PART B (BASIC)

1. (7 pts.) The (natural) *outer join* of two relations R and S is formed by taking the natural join of R and S and adding to it any tuple of R that joins with no tuple of S. Such a tuple is extended with the value `NULL` in all attributes of S that are not attributes of R. Likewise, we add to the result of the outer join any tuple of S that joins with no tuple of R, padded with `NULL` in attributes of R that are not attributes of S. For example, if:

```
          A       B                   B       C
          -------------               --------------
R=        1       2           S=      2       3
          4       5                   6       7
```

then the outer join of R and S is

```
A       B       C
------------------------
1       2       3
4       5       NULL
NULL    6       7
```

Write an expression of relational algebra that computes the outer join of arbitrary relations $R(A, B)$ and $S(B, C)$. You may use a relation N whose value is {`NULL`}, that is, a single one-attribute tuple, with value `NULL`. For simplicity, you may write your expression as a sequence of assignment steps, in which temporary relations are assigned simple expressions. For example, one such step might be:

$$T(A, B, C) := E$$

where E is some expression of relational algebra that produces a 3-attribute relation.

2. (8 pts.) Let $R(A, B, C)$ be a relation. Give examples of sets of dependencies (functional and/or multivalued) for R such that:

   (a) R is in Fourth Normal Form.

   (b) R is in Boyce-Codd Normal form but not in Fourth Normal Form.

   (c) R is in Third Normal form but not in Boyce-Codd Normal Form.

   (d) R is not in Third Normal Form.

   For each part, give the key(s) and explain why there is or is not a violation of the normal form or forms mentioned.

PART B (ADVANCED)

3. (6 pts.) A set of attributes X is said to be *closed* if $X \rightarrow A$ is false for every attribute A that is not in X. Suppose $R(A, B, C)$ is a relation whose only closed sets of attributes are the empty set and ABC. Prove that $A \rightarrow B$ holds for R.

4. (7 pts.) Let Q be the conjunctive query with negation:

$$Q: \quad \texttt{p(X,Y) :- a(X,Y) \& NOT b(X,Y)}$$

   Give an example of another conjunctive query (possibly with negation) that is contained in Q and is not equivalent to the empty query [i.e., the query `p(X,Y) :- false`], but does not contain Q. Prove that your query satisfies each of these conditions.

5. (7 pts.) Suppose we have a relation `Sells(bar,beer,price)`. A tuple (b, r, p) in this relation means that bar b sells beer r at price p. Write an SQL query that finds all bars b mentioned in `Sells` such that every beer that bar b sells is sold by some other bar at a lower price. To help us understand your query, indicate briefly what each subquery produces.

6. (10 pts.) Let p be a predicate, and suppose the two views v and w are defined in terms of p by:

```
v(X,Y) = p(X,Z) & p(Z,Y)
w(X) = p(X,X)
```

   Suppose we wish to answer the query

$$Q: \quad \texttt{ans(X) :- p(X,Y) \& p(Y,X)}$$

   by selecting a set S of conjunctive queries of the form

```
ans(X) :- B
```

   where B is the body of a conjunctive query whose subgoals have predicates v or w only. There are two constraints we want to put on the selected bodies B:

   (a) `ans(X) :- B` should be contained in the query Q, in the sense that if we expand each subgoal of B using the definition of the rule v or w, as appropriate (being careful to introduce a new variable for Z each time we expand v) then the result is a conjunctive query involving predicate p that is contained in Q.

   (b) There should be no other query `ans(X) :- C` in the set of chosen queries S, such that `ans(X) :- B` is contained in `ans(X) :- C` *before* we expand v and w.

   Find a maximal set S of conjunctive queries to satisfy the above conditions for the given query Q. Give a brief demonstration (e.g., a containment mapping) for each containment that you claim.