CSC 370: Database Systems
Spring 2024 (A01: CRN 20791; A02: CRN 20792)
Midterm #1: Thursday, 8 February 2024

Sample Solution

Exam duration: 50 minutes **Instructor:** Michael Zastre

Students must check the number of pages in this examination paper before beginning to write, and report any discrepancy immediately.

- **All answers are to be written on this exam paper.**
- The exam is closed book
- When answering questions, please do not detach any exam pages!
- No electronic devices are permitted. Phones must be turned off.
- There are fourteen (14) printed pages in this document, including this cover page; four numbered pages are intentionally left blank.
- The exam is out of a total of 100 marks.
- We strongly recommend you read the entire exam through from beginning to end before starting on your answers.
- **Please have your UVic ID card available for inspection by an exam invigilator.**

For Questions 1 to 10, place an X beside all answers that apply. **Each of these ten questions is worth three (3) marks. *Partial marks are not given for incomplete answers*.**

Question 1: Elements of a *data model* as described in lectures include:

__X__ operations.

__X__ structure.

____ functions.

____ scope rules.

____ None of the above.

Question 2: A *relation instance*:

____ is another name for a database.

__X__ may be represented by a table.

____ is normally shown as a linked list.

__X__ may be represented by a set of tuples.

____ None of the above.

Question 3: A *tuple* as described in this course:

____ is equivalent to the column in a relation table.

__X__ is normally a member of some set corresponding to a relation instance.

____ may be represented by an empty set.

____ is normally an operand for relational-algebra operators.

____ None of the above.

Question 4: As discussed in lectures, the ***relational algebra*** includes (amongst others) the operators:

____ $\alpha(), \beta()$ and $\gamma()$

____ $\rho(), \pi()$ and $\theta()$

____ $\alpha(), \theta()$ and $\bowtie$

____ $\cup, \sigma()$ and $\alpha()$

__X__ None of the above.

Question 5: Consider the two relation schemas *R(A, B, C)* and *S(A, D)*. The relation schema *T(A, B, C, D)* would result from the expression:

____ $R \times S$

____ $R - S$

__X__ $R \bowtie S$

____ $R \cup S$

____ None of the above.

Question 6: In the context of relational databases, a *key* for a relation:

__X__ may consist of one attribute.

__X__ may consist of many attributes.

__X__ is a subset of a superkey.

____ is another name for the main tuple of that relation.

____ None of the above.

Question 7: A *superkey* as described in lectures is called *super* because:

____ it is the most powerful of all relational operators.

____ it must apply to more than one relation schema.

____ it gives you root privileges to the database.

____ it is the key to the SuperBowl.

__X__ None of the above.

Question 8: Consider a relation and its schema *R(A, B, C, D, E)*. If we have some instances *R1* and *R2* of this schema, and want to create new relations based on these instances where only attributes *a*, *d*, and *e* were present, we would use:

____ Use the *extraction* operator on *R1* and *R2*.

____ Use the *delete attribute* operator on *R1* and *R2*.

____ Use the *reduce relation* operator on *R1* and *R2*.

__X__ Use the *projection* operator on *R1* and *R2*.

____ None of the above.

Question 9: *Schema decomposition* is:

__X__ normally guided by one of the normal forms (e.g., BCNF or 3NF).

_(X)_ used to partition tuples from a single relation into subsets of tuples from that single relation.

__X__ often used as a way of eliminating update and deletion anomalies.

____ often used as a way of eliminating projection and selection anomalies.

____ None of the above.

Question 10: As shown in lectures, the *chase test* is one method for:

____ determining the closure of a set of functional dependencies.

____ performing select operations on a bag of tuples.

__X__ proving a lossless join is possible for a given decomposition.

____ deriving the basis set of a schema's functional dependencies.

____ None of the above.

For questions 11 and 12, assuming the following three relation schemas.

Suppliers(<u>*sid*: integer</u>, *sname*: string, *address*: string)
Parts(<u>*pid*: integer</u>, *pname*: string, *colour*: string)
Catalog(<u>*sid*: integer, *pid*: integer</u>, *cost*: real)

.

Question 11: Find the *snames* of suppliers who supply some red part. Express your answer using the "sequence of assignment statements" form of relational algebra. ***Explain your answer.*** [10 marks]

$R_1 = \sigma_{colour=\text{red}}(Parts)$

$R_2 = R_1 \bowtie Catalog$

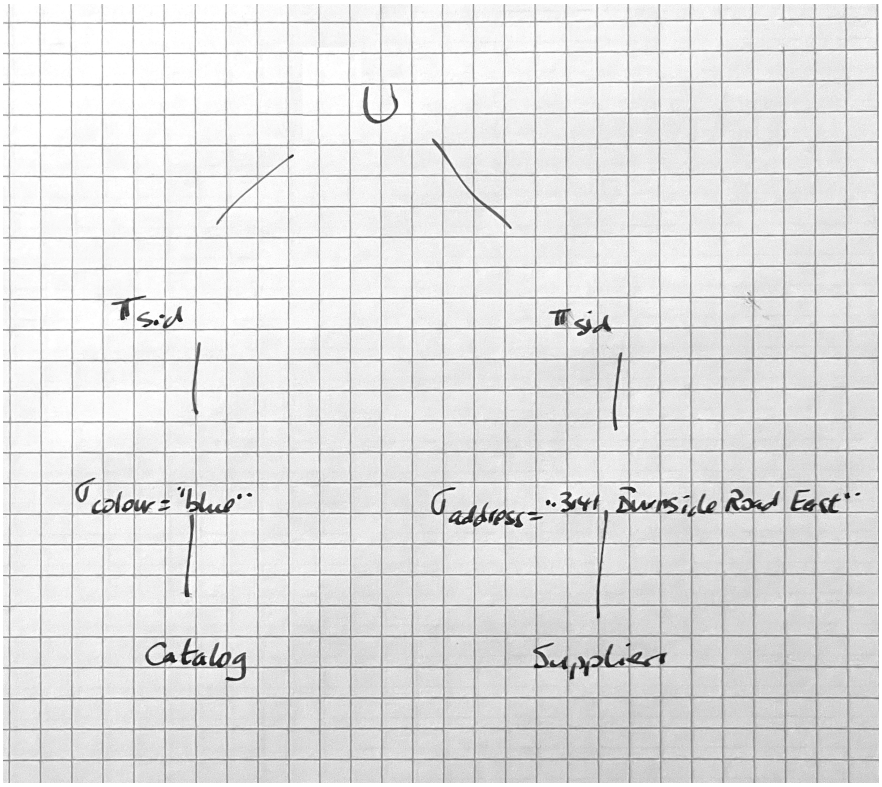
$R_3 = R_2 \bowtie Suppliers$

$R_4 = \pi_{sname}(R_3)$

The select is used to obtain only those parts in the Part relation with the colour red, but we need to combine that information with Catalog and Suppliers in order obtain the supplier name. Therefore we join the result of the select with Catalog, then Suppliers, and then project out the supplier name.

Relational algebra queries (in correct "sequence" form): 6 marks
Explanation: 4 marks

Question 12: Find the *sids* of suppliers who supply some blue part or are at 3141 Burnside Road East. Explain your answer. Express your answer using the "expression tree" form of relational algebra. ***Explain your answer.*** [15 marks]



The parts that are blue, and the suppliers that are at the given address, are obtained via separate select statements on Catalog and Supplier respectively. Since both of the relations have information on supplier ID (i.e. sid) we need not perform any joins with the select results and other relations. Therefore we simply project out the sids from both select results, and combine them together using a set union.

Relational algebra query (in "expression tree" form) where set union is used: 10 marks

Relational algebra query (in "expression tree" form) where set union is *not* used: 6 marks.

Explanation: 5 marks

Question 13 (15 marks)

Given the relation $R(A, B, C)$ shown below, state for each of the lettered dependencies a) through e) whether or not the dependency is satisfied by the relation. ***Briefly explain each answer.***

A	B	C
1	4	2
3	5	6
3	4	6
7	3	8
9	1	0

a) $A \rightarrow B$

NOT SATISFIED. Tuples 2 and 3 agree on A (value of 3) but differ on B (values of 4 & b)

b) $A \rightarrow C$

SATISFIED. The only two tuples with the same A value are tuples 2 and 3, and they also agree on their C values (i.e. 6)

c) $AB \rightarrow C$

SATISFIED but in a trivial way – that is, all AB values in the relation are unique.

d) $C \rightarrow A$

SATISFIED. Tuples 2 and 3 have the same values for C (i.e. 6) and in each tuple they also have the same values for A.

e) $AC \rightarrow B$

NOT SATISFIED. Tuples 2 and 3 agree on their AC values (3 and 6) but different on their B values (5 and 4).

Each part: 3 marks

This page intentionally left blank.

Question 14 (15 marks)

Given the relation:

$R(A, B, C, D, E, F, G, H)$

and the set of functional dependencies:

$\{ BC \rightarrow GH,\ AD \rightarrow E,\ A \rightarrow H,\ E \rightarrow BCF, G \rightarrow H \}$

show that AD is a key for the relation, ***and then*** decompose R into BCNF by decomposing the relation *in the order of the given functional dependencies*. ***Show all work.***

(7 marks) To show AD is a key, we compute its closure in order to find all of the attributes.

- $AD^+ \Rightarrow AD(initial)$
- $\Rightarrow ADE$ from $AD \rightarrow E$
- $\Rightarrow ADEH$ from $A \rightarrow H$
- $\Rightarrow ABCDEFH$ from $E \rightarrow BCF$
- $\Rightarrow ABCDEFGH$ from $BC \rightarrow GH$

And $ABCDEFGH$ is the whole relation, therefore AD serves as a key as all attributes have values depending on AD (and no attributes can be removed from AD such that it remains a key).

(8 marks) BCNF decomposition proceeds by identifying an FD and a relation which violates BCNF.

- Given $ABCDEFGH$ and $BC \rightarrow GH$ we have an FD that violates BCNF as the LHS is not a superkey for the relation. The two new relations are $BCGH$ (from the FD) and $ABCDEF$.
- We can ignore next $AD \rightarrow E$ because the LHS is a key.
- We can also ignore next $A \rightarrow H$ as neither $BCGH$ nor $BCDEF$ contain all of the FDs attributes.
- Given $ABCDEF$ and $E \rightarrow BCF$ we have an FD that violates BCNF as the FD involves attributes in the relation, and the LHS of the FD is not a superkey for this relation. The two new relations are $EBCF$ (from the FD) and AED.

- Given $BCGH$ (which we created in the first decomposition step above), and the FD $G \rightarrow H$ (which involves attributes on the relation), we have an FD that violates BCNF as the LHS is not a superkey for this relation. The two new relations are GH (from the FD) and BCG
- Final set of relations: $ADE, BCEF, BCG, and\ GH$

This page intentionally left blank.

Question 15 (15 marks)

Given the relation:

$R(A, B, C, D, E, F)$

and the set of functional dependencies:

$\{ AD \rightarrow B, A \rightarrow E, C \rightarrow E, DEF \rightarrow A, F \rightarrow D \}$

find a key for R that has two attributes. ***Show all work; part marks will be given for your work even for showing how you eliminate candidates for the key.***

Why is AD not a key? Because its closure can only ever grow to be ABD (i.e. only $AD \rightarrow B$ can be applied to the set of attributes).

Any key will need to involve attributes in the LHS, and so one could just exhaustively examine all such combinations. However, only one results in a closure to contains all five attributes.

- $CF + \Rightarrow CF$
- $\Rightarrow CEF$ from $C \rightarrow E$
- $\Rightarrow CDEF$ from $F \rightarrow D$
- $\Rightarrow ACDEF$ from $DEF \rightarrow A$
- $\Rightarrow ABCDEF$ from $AD \rightarrow B$

15 marks: Rejection of AD as a key, plus explanation of how CF is the single key (with possible work showing other possibilities that do not work)

10 marks: Only CF is given as the single key, with work shown (but no further work shown).

5 marks: CF is given as the single key, but with an unpersuasive explanation.

This page intentionally left blank.

This page intentionally left blank.

END OF EXAM