PART B (BASIC)

1. (10 pts.) Define a *standard* functional dependency to be one that has a left side of one or more attributes and a right side consisting of exactly one attribute; that attribute must not be on the left side. Suppose R is a relation with a schema consisting of n attributes.

   (a) (4 pts.) As a function of n, how many standard FD's for R are there? Justify your answer if you want to be considered for partial credit.

   (b) (6 pts.) Suppose the instance of R has exactly two (distinct) tuples. As a function of n, what is the *maximum* number of standard FD's that this instance can *violate*? Again, justify your answer for partial credit.

2. (10 pts.) This question is based on the following information about companies, stocks, stock exchanges, and analysts. Companies issue one or more kinds of stock (e.g., common, preferred). A stock has a ticker symbol (e.g., `IBM` for IBM-common or `IBMX` for IBM-preferred). Each stock is listed on one exchange, and ticker symbols are unique for each exchange. A stock has a history consisting of the closing price on each day in the past when the stock was traded. Each analyst has a name (assumed unique) and works for a company. An analyst makes recommendations for certain stocks; each recommendation consists of a rating and a predicted price.

   (a) (7 pts.) Give an Entity/Relationship diagram for the information above. Indicate keys for your entity sets, and indicate relationships that are many-one or one-one.

   (b) (3 pts.) The ODL object-oriented data model allows certain things that E/R does not allow, but also forbids some constructs that E/R allows. Give two examples of how ODL makes the above design easier or harder.

PART B (ADVANCED)

3. (15 pts.) The following question is based on a relation $R(A, B, C, D)$.

   (a) (5 pts.) Give an expression of relational algebra whose result is empty if and only if the multivalued dependency $A \twoheadrightarrow B$ holds in R.

   (b) (7 pts.) The *embedded multivalued dependency* $A \twoheadrightarrow B|C$ means that if there are two tuples t and u in R that agree on A, then there is also in R a tuple s that agrees with t on A and B and agrees with u on A and C. Note that in comparison, if we were to describe what $A \twoheadrightarrow B$ means, we would replace "agrees with u on A and C" by "agrees with u on A, C, and D." Does $A \twoheadrightarrow B$ logically imply $A \twoheadrightarrow B|C$ (i.e., every instance of R that satisfies $A \twoheadrightarrow B$ also satisfies $A \twoheadrightarrow B|C$)? Does $A \twoheadrightarrow B|C$ logically imply $A \twoheadrightarrow B$? Justify your answers with brief proofs or counterexamples.

   (c) (3 pts.) Give a relational-algebra expression that is empty if and only if $A \twoheadrightarrow B|C$ holds in R.

4. (10 pts.) Consider the following conjunctive queries:

```
Q1: p(X,Y) :- q(X,Z) & q(Z,Y) & lt(X,Z) & lt(Z,Y)
Q2: p(X,Y) :- q(X,Z) & q(Z,Y) & lt(X,Y)
Q3: p(X,Y) :- q(X,Z) & q(Z,Y) & X<Z & Z<Y
Q4: p(X,Y) :- q(X,Z) & q(Z,Y) & X<Y
```

   You should assume that p, q, and lt are uninterpreted (relational) predicates, while $<$ has its usual arithmetic meaning. You may also assume that the domains of all predicate arguments are the real numbers. Determine which, if any, containments among the four conjunctive queries hold. Give brief arguments to support your conclusions.

5. (15 pts.) If a hypergraph of n hyperedges is acyclic, we know that there is a sequence of $2n-2$ semijoins that is a full reducer (i.e., after executing the semijoins, the relations corresponding to the hyperedges are globally consistent — each tuple of each relation participates in some join of all the relations). Prove that if $n \geq 2$, then any sequence of fewer than $2n-2$ semijoins cannot be a full reducer. *Hint*: Remember that you may assume the hypergraph is acyclic. Do an induction on n, and take advantage of the fact that an acyclic hypergraph with more than one edge always has an ear. Also remember that to prove a sequence is *not* a full reducer, you have only to show that there is some database instance it does not make globally consistent.