Qualifying Exam: Database
Closed books, 60 min., 2013.1

1. Answer the following: (28%)  
   1) What is the weak entity? Give an example.  
   2) What is the data independence?  
   3) What is the completeness of the D-schema with respect to the F-schema?  
   4) What is the referential integrity?  
   5) What is the clustering index?  
   6) Bayesian classifier requires the computation of p(d|c_j). How does the naïve Bayesian classifier estimate p(d|c_j)? Assume that d has n dimensions and denote the ith attribute value of d by d_i.  
   7) What is the support for association rules?

2. Select one for each of the following sentences.: (24%)  
   1) 3NF decomposition is (lossless, lossy) join decomposition.  
   2) BCNF decomposition is (lossless, lossy) join decomposition.  
   3) 3NF decomposition (is, is not always) dependency preserving.  
   4) BCNF decomposition (is, is not always) dependency preserving.  
   5) The conceptual design is (dependent, independent) on the data model.  
   6) The physical design is (dependent, independent) on DBMS.  
   7) The logical design is (dependent, independent) on the data model.  
   8) The logical design is (dependent, independent) on DBMS.

3. Consider the following relational schema where the primary keys are underlined.  
   Suppliers(sid, sname, address)  
   Parts(pid, pname, color)  
   Catalog(sid, pid, cost)

For a request “Find the sid’s of suppliers who supply all the blue parts”  
1) Give a relational algebra expression using DIVISION. (8%)  
2) Formulate an SQL query using some from EXISTS, NOT EXISTS, EXCEPT. (10%)  
3) What is the result of the query if there is no blue part? (6%)
4. Answer the following:
1) Set S consists of three classes. The fractions of instances of the three classes are 1/6, 1/2, 1/3, respectively. What is the Gini measure for the purity of S? Show the calculation steps. (6%)
2) How to translate a multivalued attribute in the relational logical design when the order of values is important? Provide two solutions and explain your answer with examples. (18%)