Introduction to Database Design: Exercise 5: Report

Björn Þór Jónsson

Introduction

This report accompanies Exercise 5, and presents assumptions and other information that is not represented in the ER diagram and SQL script.

Notes

Data types: VARCHAR(50) is used for names, VARCHAR(200) to refer to files, and VARCHAR(2000) to refer to test case input and output. It is assumed that these lengths suffice. Amounts are assumed to be in whole units of currency and stored as INT.

Item 1: This requirement is not enforceable in the ER diagram. The requirement that gender is either M or F is currently enforced using a CHECK constraint. If a more flexible definition of gender is required, a lookup table could be used.

Item 2: The level of a degree can be stored as BSc, MSc or PhD, thus requiring only three characters.

Item 4: Since the end date of a current membership period is not known, the end attribute is allowed to contain NULL. Each interval of club membership should not overlap with the others. This is not enforced; a trigger could be implemented to enforce this when a record is inserted, or the start or end columns modified.

Item 5: The combination (venue, date) is potentially unique. More information is needed to determine this, and therefore this constraint is not enforced.

Item 7: A tournament may have 0 leagues; typically in the beginning. A league, however, is assumed to belong to exactly one tournament.

Why is league NOT a weak entity? Here, the text assigns them an ID, so they do have their own keys. But designing them as a weak entity isn't really wrong either, as we know that "the combination of league number and gender is unique within each tournament", which is a classic partial key for weak entity. Here, the choice for a regular entity is made due to the given existence of a globally unique key.

Item 8: Since the rank is determined at the end of a tournament, the corresponding column must be allowed to contain NULL.

Item 9: Note that there is no step in the PPTX file, as this requirement is not enforceable in the ER diagram. The requirement that gender is M, F or X is currently enforced using a CHECK constraint.

As with item 1, a lookup table could be used. However, since X is used for leagues but not applicable to danes, some additional mechanism would be required to prevent X gender in danes.

Furthermore, during insertion to the Participates table, it is required to enforce that either the gender of the league is the same as that of the dane, or the gender of the league is 'X'. This is best done using a trigger.

Note that should a dane decide to change genders, the rule represented here would be violated.

Item 10: As specified in the requirements, the database does not enforce payment for a dane before the league starts, or in fact at any time.

The wording "A club" is taken to mean "exactly one club". Although the Participates and Pays table therefore have the same primary key, and thus could be merged, this would require a) allowing two columns to have NULL values in Participates, and b) requiring that either both are NULL or neither is NULL. Enforcing that is difficult, and hence the tables are best kept separate. Note that IF the payment MUST be made at the time of registration, merging the tables would make sense.

The key result of the aggregation is that the table Pays has *one foreign key*, to the Participates table, rather than two to the Dane and League tables. This means that a club cannot pay for a dane that is not registered.

Item 12: We assume that the text of a problem is a reference to a document containing the text, just as the name of the dance step is a reference to a dance step program.

The database schema cannot easily enforce the requirement that at least one PDF employee authors a problem. Such a relationship can be maintained with triggers in some systems.

Item 13: Here we note that the problem must have multiple test cases (presumably 2+). We cannot capture in the ER diagram that is must have multiple cases, only that it can have at least one. And in the SQL code we cannot enforce even the ER diagram requirement easily.

Why is test case a weak entity? The text says: "Each test case has a name and an order indicator (both are distinct but for one particular problem)". This indicates that the test cases do not have their own key, and hence they are weak entities.

The input and output attributes could also refer to files rather than actual text.

Item 16: Gender is said not to be required for club rankings. We can assume that age is not required either. If registered, the natural checks are applied.

The requirements for Covering and Disjoint sub-classes of Ranking are not enforced. Doing so would require using deferred triggers and a system that allows those.

Based on the text of item 18, we can assume that nationality is unique across club rankings. The fact that there should be a single global ranking is not

represented in the database design. This could be implemented using a trigger as it only relates to a single table.

Item 17: There is no requirement to register each dane each year; only the current rank is required. This requirement is captured in the proposed design.

Item 18: Identical considerations apply to clubs as to danes in item 17.