Create Merging Table

Description

Given a database, this generates a spreadsheet that can be annotated to indicate how entities should be merged. The user selects the type of entity.

Usage Hints

After loading a database, if manual merging of some of the entities is required to perform the desired analysis, run this algorithm, annotate the merges, and run Merge Entities on the annotated results to merge the annotated entities together.

After running this algorithm, the output table will have three additional columns, named Merge Group Identifier, Primary Entity, and From table: TABLE. NAME. The last of the three columns has no value. The column name for that column should not be modified. That is used to indicate to the merging algorithm which type of entity this data is for merging together.

Every row represents an entity in the database. Initially, every row will have a merge group identifier that is different, and a star in the primary entity column, indicating it is the primary entity for its merge group. None of the columns should be altered except for Merge Group Identifier and Primary Entity, though altering any columns other than the primary key will have no effect. It is especially important not to modify any columns that are part of the primary key (usually called pk).

In order to modify this table, it should be saved outside of the tool. After being modified, it should be loaded back into the tool.

To indicate two entities should be merged together, give them the same Merge Group Identifier. This can be anything, so long as it is different from all other Merge Group Identifiers. For every Merge Group Identifier, one and only one of the rows in the table should be marked as the Primary Entity (with a *). All other rows with that Merge Group Identifier should have no value in the Primary Entity column. After merging, the row with the * will be the only one of the rows from the merge group that remains in the database.

Feel free to sort and otherwise reorder the data, provided the column names and values (other than the ones that can be changed as indicated above) remain the same.

Implementation Details

An entity is available for selection if and only if the table for it in the database has a primary key.