

1 Databases (tgg22)

Suppose we have a relational database with five tables.

| table | key attributes |
|----------------|----------------|
| S(sid, A) | sid |
| T(tid, B) | tid |
| U(uid, C) | uid |
| R(sid, tid, D) | sid, tid |
| Q(tid, uid, E) | tid, uid |

Here R implements a many-to-many relationship between the entities implemented with tables S and T, and Q implements a many-to-many relationship between the entities implemented with tables T and U.

- (a) Write an SQL query that returns all records of the form sid, uid where sid is the key of an S-record and uid is the key of a U-record and these two records are related through the relations R and Q. Use SELECT and not SELECT DISTINCT in your query. [3 marks]
- (b) Write an SQL query that returns records of the form A, C where the A-value is from an S-record and the C-value is from a U-record and these two records are related through the relations R and Q. Use SELECT and not SELECT DISTINCT in your query. [3 marks]
- (c) Could one of your queries from parts (a) and (b) return more records than the other? If so, which one? Justify your answer. [3 marks]
- (d) Suppose you replaced SELECT with SELECT DISTINCT in your queries from parts (a) and (b). Could one of these modified queries return more records than the other? If so, which one? Justify your answer. [3 marks]
- (e) Consider again your query from part (a). If pair sid, uid is returned by this query then there must exist at least one “path” that goes from from table S to table T (via relation R) and then from table T to table U (via relation Q). Note that there can be many such paths for a given pair sid, uid. Write an SQL query that returns records of the form tid, total where tid is a key of a record from table T and total indicates the total number of such paths that “go through” that record. [8 marks]