THE UNIVERSITY OF BRITISH COLUMBIA
CPSC 304: MIDTERM ANSWER KEYS

Ques 1. What does the term data independence mean, and why is it an important goal? **[4 marks]**

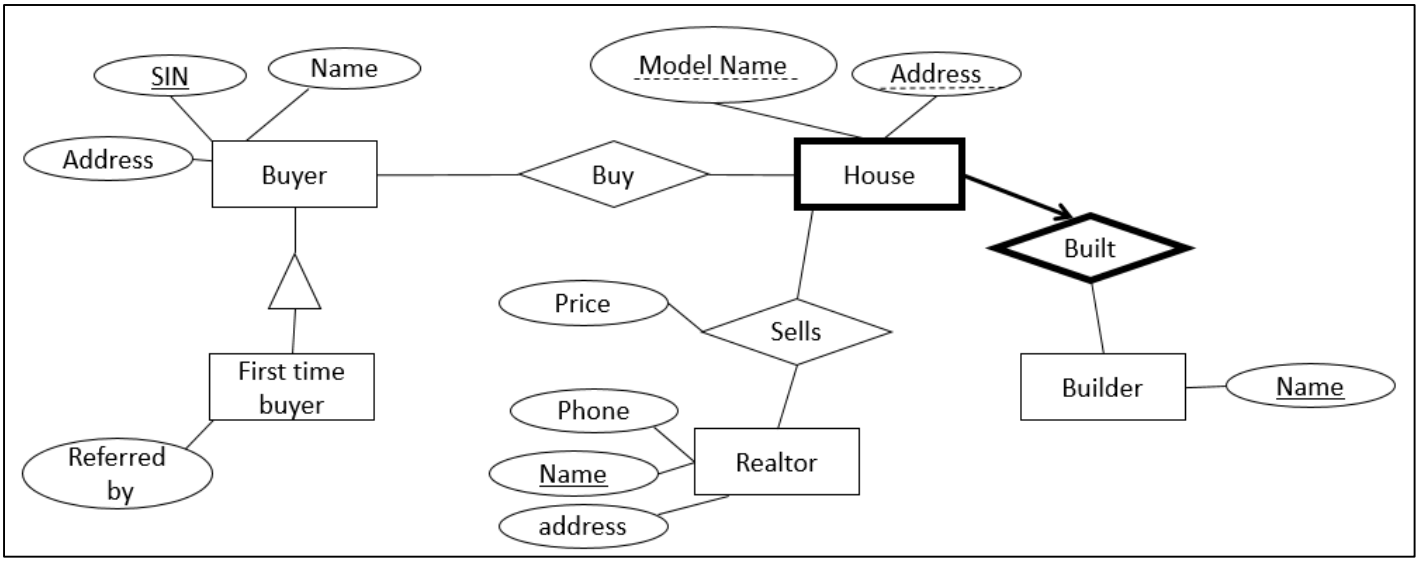
Data independence is the layering of a database into abstractions so changes in one layer will not affect the other layers.

Physical data independence - The ability to modify the physical schema without changing the logical schema.

Logical data independence - The ability to change the logical schema without changing applications.

It is an important goal because we don't want to have to change every single part of the database when we make a change to only one part. For example if you moved your database from harddisks to ssds, you don't want to have to change all your schemas because of this.

Ques 2. Consider the below ER diagram. The ER diagram represents the buying and selling of the house. Provide answers to the questions based on the diagram **[6 Marks]**

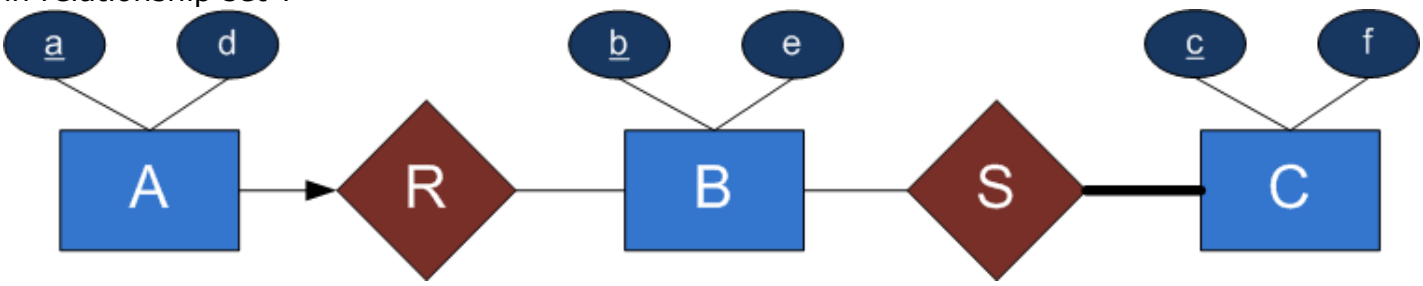


(a) If we delete the builder from the database, then what would happen to the houses that the builder had built? Justify your argument.
All information about these new houses are deleted from the database as well. This is due to the fact that new house is a weak entity of builder entity and instances of new house cannot exist without the existence of their strong entity instance. i.e Builder

(b) If we delete the realtor that sold a house, do we need to delete that house as well? Why or why not?"
No, because the house is not a weak entity to realtor. In fact, the house entity and realtor entity participate in a many to many relationships with no participation constraints. Thus, there could be other realtors that sold the house, or realtor can just be NULL for the associated house.

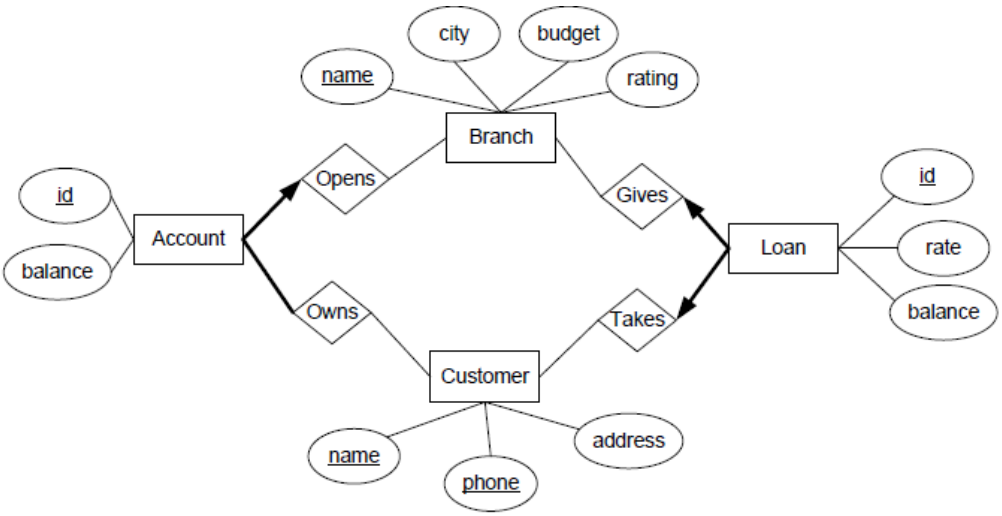
Ques 3. Answer the following multiple-choice questions. **[3 Marks]**

a. Suppose that a1 and a2 are the only entities of A, b1 and b2 are the only entities of B, and c1 and c2 are the only entities of C. Which of the following relationship sets for R and S are possible according to the diagram, where T = {(e1,f1)} means a relationship between e1 and f1 exists in relationship set T



A. R = {}; S = {}
B. R = {(a1,b1)}, S = {(b2,c2)}
C. R = {(a1,b1), (a1,b2)}; S = {(b1,c1), (b2,c2)}
D. R = {(a1, b2)}; S = {(b1,c2), (b2, c1), (b1,c1)}
E. None of the above

b. How many tables will we need to store the info provided by the following diagram?



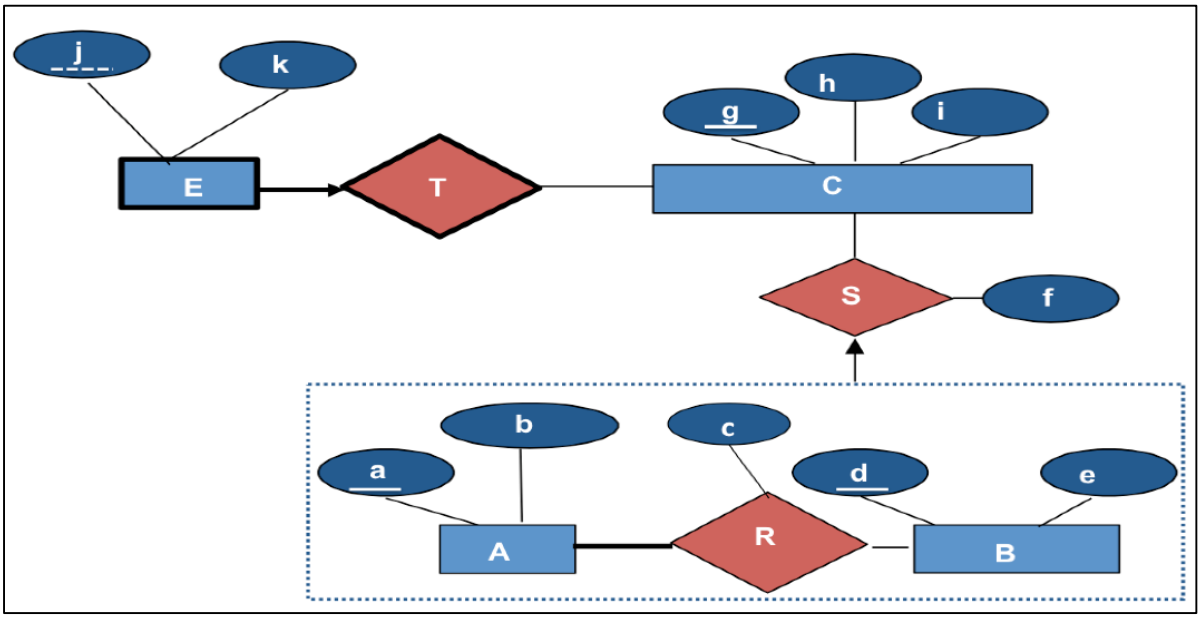
A. 4
B. 5
C. 6
D. 7
E. Some other number

c. Consider the below table definition, in which numbers denote line numbers. Which of the following is not a legal addition?

```
CREATE TABLE Emps (
        (1) id INT,
        (2) sin INT,
        (3) name CHAR(20),
        (4) managerID INT );
```

A. Add PRIMARY KEY before the comma on line (1), add FOREIGN KEY (managerID) REFERENCES Emps(id) before the) on line (4).
B. Add PRIMARY KEY before the comma on line (1), add FOREIGN KEY (managerID) REFERENCES Emps(sin) before the) on line (4).
C. Add, PRIMARY KEY (id, sin), and FOREIGN KEY (id, sin) REFERENCES Emps(id, sin) before the) on line (4).
D. All are legal
E. None are legal

Ques 4. Transform the ER diagram into a relational schema. State any assumptions that you make – but your assumptions cannot contradict the facts given. You only need to show the schema. (Use underline to show primary key and circle to show foreign key constraints) **[5 Marks]**



Primary keys are underlined and foreign keys are bold.

C(**G** , H, I)
ET(**G** , J , K)
A(A , B)
B(D , E)
RS(**A** , **D** , C, F, **G**)

Ques 5. Consider the two tables shown below. The TerID column in the SALES_REP relation is a foreign key referring to the primary key column TerID in the TERRITORY relation**.**
[6 Marks]

SALES_REP		
SRID	SRName	TerID
1	Joe	E
2	Sue	E
3	Meg	C
4	Bob	S
5	Joe	N
6	Pat	N
7	Lee	N
8	Joe	E

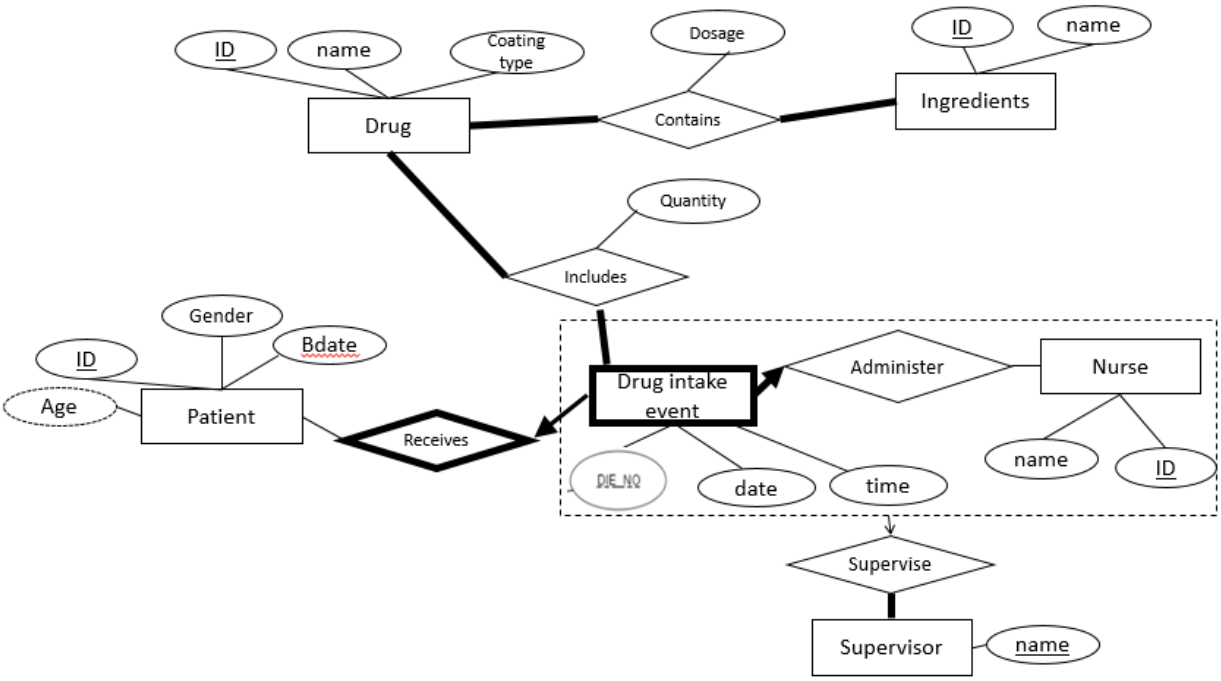
TERRITORY	
TerID	TerName
E	East
W	West
S	South
N	North
C	Central

a. If a DBMS enforces a DELETE CASCADE option on the referential integrity constraint between SALES_REP and TERRITORY, show the records in tables SALES_REP and TERRITORY after a user tries to delete the second record (W, West) from TERRITORY.
The record (W, West) from TERRITORY would be deleted. No other records would be deleted because there are no records in the SALES_REP table with a TerID of W.

b. If a DBMS enforces a DELETE RESTRICT option on the referential integrity constraint between SALES_REP and TERRITORY, show the records in tables SALES_REP and TERRITORY after a user tries to delete the fourth record (N, North) from TERRITORY. The record would not be deleted because there are records in the SALES_REP table that have N as a foreign key.

Ques 6. Midtown Memorial Patients Drug Dispense Database keeps track of the following information. Create an ER diagram based on the requirements **[15 marks].**

- For each patient it keeps track of a unique patient identifier, patient birthdate, gender and the patient age which is derived from patient birthdate.
- For each nurse it keeps track of a unique nurse ID and the nurse name.
- For each ingredient it keeps track of the unique ingredient identifier and the ingredient name.
- For each drug it keeps track of a unique drug ID, the drug name and the drug coating type.
- For each drug intake event by a patient it keeps track of the drug intake event number (unique within a patient, but different patients could have the same drug intake event number for their separate drug intake events), the drug intake event date and the time.
- Each ingredient must be contained in at least one drug, but it may be contained in many drugs. Every drug can contain many ingredients, but they must contain at least one. For every instance of an ingredient contained in a drug it keeps track of the dosage of the ingredient in the drug.
- Each drug must be included in at least one drug intake event, but it may be included in many. Each drug intake event must include at least one drug, but it may include many.
- For each instance of a drug included in a drug intake event it keeps track of the quantity.
- Each nurse can administer several drug intake events, but it does not have to administer any. Each drug intake event must be administered by one nurse.
- When a nurse administers the event, a supervisor must supervise her work in the event. The supervisor can be identified by a name.
- Each drug intake event is related to one patient. A patient can have many drug intake events, but they may not have any.


ID
name
Coating type
Drug
Dosage
Contains
ID
name
Ingredients
Quantity
Includes
Gender
ID
Bdate
Age
Patient
Receives
Drug intake event
Administer
Nurse
DIE_NO
date
time
name
ID
Supervise
Supervisor
name

Ques 7. Consider the two tables T1 and 2. Show the results of the following relational algebra operations: **[6 Marks]**

Table T1		
P	Q	R
10	A	5
15	B	8
25	C	6

Table T2		
A	B	C
10	b	6
25	c	3
10	b	5

a) $T1 \bowtie_{T1.P=T2.B} T2$

Nothing
The domains are different

b) T1 ∩ T2

Nothing
There are no rows that are the same between the two tables

Ques 8. Consider the following schema
Student (StudID, name, address, gender, major)
Course (CNo, CName)
Enrolled (StudID,CNo)
where the primary keys are underline. Give an expression in the **relational algebra** to express each of the following queries: **[9 Marks]**

a. List course numbers for which no student is enrolled.
$\pi_{CNo}(Course) - \pi_{CNo}(Enrolled)$

b. List ids of the students who are enrolled in 'Database' or 'Data Mining' course.

$$\pi_{StudId}(Student \bowtie Enrolled \bowtie (\sigma_{CName='Database'}Course) \cup \pi_{StudId}(Student \bowtie Enrolled \bowtie (\sigma_{CName='Data\ Mining'}Course)$$

c. List the name of the courses and the student ids in which all 'COMP' major students are enrolled.

$$allCompSciStudents \leftarrow \sigma_{major='COMP'}Student$$
$$\pi_{StudId,\ CName}((Enrolled / allCompSciStudents) \bowtie Course \bowtie Enrolled)$$