Five Normal Forms in Relational Database Theory

Stephanie Dahl and Erik Kringen

Chapter 7: Normalization of Relational Tables

Normalization

- Guidelines for removing redundancies in data storage
- Prevents data inconsistencies caused by frequent updates
- Penalizes data retrieval
First and Second Normal Forms

- All occurrences of a record type must contain the same number of fields.
- Relational database theory doesn't allow a variable number of fields.

- All non-key fields are not facts about a sub-set of a key.
- Ex:
  - |Part | Warehouse | Warehouse Address|
  - |Part | Warehouse |

  Correct:
  - |Part | Warehouse | Warehouse Address|
  - |Part | Warehouse |

  Part and Warehouse are a key.
Third Normal Form

- All non-key fields are not facts about another non-key field
- _______________________________________________________
- |Part | Warehouse | Warehouse Address|
- |-------------|----------------|
- Correct:
- -----------------------------------------------
- |Part | Warehouse| |Warehouse | Warehouse Address|
- |-------------|-------------|
- If there is no Functional Dependancy, it does not violate Third Normal Form
Fourth Normal Form

- A table shouldn’t have more than one many to one relationship
- | EMPLOYEE | SKILL | LANGUAGE |
  |--------------------------------|
- Instead, they should be represented in the two records
- | EMPLOYEE | SKILL | | EMPLOYEE | LANGUAGE |
  |-------------------------------|-------------------------------|
- Minimizes problems with updates
- Eliminates confusing cross product representation
Fifth Normal Form

- Cannot be reconstructed into smaller tables that contain the same information

<table>
<thead>
<tr>
<th>AGENT</th>
<th>COMPANY</th>
<th>COMPANY</th>
<th>PRODUCT</th>
<th>AGENT</th>
<th>PRODUCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith</td>
<td>Ford</td>
<td>Ford</td>
<td>car</td>
<td>Smith</td>
<td>car</td>
</tr>
<tr>
<td>Smith</td>
<td>GM</td>
<td>Ford</td>
<td>truck</td>
<td>Smith</td>
<td>truck</td>
</tr>
<tr>
<td>Jones</td>
<td>Ford</td>
<td>GM</td>
<td>car</td>
<td>Jones</td>
<td>car</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>GM</td>
<td>truck</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- if an agent sells a certain product, and he represents a company making that product, then he sells that product for that company.
Conclusion

- Fifth normal form is considered the ultimate form
- Still basic guidelines- not complete or necessarily advantageous
- Some types of redundancy are unavoidable