NATIONAL UNIVERSITY OF SINGAPORE

SCHOOL OF COMPUTING
FINAL ASSESSMENT FOR
Semester 1 AY2024/2025

IT5008 Database Design and Programming

November 2024 Time Allowed 120 minutes

INSTRUCTIONS TO CANDIDATES

1. This assessment paper contains **23** questions and comprises **8** printed pages, including this page.
2. More specific instructions *may* be given at the beginning of each section. Please read them carefully.
3. Write all your answers in the answer sheet provided. Please read the additional instruction for multiple choice (MCQ) and multiple response (MRQ) questions.
   - Any answer not in the space provided will not be graded.
   - For MCQ and MRQ, **shade your answer in the corresponding bubble on the answer sheet**.
   - For MCQ, if multiple answers are equally appropriate, pick one and shade **ONLY** the chosen answer on the answer sheet. Do **NOT** shade more than one answer.
   - For MRQ, shade **ALL** correct answers. Partial marks *may* be given, but not guaranteed.
   - If there are no correct answer or no appropriate choices, shade X in the answer sheet. No partial mark given if X is chosen.
4. Shade and write your student number in the answer sheet. Do **NOT** write your name.
5. The total marks for this assessment is 80. Answer **ALL** questions.
6. This is a **CLOSED-BOOK** assessment. You are only allowed to refer to one double-sided A4-size paper.
7. All SQL query in this assessment paper are run on PostgreSQL 16.
8. We will use the monospace `NULL` to represent NULL-values. We will represent string with single-quote `'string'`.
9. The last page is left blank. You may use them for your own work.

Question	Points
1 - 10	30
11 - 20	30
21 - 22	8
23	12
TOTAL	80

Functional Dependencies and Normal Forms

For the next **ten (10)** questions, we will be using the following relation and set of functional dependencies.

$$R = \{A, B, C, D, E\}$$

$$\Sigma = \{\ \{C, E\} \rightarrow \{A, B\}, \{A, B, C\} \rightarrow \{A, D\}, \{D, E\} \rightarrow \{A, B, E\}, \{D\} \rightarrow \{A, C\},$$
$$\{A, C\} \rightarrow \{A, B\}, \{A, D\} \rightarrow \{B\}, \{C, D\} \rightarrow \{A, B\}\}$$

Notes: We start with the following computation:

- **Keys:** $\{\{C, E\}, \{D, E\}\}$
- **Minimal Cover:** $\{\{A, C\} \rightarrow \{B\}, \{A, C\} \rightarrow \{D\}, \{C, E\} \rightarrow \{A\}, \{D\} \rightarrow \{A\}, \{D\} \rightarrow \{C\}\}$

1. (3 points) **(MRQ)** Select **ALL** the attributes in the *closure* of $\{A, C\}$ (i.e., $\{A, C\}^+$).

   A. A B. B C. C D. D E. E

   Select X if there is no answer.

   Notes: ABCD

2. (3 points) **(MRQ)** Select **ALL** the functional dependencies that are *logically entailed* by Σ.

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

   Select X if there is no answer.

   Notes: D

3. (3 points) **(MRQ)** Select **ALL** the *superkeys* of R with Σ from the choice below.

   A. $\{A, B, C\}$
   B. $\{B, C, D\}$
   C. $\{C, D, E\}$
   D. $\{A, D, E\}$
   E. $\{A, B, E\}$

   Select X if there is no answer.

   Notes: CD: $\{C, D, E\}$ and $\{A, D, E\}$ as they are the superset of keys (see below).

4. (3 points) **(MRQ)** Select **ALL** the *candidate keys* of R with Σ from the choice below.
   A. $\{A, B\}$
   B. $\{A, E\}$
   C. $\{B, E\}$
   D. $\{C, E\}$
   E. $\{D, E\}$

   Select X if there is no answer.

   Notes: DE: $\{C, E\}$ and $\{D, E\}$ only. This should be the starting point of your working.

5. (3 points) **(MRQ)** Select **ALL** the *prime attributes* or R with Σ.

   A. A B. B C. C D. D E. E

   Select X if there is no answer.

   Notes: CDE: see above. The answer is $\{C, E\} \cup \{D, E\} = \{C, D, E\}$.

6. (3 points) **(MRQ)** Select **ALL** the functional dependencies below that are in the *projection* of R into $R_1 = \{B, D, E\}$.
   A. $\{B\} \rightarrow \{D\}$
   B. $\{D, E\} \rightarrow \{B\}$
   C. $\{D\} \rightarrow \{B\}$
   D. $\{E\} \rightarrow \{B\}$
   E. $\{E\} \rightarrow \{D\}$

   Select X if there is no answer.

   Notes: BC: $\{D, E\} \rightarrow \{B\}$ and $\{D\} \rightarrow \{B\}$. We need $\{D, E\} \rightarrow \{B\}$ as it is in Σ^+ (i.e., the functional dependency closure of Σ). We do not often use it in further computation as it is logically entailed by $\{D\} \rightarrow \{B\}$. But this follows from the definition used for projection.

   Assignment 2 (for 2024-2025 Sem 1) also tested on this, so hopefully you are familiar with this. That is partly why we release assignment 2 and the answer.

7. (3 points) **(MRQ)** Select **ALL** the functional dependencies below that violates the BCNF property of R with Σ.
   A. $\{A, B, C\} \rightarrow \{D\}$
   B. $\{D, E\} \rightarrow \{B\}$
   C. $\{D\} \rightarrow \{C\}$
   D. $\{A, C\} \rightarrow \{A\}$
   E. $\{C, D\} \rightarrow \{B\}$

   Select X if there is no answer.

   Notes: ACE: Here we select all LHS that are not superset of keys (see above) **AND non-trivial.** Note that we should also check if the functional dependencies are logically entailed by Σ. But a quick look shows that they are. We simply make the RHS singular from some of the FD in Σ.

8. (3 points) **(MRQ)** Select **ALL** the functional dependencies below that violates the 3NF property of R with Σ.
   A. $\{A, B, C\} \rightarrow \{D\}$
   B. $\{D, E\} \rightarrow \{B\}$
   C. $\{D\} \rightarrow \{C\}$
   D. $\{A, C\} \rightarrow \{A\}$
   E. $\{C, D\} \rightarrow \{B\}$

   Select X if there is no answer.

> **Notes:** E: Similar as above but we further exclude cases where the RHS is a prime attributes (i.e., not C, D, or E).

9. (3 points) Find one lossless-join decomposition of R in BCNF using the BCNF decomposition algorithm introduced in the course.

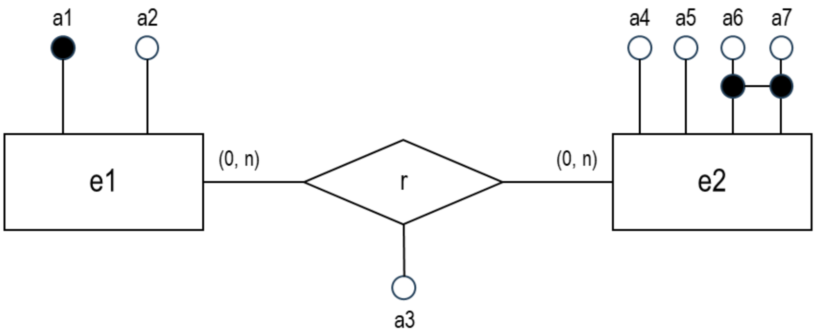
> **Notes:** $R_1 = A, B, C, D$ and $R_2 = A, C, E$ (i.e., $\{\{A, B, C, D\}, \{A, C, E\}\}$).
>
> This is just one possible answer. This can be obtained via the following steps:
>
> - $\{A, C\} \rightarrow \{B, D\}$ (from Question 1, so you do not have to do much recomputation) violates BCNF property of R with Σ. We decompose into:
>   - $R_1 = \{A, B, C, D\}$ from Question 1, we know $\{A, C\}^+ = \{A, B, C, D\}$.
>   - With $\Sigma_1 = \{\{A, C\} \rightarrow \{B, D\}, \{D\} \rightarrow \{A, B, C\}\}$.
>   - R_1 is in BCNF.
>   - $R_2 = R - \{A, C\}^+ \cup \{A, C\} = \{A, C, E\}$.
>   - With $\Sigma_2 = \{\{C, E\} \rightarrow \{A\}\}$.
>   - R_2 is in BCNF.

10. (3 points) Find one dependency preserving lossless-join decomposition of R in 3NF using the 3NF synthesis algorithm introduced in the course.

> **Notes:** $R_1 = A, B, C, D$ and $R_2 = A, C, E$ (i.e., $\{\{A, B, C, D\}, \{A, C, E\}\}$).
>
> There are other minimal cover. Using the one we computed before, we get the following.
>
> - **Minimal Cover:** $\{\{A, C\} \rightarrow \{B\}, \{A, C\} \rightarrow \{D\}, \{C, E\} \rightarrow \{A\}, \{D\} \rightarrow \{A\}, \{D\} \rightarrow \{C\}\}$
> - **Canonical Cover:** $\{\{A, C\} \rightarrow \{B, D\}, \{C, E\} \rightarrow \{A\}, \{D\} \rightarrow \{A, C\}\}$
> - **Construction:** $\{\{A, B, C, D\}, \{A, C, E\}, \{A, C, D\}\}$
> - **Subsumption:** $\{\{A, B, C, D\}, \{A, C, E\}\}$ *(because $\{A, C, D\} \subset \{A, B, C, D\}$)*
> - **Adding Keys:** $\{\{A, B, C, D\}, \{A, C, E\}\}$ *(because one key $\{C, E\} \subset \{A, C, E\}$)*
>
> We accept alternative answers as long as many conditions are satisfied. Also, the minimal cover can be more complex which leads to more potential solution. We ignore subsumption as we also mentioned that there are some *exceptions* although we verbally suggest to simply do subsumption.

Schema Translation

For the next **two (2)** questions, we will be using the following relation entity-relationship diagram.



During the translation, if t_1 and t_2 are entity/relationship sets, if the table for t_1 and t_2 are merged, if the statement is correct for table t_1, then the same statement is true for table t_2.

Notes:

```
CREATE TABLE e1 (
  a1  INT  PRIMARY KEY,
  a2  INT
);

CREATE TABLE e2 (
  a4  INT,
  a5  INT,
  a6  INT,
  a7  INT,
  PRIMARY KEY (a6, a7)
);

CREATE TABLE r (
  a1  INT  REFERENCES e1(a1),
  a3  INT,
  a6  INT,
  a7  INT,
  PRIMARY KEY (a1 a6, a7),
  FOREIGN KEY (a6, a7) REFERENCES e2(a6, a7)
);
```

This follows our 3 translation rules and 3 exceptional rules.

11. (3 points) **(MRQ)** Select **ALL** the correct statement.

A. A possible primary key of the table for r is a_3.

B. A possible primary key of the table for r is a_1.

C. A possible primary key of the table for r is (a_6, a_7).

D. A possible primary key of the table for r is (a_1, a_6, a_7).

E. A possible primary key of the table for r is (a_1, a_3, a_6, a_7).

Select X if there is no answer.

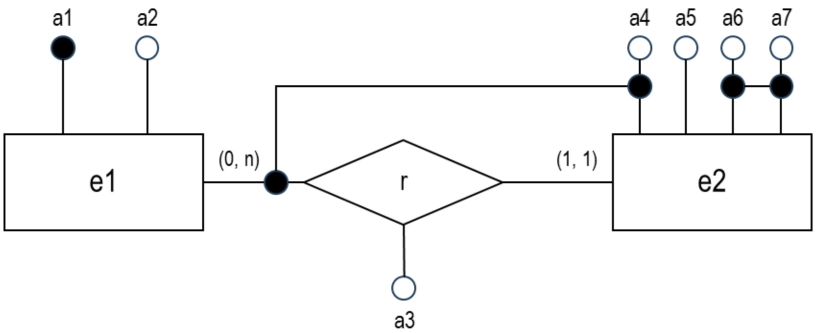
Notes: D

12. (3 points) **(MRQ)** Select **ALL** the correct statement.
    - A. a_1 is the foreign key in the table r referencing a_1 in table e_1.
    - B. a_1 is the foreign key in the table e_1 referencing a_1 in table r.
    - C. a_1 is the foreign key in the table e_2 referencing a_1 in table e_1.
    - D. (a_6, a_7) is the foreign key in the table r referencing (a_6, a_7) in table e_2.
    - E. (a_6, a_7) is the foreign key in the table e_2 referencing (a_6, a_7) in table r.

    Select X if there is no answer.

Notes: AD

For the next **two (2)** questions, we will be using the following relation entity-relationship diagram.



During the translation, if t_1 and t_2 are entity/relationship sets, if the table for t_1 and t_2 are merged, if the statement is correct for table t_1, then the same statement is true for table t_2.

Notes:

```
CREATE TABLE e1 (
  a1  INT  PRIMARY KEY,
  a2  INT
);

CREATE TABLE e2 (  /* merged for both r and e2 */
  a1  INT  NOT NULL  REFERENCES e1(a1),
  a4  INT  NOT NULL,
  a5  INT,
  a6  INT,
  a7  INT,
  PRIMARY KEY (a6, a7),
  UNIQUE (a1, a4)
);
```

This follows our 3 translation rules and 3 exceptional rules.

13. (3 points) **(MRQ)** Select **ALL** the correct statement.
    - A. The table e_1 is merged with the table r.
    - B. The table e_2 is merged with the table r.
    - C. The table e_1 is merged with the table e_2.
    - D. The tables e_1, e_2 and r are all merged.
    - E. No tables are merged.

Select X if there is no answer.

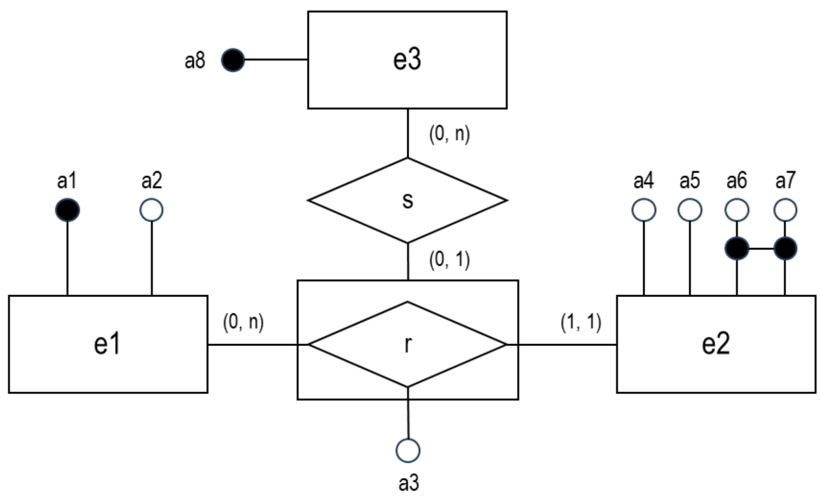
Notes: B

14. (3 points) **(MRQ)** Select **ALL** the correct statement.
    A. A possible primary key of the table r is a_1.
    B. A possible primary key of the table r is a_4.
    C. A possible primary key of the table r is (a_1, a_4).
    D. A possible primary key of the table r is (a_6, a_7).
    E. A possible primary key of the table r is (a_1, a_6, a_7).

Select X if there is no answer.

Notes: CD

For the next **three (3)** questions, we will be using the following relation entity-relationship diagram.



During the translation, if t_1 and t_2 are entity/relationship sets, if the table for t_1 and t_2 are merged, if the statement is correct for table t_1, then the same statement is true for table t_2.

Notes:

```
CREATE TABLE e1 (
  a1  INT  PRIMARY KEY,
  a2  INT
);

CREATE TABLE e2 (  /* merged for both r and e2 */
  a1  INT  NOT NULL  REFERENCES e1(a1),
  a3  INT,
  a4  INT,
  a5  INT,
  a6  INT,
  a7  INT,
  PRIMARY KEY (a6, a7)
);

CREATE TABLE e3 (
  a8  INT  PRIMARY KEY
);

CREATE TABLE s (  /* not merged as it is (0, 1) */
  a6  INT,
  a7  INT,
  a8  INT  NOT NULL  REFERENCES e3(a8),
  PRIMARY KEY (a6, a7),  /* no a8 because (0, 1) */
  FOREIGN KEY (a6, a7) REFERENCES e2(a6, a7)
);
```

This follows our 3 translation rules and 3 exceptional rules.

15. (3 points) **(MCQ)** How many tables are created.

    A. 6

    B. 5

    C. 4

    D. 3

    E. 2.

    Select X if there is no answer.

    Notes: C

16. (3 points) **(MRQ)** Select **ALL** the correct statement.

    A. The table r and e_1 are merged.

    B. The table r and e_2 are merged.

    C. The table r and s are merged.

    D. The table s and e_3 are merged.

    E. No tables are merged.

    Select X if there is no answer.

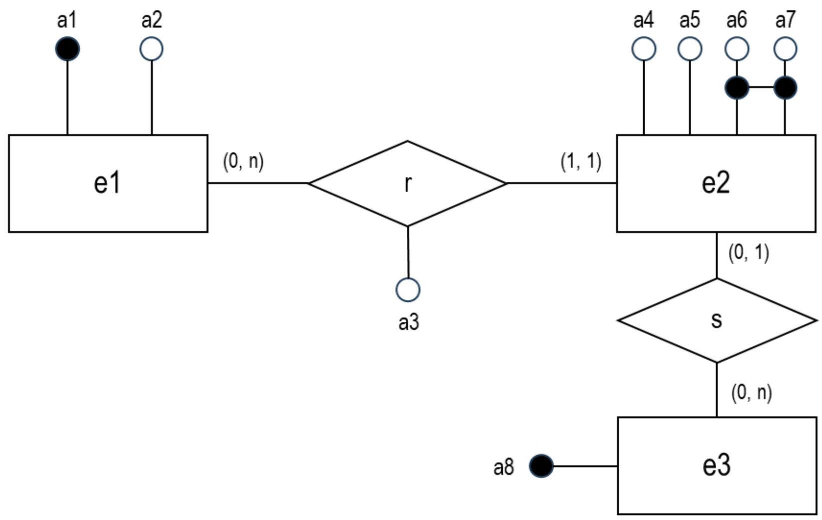
    Notes: B

17. (3 points) **(MRQ)** Select **ALL** the correct statement.
    - A. A possible primary key of the table s is a_1.
    - B. A possible primary key of the table s is (a_6, a_7).
    - C. A possible primary key of the table s is (a_1, a_6, a_7).
    - D. A possible primary key of the table s is (a_6, a_7, a_8).
    - E. A possible primary key of the table s is (a_1, a_6, a_7, a_8).

    Select X if there is no answer.

    Notes: B

For the next **three (3)** questions, we will be using the following relation entity-relationship diagram.



During the translation, if t_1 and t_2 are entity/relationship sets, if the table for t_1 and t_2 are merged, if the statement is correct for table t_1, then the same statement is true for table t_2.

Notes:

```
CREATE TABLE e1 (
  a1  INT  PRIMARY KEY,
  a2  INT
);

CREATE TABLE e2 (  /* merged for both r and e2 */
  a1  INT  NOT NULL  REFERENCES e1(a1),
  a3  INT,
  a4  INT,
  a5  INT,
  a6  INT,
  a7  INT,
  PRIMARY KEY (a6, a7)
);

CREATE TABLE e3 (
  a8  INT  PRIMARY KEY
);

CREATE TABLE s (  /* not merged as it is (0, 1) */
  a6  INT,
  a7  INT,
  a8  INT  NOT NULL  REFERENCES e3(a8),
  PRIMARY KEY (a6, a7),  /* no a8 because (0, 1) */
  FOREIGN KEY (a6, a7) REFERENCES e2(a6, a7)
);
```

This follows our 3 translation rules and 3 exceptional rules.

18. (3 points) **(MCQ)** How many tables are created.

    A. 6

    B. 5

    C. 4

    D. 3

    E. 2.

    Select X if there is no answer.

    Notes: C

19. (3 points) **(MRQ)** Select **ALL** the correct statement.

    A. A possible primary key of the table s is (a_1, a_6, a_7).

    B. A possible primary key of the table s is (a_3, a_6, a_7).

    C. A possible primary key of the table s is (a_6, a_7, a_8).

    D. A possible primary key of the table s is (a_1, a_3, a_6, a_7).

    E. A possible primary key of the table s is $(a_1, a_3, a_6, a_7, a_8)$.

    Select X if there is no answer.

    Notes: X

20. (3 points) **(MRQ)** Select **ALL** the correct statement.
    A. A possible primary key of the table e_1 is a_1.
    B. A possible primary key of the table e_2 is (a_6, a_7).
    C. A possible primary key of the table e_3 is a_8.
    D. A possible primary key of the table r is a_1.
    E. A possible primary key of the table r is (a_1, a_6, a_7).

    Select X if there is no answer.

    Notes: ABC

Relational Algebra

For the next **two (2)** questions, we will consider the following schema.

- `students(matric, sname)` with `matric` as the primary key.
- `workings(pid, matric, since)` with `(pid, matric)` as the primary key, with `workings.pid` referencing `projects.pid`, and `workings.matric` referencing `students.matric`.
- `projects(pid, pname)` with `pid` as the primary key.
- `kinds(pid, cname)` with `(pid, cname)` as the primary key and with `kinds.pid` referencing `projects.pid`.

Note that you may use the dot notation in this relational algebra.

21. (3 points) Write a relational algebra expression to find the different pairs of students (i.e., the matric numbers (`matric1`, `matric2`)) that are working on the same project. Note that since the same student s is always working on the same project as himself/herself, we also exclude the pair (s, s).

    Notes: $\pi_{[W_1.matric, W_2.matric]}($
    $\sigma_{[(W_1.pid = W_2.pid) \wedge (W_1.matric \neq W_2.matric)]}(\rho_{[W_1]}(Workings) \times \rho_{[W_2]}(Workings))$
    $)$

    We allow inclusion of `students` but it is not necessary as students are identified by `matric`. This can be deduced from the primary key of `students` which is `matric` and not `sname`.

22. (5 points) Write a relational algebra expression to find the oldest project id (i.e., `pid`) in the database. The oldest project is the project with the smallest `since` value. Assume you can compare `since` using the typical relational operator <, >, <=, >=, =, or `!=`.

    Note that there may be more than one oldest project. Your relational algebra expression should produce all.

    Notes: $Q_1 := \pi_{[W_2.pid, W_2.since]}(\sigma_{[(W_2.since > W_1.since)]}(\rho_{[W_1]}(Workings) \times \rho_{[W_2]}(Workings)))$
    $\pi_{[W_3.pid]}(\pi_{[W_3.pid, W_3.since]}(\rho_{[W_3]}(Workings)) - Q_1)$

    We want to find the project W with the *smallest* value of `since`. The idea is to find all projects W_2 that has larger `since` than at least one project W_1. We then exclude this W_2 from the answer via $-$.

    The difficulty lies in the need for projecting the correct attributes so that set difference via $-$ can be done properly.

Miscellaneous

23. (12 points) Consider the relation and set of functional dependencies.

$$R = \{A, B, C, D\}$$
$$\Sigma = \{\ \{A\} \rightarrow \{B\}, \{B\} \rightarrow \{C\}, \{C\} \rightarrow \{A\}, \{D\} \rightarrow \{A\}\ \}$$

(a) (4 points) Using only 3NF synthesis algorithm introduced in the lecture starting from Σ when computing the minimal cover, produce one lossless-join dependency preserving decomposition in 3NF of R with Σ.

> **Notes:** $\{\{A, B\}, \{B, C\}, \{A, C\}, \{A, D\}\}$

(b) (4 points) Is there a lossless-join dependency preserving decomposition in 3NF of R with Σ such that all of the following is satisfied?

- The number of fragments is fewer than your answer in part (a) above.
- Each fragment is in 3NF.

If there is, provide such decomposition (i.e., provide the fragments). Otherwise, briefly explain why those are not possible.

> **Notes:** $\{\{A, B, C\}, \{A, D\}\}$

(c) (4 points) Provide a `CREATE TABLE` statement to enforce the functional dependencies for your answer in part (b) if your answer in part (b) is different from part (a). Otherwise, provide a `CREATE TABLE` statement to enforce the functional dependencies for your answer in part (a).

Notes: The type does not really matter. We are only concerned about the constraints.

```
CREATE TABLE R1 (
  A  INT PRIMARY KEY,
  B  INT UNIQUE NOT NULL,
  C  INT UNIQUE NOT NULL
);

CREATE TABLE R2 (
  A  INT  REFERENCES R1(A),
  D  INT  PRIMARY KEY
);
```

We need `A` to be the primary key of `R1` as we want our solution to be *portable* and we can only reference the primary keys. But `B` and `C` also need to be candidate keys so we use `UNIQUE` and `NOT NULL`.
Lastly, do not forget to add foreign key from `R2.A` to `R1.A`. Otherwise, we may add `A` that is not satisfying the constraints.

This page is intentionally left blank.

END OF PAPER