Database Systems, CSCI 4380-01 Exam #1 Answers
Thursday February 17, 2010 at 2 pm

1 (24)	2 (16)	3 (12)	4 (8)	5 (8)	6 (6)	7 (12)	8 (14)	TOTAL

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 authors of the books published in 2010 on topic 'Databases'. Return their id and name.

Answer.

$$
\begin{aligned}
R1 &:= \pi_{book\_id}((\sigma_{Title='Database'}Topics) \bowtie_{id=topic\_id} BookTopic) \\
R2 &:= \pi_{book\_id,person\_id}(R1 \bowtie BookAuthor \bowtie_{book\_id=id} (\sigma_{publicationYear=2011}Book)) \\
R3 &:= \pi_{id,name}(R2 \bowtie_{person\_id=id} Person)
\end{aligned}
$$

(b) Find books that have both at least one editor and one author, but have no topic listed in the database. Return the id and title of the books.

Answer.

$$
\begin{aligned}
R1 &:= (\pi_{book\_id}(BookAuthor) \cap \pi_{book\_id}(BookEditor)) - (\pi_{book\_id}(BookTopics) \\
R2 &:= \pi_{id,title}(R1 \bowtie_{book\_id=id} Books)
\end{aligned}
$$

(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.

$$\begin{aligned} B2(bid2, pid2) &:= BookAuthor \\ B3 &:= B2 \bowtie_{book\_id=bid2 \text{ and } person\_id<>pid2} BookAuthor \\ B4 &:= \pi_{book\_id}(BookAuthor) - \pi_{book\_id}(B3) \\ B5 &:= \pi_{id,title}(B4 \bowtie_{book\_id=id} Books) \end{aligned}$$

Answer. Returns books that have exactly one author.

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

(a) BCNF decomposition guarantees that the resulting decomposition is dependency preserving and lossless.

Answer. False. BCNF decompositions are lossless, but not necessarily dependency preserving.

(b) It is possible that $R \bowtie S = R \cap S$.

Answer. True, this is possible if R and S have exactly the same schema.

(c) The relations $R(A, B, C)$ and $R(B, A, C)$ are equivalent (have the same meaning).

Answer. True. The order of attributes in relational data model is not significant.

(d) If an attribute only appears on the right hand side of functional dependencies, then it must appear in a key.

Answer. False. If the dependency appears only one the right hand side, then other attributes can be used to infer this attribute and hence it is not needed in the key.

Question 3 (12 points). Are the following relations in 3NF or BCNF? Briefly describe why or why not by listing which functional dependencies violate which normal form.

1. $R2(A, B, C, D, E)$, $F2 = \{A \to B, BD \to C\}$

   Answer.

   Key. ADE. None of the functional dependencies have a superkey on the left, so they both violate BCNF.

   As none of the functional dependencies also do not have a prime attribute on the right, they also both violate 3NF.

   Neither 3NF, nor BCNF.

2. $R2(A, B, C, D, E)$, $F1 = \{ABC \to DE, CD \to A\}$

Answer. Keys: ABC, BCD. $ABC \to DE$ satisfies both BCNF and 3NF as it has a superkey on the left.

$CD \to A$ violates BCNF as it does not have a superkey on the left. However, this satisfies 3NF as A is a prime attribute.

Not BCNF, but in 3NF.

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

Company(name, branchNo, address, phoneNo, account)

name, branchNo → address
address → name
name, address → phoneNo

Answer.

Keys: address, branchNo, account and name, branchno, account.

This is not in minimal basis, but the following is:

name, branchNo → address
address → name, phoneNo

3NF Decomp:

R1(name, branchno, address), key name, branchno

R2(address, name, phoneNo), key address

R3(address, branchno, account), key all attributes

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

$F = \{AB \rightarrow C, BC \rightarrow DE\}$
Decomposition: $R1(A, B, C)$, $R2(B, C, D, E)$, $R3(A, B, D, F)$

Answer.

A	B	C	D	E	F
a	b	c	d1	e1	f1
a2	b	c	d	e	f2
a	b	c3	d	e3	f

Given $AB \rightarrow C$, c3 becomes c.

Given $BC \rightarrow DE$, e3 becomes e. The last line does not have any subscripts, and hence this decompositions is lossless.

Question 6 (6 points). Are the two models (a) and (b) in Figure 3 equivalent? In other words, can (a) store all the data stored in (b) and vice versa. Explain why or why not.

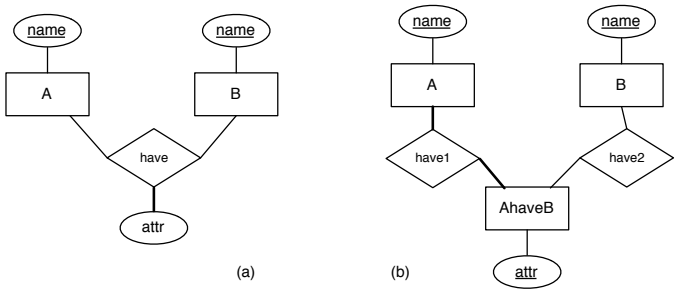

Figure 1: E-R Diagram for Question 6

Answer. No. Two reasons:

1. In (b), we can only store one instance of attr. In (a), different A, B combinations can have the same attr.
2. In (a), we store attr for a pair of A and B. In (b), we can store attr even if there is no A or B associated with it, or if there is one of A or B, or both.

So, (a) is not more general than (b), and vice versa.

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

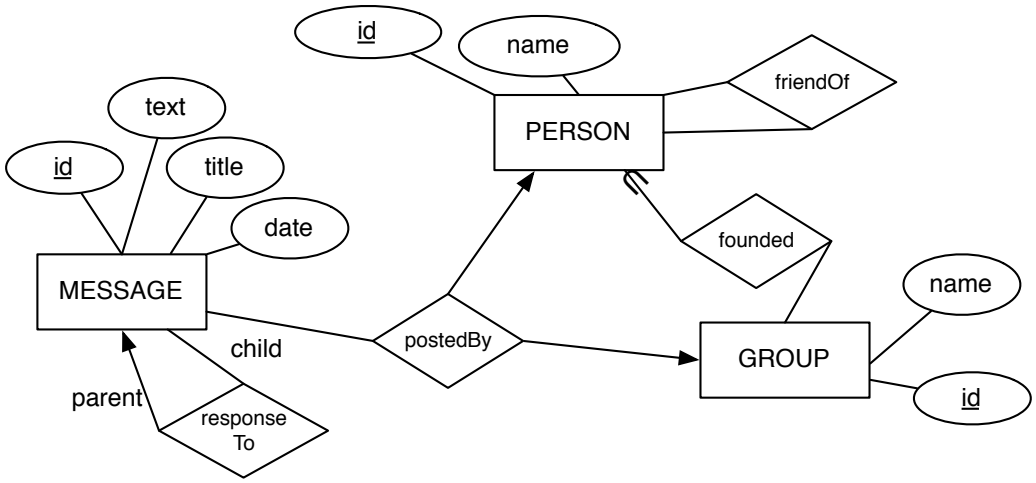

Figure 2: E-R Diagram for Question 7

Answer.

Message(id, text, title, date, parentMessageid, posterPersonid, posterGroupid)

Person(id, name)

Friend(person1id, person2id)

Group(id, name, founderPersonid)

Question 8 (14 points). This question was given to me by Watson who wants to know who should be eliminated first when he takes over.

Create an ER diagram for the following data model.

Suppose you are creating a database to track people's achievements in life. You are especially interested in their education, their total worth and the types of awards they got. For each person, you will store their SSN, name, total worth, birthdate. Also, you would like to know all the college degrees they got at Bachelor's, Master's and Doctoral levels. People may get many such degrees or none. The name of the school, the name and type of the degree and the year the degree is obtained is stored. Also, you would like to store various awards they got. For each award, store the name of the award and the year they got it. A person may get the same award many times, even in the same year. Finally, this is the most important, you would like to know all addresses for this person: their work addresses and their home addresses. There could be many of each type. For each address, store street, state, city, zip, latitude and longitude.

Watson thanks you for the convenience.

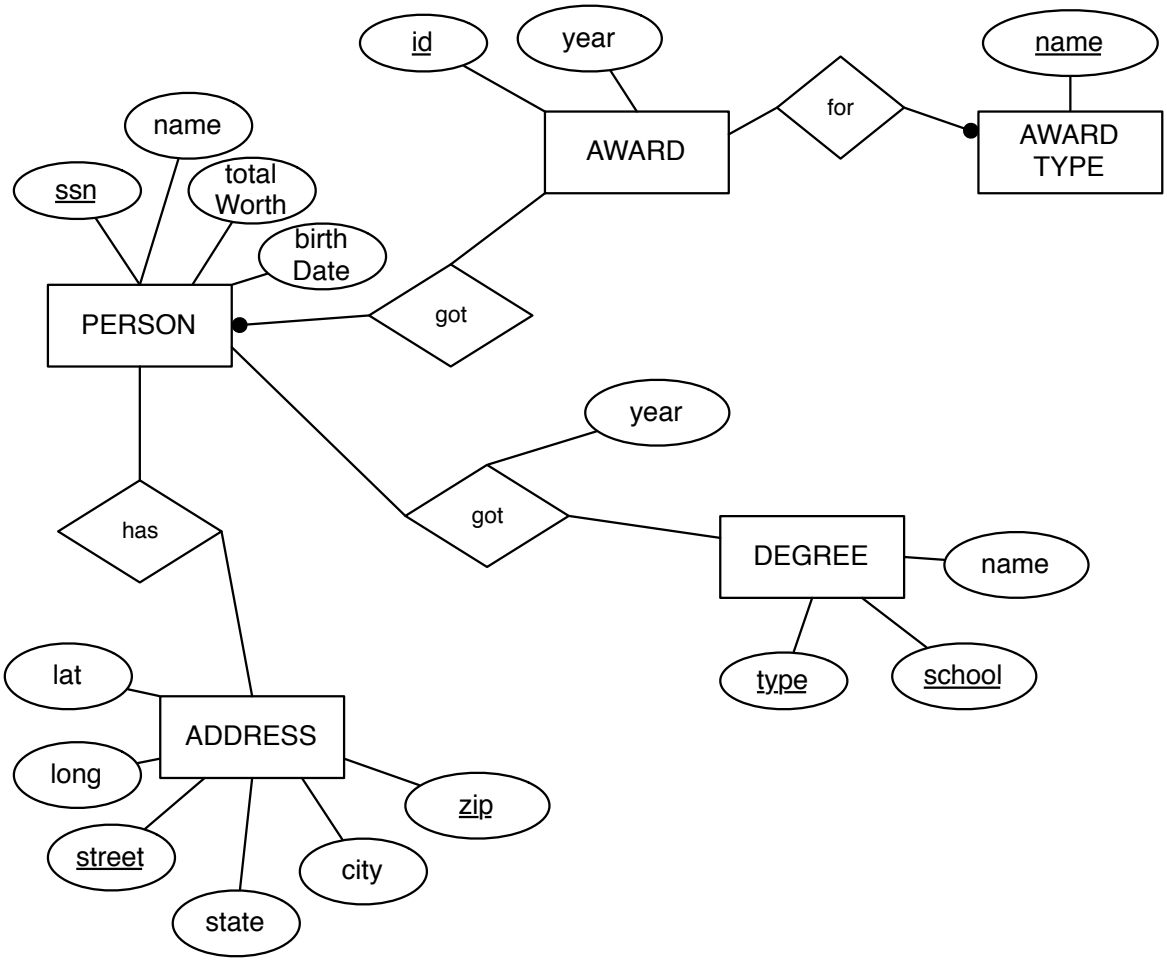

Figure 3: E-R Diagram for Question 8 Answer

Appendix

DATA MODEL.

Person(id, name, countryOfBirth, yearOfBirth)
Topics(id, title)
Books(id, title, publisher, isbn, versionNo, publicationYear)
BookAuthor(book_id, person_id)
BookEditor(book_id, person_id)
BookTopic(book_id, topic_id)