Exam #1 Answers
Thursday October 2, 2010 at 2 pm

Database Systems, CSCI-4380-01

Note. The exam is open book and open notes. Use your own book and notes only, sharing is not allowed. Electronic gadgets are NOT allowed during the exam. Write your answers clearly, legibly and explain your reasoning as much as you can. If I cannot read or understand your answers, you will not get points.

Question 1 (24 points). Answer this question using the data model in the appendix. Answer (1) and (2) below using relational algebra.

(a) Find the name of all artists who contributed (wrote and composed) to the song titled: "Bohemian Rhapsody".

Answer.

$$
\begin{aligned}
T1 &:= SongComposedBy \cap SongWrittenBy \\
T2 &:= \Pi_{songId,artistId}(\sigma_{name='Bohemian\ Rhapsody}(T1 \bowtie_{SongId=Id} Song)) \\
T3 &:= \Pi_{Name}(T2 \bowtie_{artistId=id} Artist)
\end{aligned}
$$

(b) Find the name of all artists who have not released an album after '12-31-20000'.

Answer.

$$
\begin{aligned}
T1 &:= \Pi_{artistId}((\sigma_{releaseDate>'12-31-2000'} Album) \bowtie_{id=albumId} AlbumReleasedBy) \\
T2 &:= (\Pi_{id} Artist) - (\rho_{T1(id)} T1) \\
T3 &:= \Pi_{name}(T2 \bowtie Artist)
\end{aligned}
$$

(c) What does the following relational algebra expression return (for T3)? Write it using a succint English sentence (i.e. do not use words like join, project, etc. in your answer.) Show your work for partial credit.

$$
\begin{aligned}
T1(id1, name1, bio1, birthYear1) &:= Artists \\
T2 &:= \Pi_{id}\ (T1 \bowtie_{birthYear1>birthYear} Artists) \\
T3 &:= \Pi_{id}(Artists) - T2
\end{aligned}
$$

Answer.

T2 finds all artists who are not oldest (have an older artist in the database) T3 finds all the oldest artists in the database.

Question 2 (16 points). Are the following true or false? Explain with a short sentence.

(a) Whenever $A \to BC, CD \to E$ are both in a set of functional dependencies F, then $AD \to CE$ is implied by F.

Answer. True, $AD+ = \{A, D, B, C, E\} \supseteq \{C, E\}$, hence this f.d. is implied.

(b) If two relations have the same keys, then they also have the same set of functional dependencies.

Answer. False, not necessarily. $\{A \to B, B \to C\}$ and $\{A \to BC\}$ have the same key but are not equivalent.

(c) All relations have a key.

Answer. True. Even if there are no functional dependencies, then all the attributes in the relation forms a key.

(d) Relation $R(A, B, C)$ and the result of the query $\Pi_{A,B}R$ always have the same number of tuples.

Answer. False. The projection may create and remove duplicates, resulting in fewer tuples than R.

Question 3 (12 points). Are the following relations in 3NF, BCNF or 4NF? Briefly describe why or why not.

1. $R1(A, B, C, D, E)$, $F1 = \{AB \to CDE, D \to A\}$

   Answer. Keys: AB, BD. Not in BCNF or 4NF, but in 3NF (AB superkey, A prime attribute).

2. $R2(A, B, C, D, E)$, $F2 = \{AB \to C, CD \to E\}$

   Answer. Key: ABD. Not in 3NF, BCNF or 4NF. AB, CD are not superkeys, C,E are not prime attributes.

3. $R3(A, B, C, D, E)$, $F3 = \{ABC \twoheadrightarrow DE\}$

   Answer. Key: all attributes. In 3NF, BCNF and 4NF. The only mvd is trivial.

Question 4 (8 points). Decompose the following relation to find a 3NF decomposition.

$R(A, B, C, D, E, F)$, $F = \{AB \to DE, CD \to AF\}$

Answer. Keys: ABC, BCD.

We get:

R1(A,B,D,E) R2(A,C,D,F) R3(A,B,C) (or for the other key, R3(B,C,D))

Question 5 (8 points). Is the following decomposition of $R(A, B, C, D, E, F, G)$ lossless? Show your work using the Chase algorithm.

$F = \{CD \rightarrow F, D \rightarrow E, C \rightarrow D, BF \rightarrow AC\}$

$R1(A, B, C, D)$, $R2(C, D, E, F)$, $R3(B, F, G)$

Answer.

A	B	C	D	E	F	G
a	b	c	d	e1	f1	g1
a2	b2	c	d	e	f	g1
a3	b	c3	d3	e3	f	g

applying, $CD \rightarrow F$, we get f1=f

A	B	C	D	E	F	G
a	b	c	d	e1	f	g1
a2	b2	c	d	e	f	g1
a3	b	c3	d3	e3	f	g

applying $BF \rightarrow AC$, we get a3=a, c3=c

A	B	C	D	E	F	G
a	b	c	d	e1	f	g1
a2	b2	c	d	e	f	g1
a	b	c	d3	e3	f	g

applying $C \rightarrow D$, we get d3=d

A	B	C	D	E	F	G
a	b	c	d	e1	f	g1
a2	b2	c	d	e	f	g1
a	b	c	d	e3	f	g

applying $D \rightarrow E$ we get $e3 = e$

A	B	C	D	E	F	G
a	b	c	d	e1	f	g1
a2	b2	c	d	e	f	g1
a	b	c	d	e	f	g

Given the last row has no subscripts, this relation is lossless.

Question 6 (6 points). Are the two models (a) and (b) in Figure 1 equivalent? Explain why or why not.

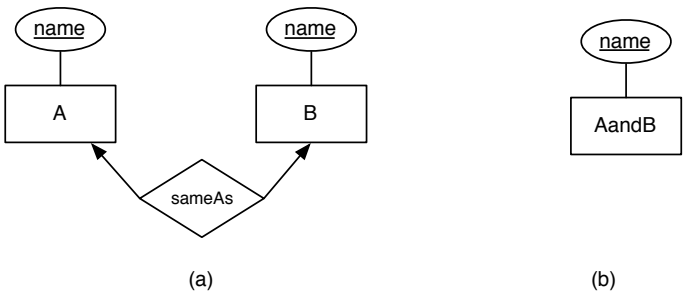

Figure 1: E-R Diagram for Question 6

Answer. No, they are not equivalent. A and B may be the same, but may have different names (apples are the same as bananas). It is also possible for an A not to be a B and vice versa. If they are combined, two name attributes are needed, one for A and one for B.

Question 7 (12 points). Convert the ER diagram in Figure 2 to the relational data model.

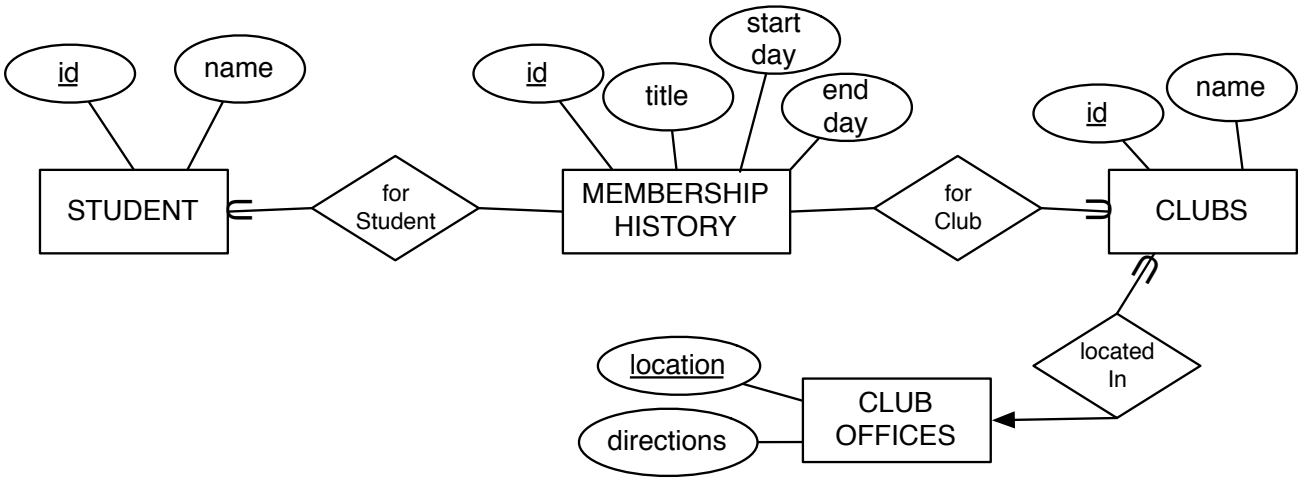

Figure 2: E-R Diagram for Question 7

Answer. Student(id, name)

Clubs(id, name)

MembershipHistory(id, title, startday, endday, studentId, ClubId)

ClubOffices(location, directions, clubId)

Question 8 (14 points). Create an ER diagram for the following data model.

Suppose you are creating a database to store the parking information at RPI. The database will store people who will obtain parking permits, their RIN, name, address (building name and room number) and status (student, faculty, etc.) In addition, there are lots and parking permit types to store. Lots have name and location, permit types have name and the types of people who can have them. People are eligible for permits due their status (for example, permit A is for students, B is for faculty and staff, etc.). Different permits types may be applicable to multiple lots and a lot may have multiple permit types. People own permits for a specific car and a specific permit type. A person may have more than one permit, each permit is for a different car. Cars are distinguished by their license number.

Answer.

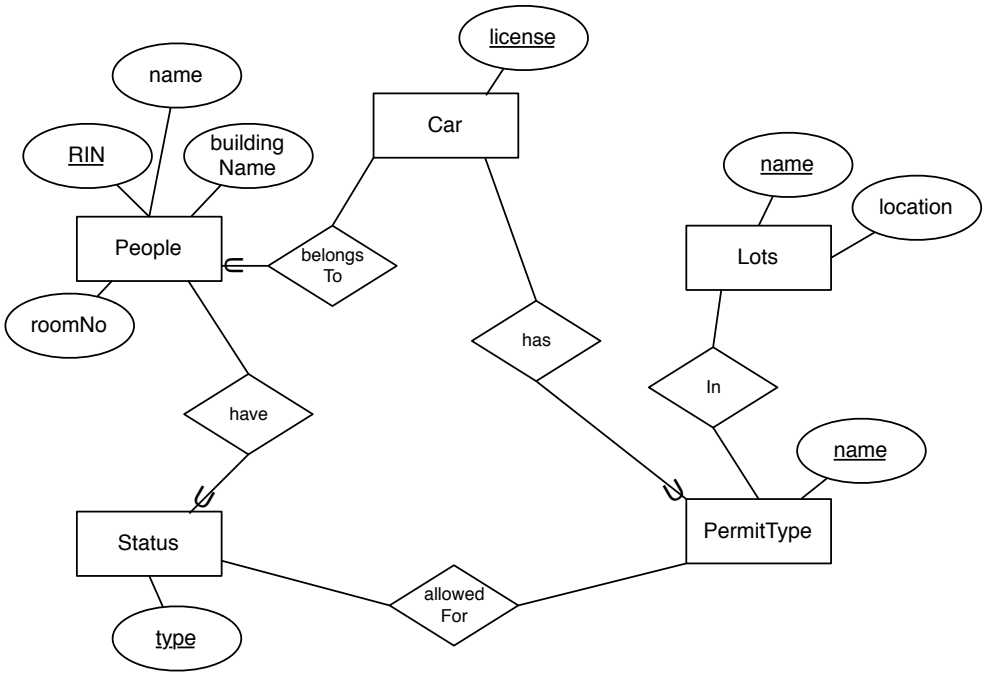

Figure 3: E-R Diagram for Question 8

Appendix

Suppose you are given the below data model for an application providing music services. This is the same data model that we have created in class (except I did not include the data relating to the users).

Note that, albumId refers to Albums(id), songId refers to Songs(id), artistId refers Artists(id), recordingId refers to Recordings(id) and characteristingsId refers to Characteristings(id).

DATA MODEL.

Artists(id, name, bio, birthYear)

Albums(id, name, releaseDate)

Songs(id, name, lyrics)

Recordings(id, trackNum, length, songId, albumId)

Characteristics(id, name)

RecordingHasCharacteristics(recordingId, characteristicId)

AlbumsReleasedBy(albumId, artistId)

SongComposedBy(songId, artistId)

SongWrittenBy(songId, artistId)