

CS411 Database Systems

Fall 2007

Midterm Solution

Problem 1

- (1) False; (2) True; (3) True; (4) False; (5) True;
(6) False; (7) False; (8) False; (9) False; (10) True;
(11) False; (12) False; (13) True; (14) False;

Problem 2

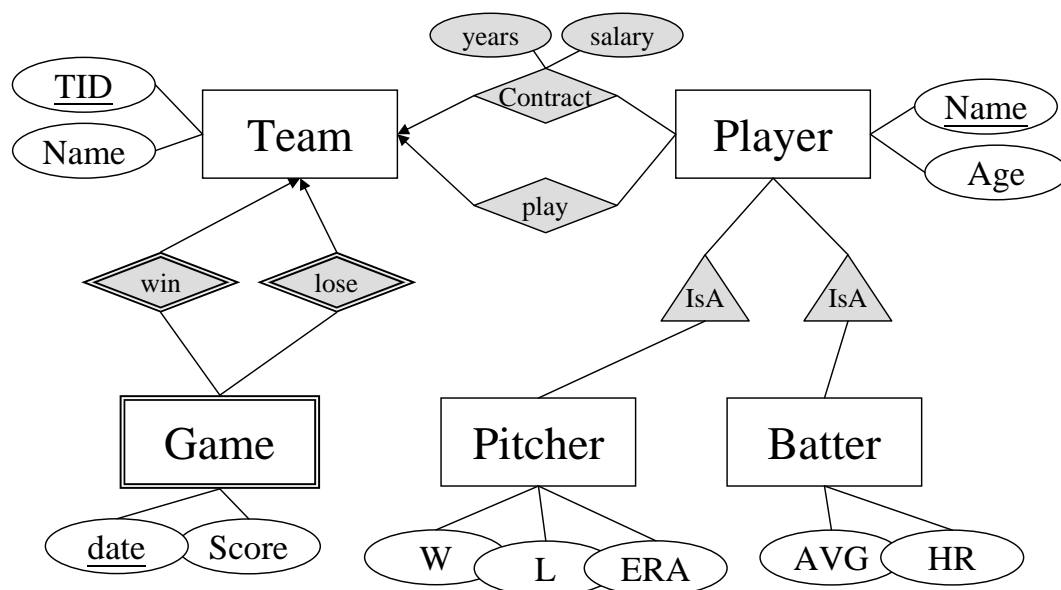


Figure 1: ER Diagram

Problem 3

Consider a relation $R(A, B, C, D, E)$, with given FD's $AB \rightarrow C$, $BC \rightarrow D$, $CD \rightarrow E$, $DE \rightarrow A$.

- (i) Determine all the keys of R .
(Hint: There are three keys and you don't need to list superkeys that are not keys)
Answer: AB , BC , and BDE . Note that B must be in any key, since it doesn't appear on the right of any FD. That fact makes the search for keys fairly easy.
- (ii) List which FDs violate 3NF if any.
None violate 3NF, because all attributes are prime.

- (iii) List which FDs violate BCNF if any.

Answer: $CD \rightarrow E$ and $DE \rightarrow A$ violate BCNF.

- (iv) Decompose R using BCNF decomposition. Indicate your working and summarize your final set of relations.

Answer: Suppose we use $CD \rightarrow E$ to decompose. Since $\{CD\}^+ = ACDE$, one of the schemes is $R_1(A, C, D, E)$ and the other is $R_2(B, C, D)$. The latter is in BCNF, since BC is the only key, and $BC \rightarrow D$ the only projected FD. However, R_1 is not in BCNF. For example, $DE \rightarrow A$ is a projected FD, but $\{DE\}^+ = ADE$, so DE is not a superkey for R_1 . Thus, we decompose R_1 into $R_3(A, D, E)$ and $R_4(C, D, E)$. The constituents of the decomposition are R_2 , R_3 , and R_4 .

Problem 4

- (i) $\pi_{Name}(Student \bowtie (\rho_{TeamName='BEE'}ProjectTeam))$
- (ii) $\pi_{UIN}Student - \pi_{UIN}ProjectTeam$
- (iii) $\pi_{TeamName}ProjectTeam - \pi_{TeamName}(ProjectTeam \bowtie \sigma_{Department \neq 'CS'}Student)$
- (iv) $\pi_{Name}(Student \bowtie (Midterm - \pi_{UIN, Grade}(Midterm \bowtie_{Grade < Grade1} \rho_{M2(UIN1, Grade1)}Midterm))))$

Problem 5

- (i)

```
SELECT Customer.cname
FROM Buy, Customer, Book
WHERE Customer.cid = Buy.cid AND Customer.state = 'Illinois' AND
      Buy.isbn = Book.isbn AND Buy.year = 2000
GROUP BY Customer.cid
Having SUM(Book.price) > 5000;
```
- (ii)

```
CREATE VIEW Sales AS
SELECT Author.assn, Author.aname, COUNT(Buy.tid) AS count
FROM Author, Book, Buy
WHERE Buy.isbn = Book.isbn AND
      Book.isbn = Author.isbn AND
      Buy.year = 2006
GROUP BY Author.assn;

SELECT Author.aname
FROM Sales
WHERE Sales.count = (SELECT MAX(Sales.count)
                     FROM Sales);
```

- (iii) `CREATE VIEW FriendsOfBob AS
 SELECT Buy.cid AS cid
 FROM Buy
 WHERE Buy.cid <> 12345 AND
 Buy.isbn IN (SELECT Buy.isbn
 FROM Buy
 WHERE Buy.cid = 12345)
 GROUP BY Buy.cid
 HAVING COUNT(Buy.isbn) > 20;`
- (iv) `SELECT Buy.isbn
 FROM FriendsOfBob, Buy
 WHERE FriendsOfBob.cid = Buy.cid AND Buy.isbn NOT IN
 (SELECT Buy.isbn
 FROM Buy
 WHERE Buy.cid = 12345);`

Problem 6

- (i) 1. $RemoteLikes(p, q) \leftarrow Likes(p, q) \text{ AND NOT } Dislikes(p, q)$
 2. $RemoteLikes(p, q) \leftarrow RemoteLikes(p, r) \text{ AND } RemoteLikes(r, q) \text{ AND NOT } Dislikes(p, q)$
- (ii) There is only one IDB predicate, *RemoteLikes*, and there is no negative self-loop arc on it although there is a positive self-loop arc. Therefore, negation in rules in Problem 6(i) is stratified.