

Workshop

Midterm

1. Note that control questions are only a sample. To study for the final midterm review all the slides and the book!
2. Material includes everything from Lecture X (SQL Libraries).
3. Bring your laptop to the final midterm.

Functional Dependency

1. Inference Test. Construct inference test with two tuples (using 0, ? and 1 symbols). Describe all the steps in your reasoning.
 - A. Assume a set of FDs $F = \{ABC \rightarrow DEF, D \rightarrow G, D \rightarrow H, GH \rightarrow IJ\}$
 - B. Is it true that $ABC \rightarrow I$?
 - C. Is it true that $DEF \rightarrow IJ$?
 - D. Is it true that $GH \rightarrow D$?

Closure Test

1. Assume set of FDs $F = \{AB \rightarrow C, AB \rightarrow D, CD \rightarrow EF, F \rightarrow G, G \rightarrow CD, I \rightarrow J\}$
2. Compute closure of AB^+ and the closure of F^+ .
3. Describe each of the steps in your computation (Basis and Induction).
4. Based on the closure information:
 - A. Is it true that $AB \rightarrow EF$?
 - B. Is it true that $AB \rightarrow G$?
 - C. Is it true that $F \rightarrow AB$?
 - D. For each of the points provide justification.

Data integration

1. Provide three examples of schema heterogeneity.
2. Assume the following table with an FD: Postal_Code \rightarrow City. Propose an algorithm to clean the data. Provide justification of your decisions.

Postal_Code	City
M4Y2W4	TO
M4Y2W4	TO
M4Y2W4	TO
M4Y2W4	MIA

Is there any alternative way to clean the data assuming data come from different sources with a level of confidence?

Data Warehousing

1. Draw a Star-Schema for Sales data warehouse of car dealership (that consists of five tables).
2. Provide description which tables are fact tables and dimension tables.
3. Which attributes are dimension attributes and dependent attributes (in the fact table)?

DTD and XML Document

1. Provide DTD and XML Document for PART of your car dealership schema (covering portion of sale table and car table)
2. In your specification include multiplicity, IDs, IDREFs, required and non-required attributes etc.

Anomalies

Provide an example of update and delete anomalies (provide a table with sample data) over the table with movies and actors.

Normalization

1. Assume set of FDs $F = \{AB \rightarrow CD, AB \rightarrow E, F \rightarrow G\}$ over relation R .
2. What does it mean that relation is in BCNF? (provide definition)
3. Is table R in BCNF? (Provide justification – prove it by closure test or inference test.
4. If answer is NO decompose table R to be in BCNF.