A FROM R WHERE B=? AND C=?
```

Which one of the following indexes would you create?

- ■ B
- B A
- ■ B, C
- D A, B, C
- E A, B
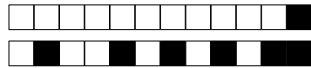
Final exam in 1DL301 Database Design I – Answer sheet

Department of Information Technology, Uppsala University

January 11, 2019, 14.00 – 19.00

Your exam code

. . . .

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!

0:☐	0:☐	0:☐	0:☐
1:☐	1:☐	1:☐	1:☐
2:☐	2:☐	2:☐	2:☐
3:☐	3:☐	3:☐	3:☐
4:☐	4:☐	4:☐	4:☐
5:☐	5:☐	5:☐	5:☐
6:☐	6:☐	6:☐	6:☐
7:☐	7:☐	7:☐	7:☐
8:☐	8:☐	8:☐	8:☐
9:☐	9:☐	9:☐	9:☐

i A:■ B:■ C:■ D:■ E:■

ii A:■ B:■ C:■

iii A:■ B:■ C:■ D:■ E:■

1 A:☐ B:☐ C:■ D:☐ E:☐

2 A:☐ B:☐ C:☐ D:☐ E:■ F:■
   G:■

3 A:■ B:☐ C:☐ D:☐

4 A:☐ B:☐ C:■ D:☐

5 A:☐ B:☐ C:☐ D:■

6 A:☐ B:☐ C:■ D:☐

7 A:☐ B:☐ C:☐ D:☐ E:■

8 A:☐ B:■ C:☐ D:☐

9 A:☐ B:■ C:☐ D:☐ E:☐

10 A:☐ B:■ C:■ D:☐ E:☐

11 A:☐ B:☐ C:☐ D:■ E:☐

12 A:☐ B:☐ C:☐ D:☐ E:■

13 A:☐ B:☐ C:■ D:☐ E:☐

14 A:☐ B:☐ C:☐ D:☐ E:■

15 A:■ B:☐ C:☐ D:☐ E:☐

16 A:■ B:☐ C:☐ D:☐ E:☐

17 A:☐ B:☐ C:■ D:☐ E:☐

18 A:☐ B:■ C:☐ D:■ E:☐ F:☐

   G:■ H:☐

19 A:■ B:■ C:☐ D:☐ E:■ F:☐
   G:☐

20 A:■ B:☐ C:☐ D:☐ E:■ F:☐
   G:■ H:☐ I:■ J:☐

21 A:■ B:■ C:■ D:■ E:■

22 A:☐ B:☐ C:☐ D:☐ E:■

23 A:☐ B:☐ C:■ D:■ E:■

24 A:☐ B:☐ C:■ D:☐ E:☐

25 A:☐ B:☐ C:■ D:☐ E:☐

26 A:☐ B:■ C:■ D:☐ E:■

27 A:☐ B:☐ C:☐ D:☐ E:■

28 A:☐ B:☐ C:☐ D:■ E:☐

29 A:☐ B:☐ C:☐ D:■ E:☐

30 A:☐ B:☐ C:☐ D:☐ E:■

31 A:☐ B:■ C:☐ D:☐ E:■ F:■
   G:■ H:☐

32 A:☐ B:■ C:☐ D:☐ E:■ F:■

33 A:■ B:☐ C:☐ D:☐ E:☐

34 A:■ B:☐ C:☐ D:☐ E:☐

35 A:☐ B:☐ C:☐ D:☐ E:☐ F:■

36 A:■ B:☐ C:■ D:☐ E:☐