Introduction to Databases (winter term 2005/2006)

Assignment 7

hand in on December 19, 2005 during the tutorial

Please hand in your solutions in groups of up to 3 students.
Do not forget to write down your name and matriculation number on the solutions you hand in. Please also add your study course (e.g. Dipl.-Inf., Master SSE, ...).

Task 7.1 (Normal Forms) (2 Points)
Let $R = (U, F)$ be a relational schema. Prove the following propositions:

a) $R$ is in 3NF $\Rightarrow$ $R$ is in 2NF.

b) $R$ is in 2NF $\neq$ $R$ is in 3NF.

c) $R$ is in BCNF $\Rightarrow$ $R$ is in 3NF.

d) $R$ is in 3NF $\neq$ $R$ is in BCNF.

Task 7.2 (Decomposition) (4 Points)
Consider the following relation schema of a book store database:

$\text{BOOKS} = (\{\text{title, author, type, listprice, affiliation, publisher}\},$

$\{\text{title }\rightarrow\text{ publisher, type; type }\rightarrow\text{ listprice; author }\rightarrow\text{ affiliation}\})$

a) Identify the key of BOOKS and prove your claim.

b) In which normal form is BOOKS? Why?

c) The decomposition algorithm is non-deterministic, since the choice of the FD that violates the BCNF and is used to split the attribute set, is not fixed.
Show two different runs of the algorithm, such that the result of the first execution has the dependency preservation property and the second one has not. Which dependency is lost?

Task 7.3 (Synthesis) (4 Points)
Consider the following relation schema:

$R = (U, F) = (\{A, B, C, D, E, F\}, \{A \rightarrow BE, AE \rightarrow BD, F \rightarrow CD, CD \rightarrow BEF, CF \rightarrow B\})$.

a) Find all candidate keys of $R$ and prove that they are candidate keys.

b) In which normal form is $R$? Why?

c) Apply the synthesis algorithm to derive a schema in third normal form.
Task 7.4 (Lossless Join Property) (2 Points)
It is obvious, that any correct and complete set of FD inference rules (like e.g. the RAP rule set) can be used to decide, whether a decomposition $D$ of a relational schema $R = (U, F)$ preserves the dependencies $F$.

The following algorithm (taken from Elmasri/Navathe ‘Fundamentals of Database Systems’) checks whether a decomposition is loss-free.

Algorithm 15.2 Testing for the lossless (nonadditive) join property

Input: A universal relation $R$, a decomposition $D = \{R_1, R_2, \ldots, R_m\}$ of $R$, and a set $F$ of functional dependencies.

1. Create an initial matrix $S$ with one row $i$ for each relation $R_i$ in $D$, and one column $j$ for each attribute $A_j$ in $R$.
2. Set $S(i,j):= b_{ij}$ for all matrix entries.
   (* each $b_{ij}$ is a distinct symbol associated with indices $(i,j)$ *)
3. For each row $i$ representing relation schema $R_i$
   {for each column $j$ representing attribute $A_j$
     {if (relation $R_i$ includes attribute $A_j$) then set $S(i,j):= a_{ij}$};
     (* each $a_{ij}$ is a distinct symbol associated with index $(j)$ *)
   } repeat the following loop until a complete loop execution results in no changes to $S$
   {for each functional dependency $X \rightarrow Y$ in $F$
     {for all rows in $S$ which have the same symbols in the columns corresponding to attributes in $X$
       {make the symbols in each column that correspond to an attribute in $Y$ be the same in all these rows as follows: if any of the rows has an “$a$” symbol for the attribute in any of the rows, choose one of the “$b$” symbols that appear in one of the rows for the attribute and set the other rows to that same “$b$” symbol in the column};}};
4. If a row is made up entirely of “$a$” symbols, then the decomposition has the lossless join property; otherwise it does not.

Consider the following relation schema:
$R = (ABCDEFGHJ, \{AB \rightarrow C, A \rightarrow DE, B \rightarrow F, F \rightarrow GH, D \rightarrow IJ\})$

Decide by using the above algorithm, whether the following two decompositions $D = ((X_1, F_1), \ldots, (X_n, F_n))$ of $R$ are lossless or not. Write down each change of the matrix in the 4th step.

a) $X_1 = ABC, \quad X_2 = ADE, \quad X_3 = BF, \quad X_4 = FGH, \quad X_5 = DIJ$
b) $X_1 = ABCDE, \quad X_2 = BFGH, \quad X_3 = DIJ$