2 Databases (tgg22)

Suppose we have a relational database with three tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Key Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>S(sid, A)</td>
<td>sid</td>
</tr>
<tr>
<td>T(tid, B)</td>
<td>tid</td>
</tr>
<tr>
<td>R(sid, tid, C)</td>
<td>sid, tid, C</td>
</tr>
</tbody>
</table>

The following referential integrity constraints hold on the table R. Every sid-value in R is the key of a record in table S and tid-value in R is the key of a record in table T.

(a) What does it mean that the attribute C is included in R’s key? [2 marks]

(b) Write an SQL query to return records of the form sid where sid is the key of an S record that is not R-related to any records in table T. [2 marks]

(c) Consider the following two queries.

\[
Q1 = \text{SELECT S1.A AS A, R1.C AS C FROM R AS R1 JOIN R AS R2 ON R2.tid = R1.tid JOIN S AS S1 ON S1.sid = R1.sid JOIN S AS S2 ON S2.sid = R2.sid}
\]

and

\[
Q2 = \text{SELECT DISTINCT S1.A AS A, R1.C AS C FROM R AS R1 JOIN R AS R2 ON R2.tid = R1.tid JOIN S AS S1 ON S1.sid = R1.sid JOIN S AS S2 ON S2.sid = R2.sid}
\]

Note that the only difference is the use of \text{DISTINCT} in Q2.

(i) If Q1 and Q2 return the same results, what can you conclude about the data in this database? Justify your answer. [4 marks]

(ii) Suppose we add this where-clause to each query:

\[
\text{WHERE R1.C = R2.C AND S1.sid <> S2.sid}
\]

If the modified Q1 and Q2 return the same results, what can you conclude about the data in this database? Justify your answer. [4 marks]

(iii) Suppose we add the where-clause \text{WHERE R1.tid <> R2.tid} to each query. If the modified Q1 and Q2 return the same results, what can you conclude about the data in this database? Justify your answer. [4 marks]

(iv) Suppose we add this where-clause to each query:

\[
\text{WHERE R1.tid <> R2.tid AND S1.sid <> S2.sid}
\]

If the modified Q1 and Q2 return the same results, what can you conclude about the data in this database? Justify your answer. [4 marks]