THE AUSTRALIAN NATIONAL UNIVERSITY

Second Semester Examination – November 2012

RELATIONAL DATABASES

(COMP2400/COMP6240)

Writing period: 3 hours duration
Study period: 15 minutes duration
Permitted materials: A4 paper (one sheet) with handwritten notes one side only

Instructions:

- This exam booklet contains 5 questions, totaling 65 marks.
- You need to answer all questions. Whenever you feel that some information is missing, add an assumption and make it explicit in your solution.
- All your answers must be written in the spaces provided in this booklet. You may be provided with scrap paper for working, but it must **not** be used to write final answers. There is additional space at the end of the booklet in case the spaces provided under questions are insufficient.
- Do not remove this booklet from the examination room.

Student Number	

Official use only:

Question	1	2	3	4	5
Mark					
Out of	17	8	15	18	7

Total: ________ /65

Question 1: SQL and the Relational Model [17 marks]

1. a General Concepts [4 marks]

1. a (i) [2 marks]

Explain the relationship of data independence with the ANSI/SPARC three level architecture.

Answer: Refer to the text book and lecture notes.

1. a (ii) [1 mark]

Which of the following statements are true for a relation?

(1) Each superkey is a candidate key.

(2) Each candidate key is a superkey.

(3) The primary key is a candidate key, but there may be a candidate key that is not a primary key.

Answer: (2) and (3)

1. a (iii) [1 mark]

Given the sets $A = \{Sue, Ali\}$, $B = \{white, black\}$ and $C = \{cat, dog\}$, what is the Cartesian product $A \times B \times C$?

Answer:

$A \times B \times C$ = {(Sue, white, cat)
(Sue, white, dog)
(Sue, black, cat)
(Sue, black, dog)
(Ali, white, cat)
(Ali, white, dog)
(Ali, black, cat)
(Ali, black, dog)}

1. b Writing SQL [4 marks]

Not relevant to the final examination this year

1. c SQL Evaluation [5 marks]

Not relevant to the final examination this year

1. d Integrity Constraints [4 marks]

1. d (i) [2 marks]

Suppose that the relation SUPERVISE was created as follows:

```
CREATE TABLE SUPERVISE (
    pssn INT REFERENCES PROFESSOR(ssn) ON DELETE NO ACTION,
    gid INT REFERENCES GRADUATE(gid) ON DELETE SET NULL,
    pid INT REFERENCES PROJECT(pid) ON DELETE CASCADE,
    PRIMARY KEY(pssn, gid, pid)
);
```

Which of the following statements are true, and which are false?

(a) If we delete a tuple from SUPERVISE, any tuples in PROJECT referred to by this tuple are also deleted.

(b) If we delete a tuple from GRADUATE, some tuples of SUPERVISE may have their values of attribute gid set to NULL.

(c) If we try to insert a tuple into PROFESSOR, with an ssn that does not exist in SUPERVISE, the operation is rejected.

(d) If we try to insert a tuple into SUPERVISE, with a gid that does not exist in GRADUATE, the operation is rejected.

Provide your answer in the following table.

Statements	(a)	(b)	(c)	(d)
True		√		√
False	√		√	

1. d (ii) [2 marks]

Consider the relation Book in Figure 1, which has the primary key $\{bid\}$ and the foreign key $[aid] \subseteq$ Author$[aid]$.

Book				
bid	title	language	date	aid
1	The Plague	French	1947	4
2	The Cat in the Hat	English	1957	2
3	The Hobbit	English	1937	1
4	The Lord of the Rings	English	1954	1

Author	
aid	name
1	J.R.R.Tolklen
2	Dr. Seuss
3	S.E.Hinton
4	Albert Camus

Figure 1: Relation Book and Author

- Write down an SQL statement to modify an existing tuple in Author which would yield a key integrity violation. The modification should not violate any other integrity constraints.

Writing SQL
is not covered
in the final exam.

  Answer:


```
UPDATE Author
SET aid = 2
WHERE name = "S.E.Hinton";
```


- Write down an SQL statement to insert a tuple into Book which would yield an entity integrity violation. The insertion should not violate the existing foreign key constraint.

  Answer:

```
INSERT INTO Book
VALUES (NULL, "Fire", English, 1980, 1);
```

Question 2: ER Modelling and Translation [8 marks]

2. a ER Modelling [4 marks]

Canberra Employment Centre (CEC) places temporary workers in companies during peak periods. CEC maintains a file of candidates who wish to work. If the candidate has worked before, that candidate has a specific job history. (Naturally, no job history exists if the candidate has never worked.)
Each candidate may have several qualifications. CEC also has a list of companies that request temporaries. Each time a company requests a temporary employee, CEC makes an entry in the openings folder. This folder contains an opening number, company name, required qualifications, starting date, anticipated ending date, and hourly pay. Each opening requires only one specific qualification.

Draw an ER diagram that captures the above information, which should include:

1. identifying the entities, relationships and their attributes;
2. indicating the key attributes which you have chosen.

Answer: See Figure 2, in which the following assumptions are made:

- One opening will match with one qualification only.
- A qualification may be good for a number of openings.
- A company may or may not have an opening.
- Qualification will be identified by QualificationID and will be written in terms of type such as Typing, IT, Management etc. One may choose a surrogate key for Qualification.

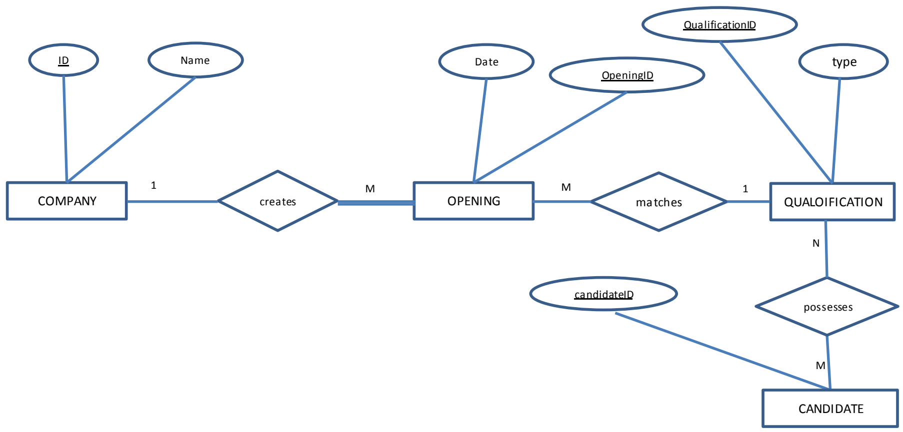

Figure 2: Answer for Q2.a

2.b ER-to-Relational Translation[4 marks]

The following ER diagram is drawn from a business case where a vendor provides a quotation for the supply of a part.

Transform the ER diagram to a relational database schema and identify the primary and foreign keys for each relation schema.

Answer:
Applying the translation rules in the lecture notes, we have:

- VENDOR(VendorID, Address) with the primary key {VendorID};
- PART(PartNo, Details) with the primary key {PartNo};
- QUOTES(VendorID, PartNo, QuoteQuantity, Price) with the primary key {VendorID, PartNo} and the foreign keys: [VendorID] $\subseteq$ VENDOR[VendorID] and [PartNo] $\subseteq$ PART[PartNo].

Question 3: Functional Dependencies and Normal Forms [15 marks]

3. a Satisfaction of Functional Dependencies [4 marks]

3. a (i) [3 marks]

Consider two relations $r_1(R)$ and $r_2(R)$ over the same relation schema $R(A, B, C, D)$.

$r_1(R)$

A	B	C	D
1	2	3	1
4	5	3	2
4	3	3	2
1	5	2	3

$r_2(R)$

A	B	C	D
1	2	3	2
1	4	5	3
3	4	2	4

The following is a table (i.e., Table 1) with a column for each of these relations and a row for a functional dependency. Enter "yes" or "no" in each cell of the table, indicating whether the relation satisfies the functional dependency.

Answer:

	$r_1(R)$	$r_2(R)$
$A \longrightarrow B$	no	no
$AB \longrightarrow C$	yes	yes
$A \longrightarrow BC$	no	no
$DC \longrightarrow B$	no	yes
$BC \longrightarrow B$	yes	yes
$AD \longrightarrow C$	yes	yes

Table 1: Functional dependencies

3. a (ii) [1 mark]

Are there any trivial functional dependencies shown in Table 1? If any, specify them and explain why they are trivial.

Answer: $BC \longrightarrow B$ is trivial.

3. b Candidate Keys and Normal Forms [4 marks]

Given a relation schema $R(A, B, C, D, E)$ with the following set Σ of functional dependencies:

$$\Sigma = \{A \longrightarrow C, CE \longrightarrow B, BC \longrightarrow AD \text{ and } D \longrightarrow E\}.$$

3. b (i) [1 mark]

Does $AB \longrightarrow E$ hold on any relation of R that satisfies Σ? If so, explain why; otherwise, give a counterexample.

Answer: Compute the closure of AB w.r.t. Σ: $(AB)^+$ = $(ABC)^+$ = $(ABCD)^+$ = $(ABCDE)^+$ =ABCDE. Because $E \in (AB)^+$ holds, $AB \longrightarrow E$ holds on any relation of R that satisfies Σ.

3. b (ii) [3 marks]

Is R in BCNF? If not, normalise R into BCNF. Explain your answer.

Answer:

- Step 1: check whether the left hand side of each FD is a superkey:

  - $(A)^+$ =AC
  - $(CE)^+$ = $(BCE)^+$ = $(ABCDE)^+$ =ABCDE
  - $(BC)^+$ = $(ABCD)^+$ = $(ABCDE)^+$ =ABCDE
  - $(D)^+$ = DE

- Step 2: $A \longrightarrow C$ and $D \longrightarrow E$ are problematic, so we decompose R along them into:
  - AC with $\{A \longrightarrow C\}$
  - DE with $\{D \longrightarrow E\}$
  - ABD

3. c Candidate Keys and Normal Forms [7 marks]

Consider the relation schema

MEETING(OfficerID, OfficerName, CustNo, CustName, Date, Time, Room),

and the following set of functional dependencies on MEETING:

- OfficerID $\longrightarrow$ OfficerName;
- OfficerID, Date $\longrightarrow$ Room;
- CustNo $\longrightarrow$ CustName;
- CustNo, Date, Time $\longrightarrow$ OfficerID;
- Date, Time, Room $\longrightarrow$ CustNo.

3. c (i) [1 mark]

Discuss the anomalies in the current schema MEETING and identify at least two potential problems.

Answer: Refer to the text book and the lecture notes about insert anomalies, delete anomalies and modification anomalies.

3. c (ii) [2 marks]

Find out all the candidate keys and prime attributes of MEETING.

Answer: Compute the closure of attributes (refer to the lecture nodes). The candidate keys are:

- {CustNo, Date, Time}
- {OfficerID, Date, Time}

- {Data, Time, Room}

The prime attributes are {CustNo, OfficeID, Date, Time, Room}.

3. c (iii) [1 mark]

As we have not discussed 1NF and 2NF in S2 2018, you can skip this question when preparing for the final exam.

What is the highest normal form of MEETING with respect to the given set of functional dependencies? Explain the reason.

Note:

- We only consider the normal forms 1NF, 2NF, 3NF and BCNF (in increasing order of strength).
- No primary keys are given, so the relevant definitions of the normal forms are the ones that refer to *all* candidate keys.

Answer: The highest normal form of MEETING is 1NF because OfficerID $\longrightarrow$ OfficerName and CustNo $\longrightarrow$ CustName are partial dependencies with respect to the candidate keys.

3. c (iv) [3 marks]

Normalise the relation schema MEETING into BCNF.

As we have not discussed 2NF in S2 2018, please ignore the sample solution to this question when preparing for the final exam.

Answer: There are several steps:

- Normalise MEETING into 2NF along OfficerID $\longrightarrow$ OfficerName and CustNo $\longrightarrow$ CustName:
  - OFFICE(OfficeID, OfficeName) with the FD: OfficerID $\longrightarrow$ OfficerName
  - CUSTOMER(CustNo, CustName) with the FD: CustNo $\longrightarrow$ CustName
  - MEETING'(OfficerID, CustNo, Date, Time, Room) with the FDs:
    * OfficerID, Date $\longrightarrow$ Room;
    * CustNo, Date, Time $\longrightarrow$ OfficerID;
    * Date, Time, Room $\longrightarrow$ CustNo.
- Normalise MEETING' into BCNF along OfficerID, Date $\longrightarrow$ Room:

  - MEETING''(OfficerID, Date, Room) with the FD: OfficerID, Date $\longrightarrow$ Room;
  - MEETING'''(OfficerID, CustNo, Date, Time) with the FD: CustNo, Date, Time $\longrightarrow$ OfficerID.

Hence, MEETING can be decomposed into the following four relations in BCNF:

- OFFICE, CUSTOMER, MEETING'' and MEETING'''

Question 4: Relational Algebra and Query Processing [18 marks]

4. a Relational Algebra Expressions [4 marks]

Consider the following relation schemas:

AUTHOR(aid, name) with the primary key {aid};

BOOK(bid, title, language, date, aid) with the primary key {bid} and the foreign key [aid]$\subseteq$ AUTHOR[aid].

Write relational algebra expressions for the following queries.

4. a (i) [1 mark]

Who wrote the book titled “The Cat in the Hat”?

Answer:

- $\pi_{name}(\sigma_{title=\text{“}The\ Cat\ in\ the\ Hat\text{”}}(\text{BOOK}) \bowtie \text{AUTHOR})$, or
- $\pi_{aid,\ name}(\sigma_{title=\text{“}The\ Cat\ in\ the\ Hat\text{”}}(\text{BOOK}) \bowtie \text{AUTHOR})$

4. a (ii) [1 mark]

List the names of authors who have published at least one book in English and one book in Japanese.

Answer:

- $\pi_{name}((\pi_{aid}(\sigma_{Language=\text{“}English\text{”}}(\text{BOOK})) \bowtie \pi_{aid}(\sigma_{Language=\text{“}Japanese\text{”}}(\text{BOOK}))) \bowtie \text{AUTHOR})$

- $\pi_{name}((\pi_{aid}(\sigma_{Language=\text{“}English\text{”}}(\text{BOOK})) \cap \pi_{aid}(\sigma_{Language=\text{“}Japanese\text{”}}(\text{BOOK}))) \bowtie \text{AUTHOR})$

4. a (iii) [2 marks]

Find out the authors who have never published a book in English.

Answer:

- $\pi_{aid,name}(\textsc{Author}) - \pi_{aid,name}(\sigma_{language=\text{``}English\text{''}}(\textsc{Book} \bowtie \textsc{Author}))$

4. b Evaluation [5 marks]

Suppose that we have the relations ANIMAL and COLOR shown in Figure 3.

ANIMAL

A	B	C
1	white	cat
2	brown	rabbit
3	white	bird
4	red	bird

COLOR

D	E
1	brown
2	white
3	blue

Figure 3: Relations ANIMAL and COLOR

Evaluate the following relational algebra expressions. Show your answer as a table, like those in Figure 3.

4. b (i) [1 mark]

Evaluate $\pi_C(\sigma_{B='white'}(\text{ANIMAL}))$.

Answer:

C				
cat				
bird				

4. b (ii) [1 mark]

Evaluate $\pi_B(\text{ANIMAL}) \cup \rho_{(B)}(\pi_E(\text{COLOR}))$.

Answer:

B				
white				
brown				
red				
blue				

4. b (iii) [1 mark]

Evaluate $\pi_{A,C,E}(\text{ANIMAL} \bowtie_{B=E} \text{COLOR})$.

Answer:

A	C	E		
1	cat	white		
2	rabbit	brown		
3	bird	white		

4. b (iv) [2 marks]

Evaluate $(\sigma_{B='white'}(\text{ANIMAL})) \times \pi_E(\text{COLOR})$

Answer:

A	B	C	E	
1	white	cat	brown	
1	white	cat	white	
1	white	cat	blue	
3	white	bird	brown	
3	white	bird	white	
3	white	bird	blue	

4. c Relational Algebra Operators [5 marks]

4. c (i) [1 mark]

List the six basic relational algebra operators that constitute a complete set in relational algebra.

Answer:

1. selection σ;
2. projection π;
3. renaming ρ;
4. union $\cup$;
5. difference $-$;
6. Cartesian product $\times$.

4. c (ii) [1 mark]

Define the operator *join* in terms of the six basic operators in relational algebra.

Answer:

- $R_1 \bowtie_\varphi R_2 = \sigma_\varphi(R_1 \times R_2)$

4. c (iii) [1 mark]

Suppose that two relations R and Q have exactly the same schema. Which of the following statements are true in relational algebra?

1. $R \cap Q = R - (R - Q)$
2. $R \cap Q = Q - (Q - R)$
3. $R \cap Q = R \times Q$
4. $R \cap Q = R \bowtie Q$

Answer:

- (1), (2) and (4)

4. c (iv) [2 marks]

Consider the following statements of relational algebra. Does each of them hold for any relation R? Justify your answer.

1. $\sigma_A(\sigma_B(R)) = \sigma_B(\sigma_A(R))$

   Answer:

   Yes, it holds by the commutativity property of σ.

2. $\pi_X(\pi_Y(R)) = \pi_X(R)$

   Answer:

   No, it only holds under the condition $X \subseteq Y$.

4. d Query Processing [4 marks]

Consider the following relation schemas:

- MOVIE(title, production_year, country) with the primary key {title, production_year};
- PERSON(id, first_name, last_name, year_born) with the primary key {id};
- DIRECTOR(pid, title, production_year) with the primary key {pid}
  and the foreign keys:

  [pid] $\subseteq$ PERSON[id];

  [title, production_year] $\subseteq$ MOVIE[title, production_year].

4. d (i) [2 marks]

Translate the following SQL query into a relational algebra expression, and then draw the query tree correspondingly.


```
SELECT MOVIE.title, PERSON.first_name
FROM MOVIE, PERSON, DIRECTOR
WHERE MOVIE.title = DIRECTOR.title
AND DIRECTOR.pid=PERSON.id
AND MOVIE.country='USA';
```


Answer:

- Relational algebra expression:
  - $\pi_{Movie.title, Person.first\_name}$
    $(\sigma_{Movie.title=Director.title \wedge Director.pid=Person.id \wedge Movie.country='USA'}$
    $(\text{MOVIE} \times \text{DIRECTOR} \times \text{PERSON}))$

4. d (ii) [2 marks]

Optimise your tree by applying at least two different transformation rules of relational algebra studied in lectures.

Answer:

- Since country is an attribute of MOVIE, by the rule $\sigma_{\varphi}(R_1 \bowtie R_2) \equiv R_1 \bowtie \sigma_{\varphi}(R_2)$, if R_1 is unaffected by φ, we have
  - $\pi_{Movie.title,Person.first\_name}$
    $(\sigma_{Movie.title=Director.title \wedge Director.pid=Person.id}$
    $(\sigma_{country='USA'}(\text{MOVIE}) \times \text{DIRECTOR} \times \text{PERSON}))$
- Since first_name is an attribute of PERSON, by the rule $\pi_X(R_1 \bowtie R_2) \equiv \pi_X(\pi_{X_1}(R_1) \bowtie \pi_{X_2}(R_2))$, where X_i contains attributes both in R_i and X, and ones both in R_1, we have:
  - $\pi_{Movie.title,Person.first\_name}$
    $(\sigma_{Movie.title=Director.title \wedge Director.pid=Person.id}$
    $(\pi_{title,pid}(\sigma_{country='USA'}(\text{MOVIE}) \times \text{DIRECTOR}) \times \pi_{id,first\_name}(\text{PERSON})))$
- Further optimization can be applied, for example, pushing the selection condition $Movie.title = Director.title \wedge Director.pid = Person.id$ down into the joins, i.e.,
  - $\pi_{Movie.title,Person.first\_name}$
    $(\pi_{title,pid}(\sigma_{country='USA'}(\text{MOVIE}) \bowtie_{Movie.title=Director.title} \text{DIRECTOR})$
    $\bowtie_{Director.pid=Person.id} (\pi_{id,first\_name}(\text{PERSON})))$

The general idea is to apply *push-down selection* and *push-down projection*.

Question 5: Transactions and Security [7 marks]

5. a [1 mark]

What are the ACID properties?

(1) atomicity, constant, isolation, durability

(2) atomicity, consistency, isolation, duration

(3) atomicity, consistency, isolation, durability

(4) atomicity, consistency, indexing, durability

(5) atomicity, constant, indexing, durability

Answer: (3)

5. b [2 marks]

Suppose that there is no concurrency control for the following transactions T_1 and T_2. What kind of problem can occur in this case?

T_1	T_2
readItem(Y)	
writeItem(Y)	
	readItem(X)
	readItem(Y)
	writeItem(Y)
readItem(X)	
abort	

Answer: The dirty read problem. The explanation about how this problem might occur in this case should be provided (refer to the text book and the lecture notes).

5. c [2 marks]

Consider the following SQL code built by an application, in which the email address tom@gmail.com was entered by the user:

SELECT name, password FROM PERSON WHERE email = 'tom@gamil.com';

Show how an SQL injection attack can happen in this case.

Answer: an SQL injection injects a string input through the Web application which changes the SQL statement to their advantage.

```
SELECT name, password
FROM PERSON
WHERE email = 'tom@gamil.com' OR 'x'='x';
```


5. d [2 marks]

Consider the table PROJECT that has been created in a relational database.

5. d (i) [1 mark]

Use SQL to give the read and update privileges on table PROJECT to Bob.

Answer:

- grant SELECT, UPDATE on PROJECT TO Bob;

5. d (ii) [1 mark]

Use SQL to cancel Bob's update privilege on table PROJECT.

Answer:

- revoke UPDATE on PROJECT from Bob;