Introduction to Functional Dependencies

by Vladimir Surkov

Rents Example

A Rents table stores information about Customers, Films, and the films a customer rents.

- Rents(CustomerID, Title, Price, Rating, Date)

CustomerID	Title	Price	Rating	Date
0001	Braveheart	4.20	PG13	2003-01-08
0001	The Patriot	3.30	R	2003-01-08
0002	The Patriot	3.30	R	2003-03-03
0003	Ransom	3.50	PG13	2003-04-01

Redundant Information

CustomerID	Title	Price	Rating	Date
0001	Braveheart	4.20	PG13	2003-01-08
0001	The Patriot	3.30	R	2003-01-08
0002	The Patriot	3.30	R	2003-03-03
0003	Ransom	3.50	PG13	2003-04-01

- Title, Price, and Kind is repeated for each film.
- Wastes space
- Potential for inconsistent data is increased
- Cannot store information about a film if no reservation for it exists (a null cannot be stored in the CustomerID)

Anomalies Resulting from A Bad Design

- Update anomaly:

  Update rental price to $4 for 'The Patriot' have to change several tuples

  Anomally if changes in some but not in all

- Insertion anomaly:

  Cannot insert information about a film if it has no rentals

- Deletion anomaly:

  If no rentals, information about the film disappears

Goals of Logical Design

- Develop good relational schemas
- Avoid redundant data
- Avoid modification anomalies
- Ensure that relationships among attributes are represented
- Facilitate the checking of updates for violation of database integrity constraints

Good design important even if ER model was used (although developing with an ER model often yields a good design)

Decomposition

Solution:

Decompose the relation schema Rents into:

`R1 = (CustomerID, Title, Date)`

`R2 = (Title, Price, Kind)`

Requirements:

All attributes of the original scheme (R) must appear in the decomposition

`attributes(R) = attributes(R1)` $\bigcup$ `attributes(R2)`

Lossless-join decomposition:

For all possible relations `r` on scheme `R`

$r = \pi_{R1}(r) \ltimes \pi_{R2}(r)$

How do we ensure that the decomposition is a losslessjoin decomposition?

Next tutorial !!!

Functional Dependencies

- Functional dependencies generalize the previously introduced notions of keys.
- They also allow us to identify possible information loss from a given decomposition.

Definition:

Let α and β be subsets of schema R. A *functional dependency* $\alpha \rightarrow \beta$ holds on R if for all legal instances r of R,

$$\forall t_1, t_2 \in r (t_1[\alpha] = t_2[\alpha] \Rightarrow t_1[\beta] = t_2[\beta])$$

Definition:

A dependency $\alpha \rightarrow \beta$ is trivial if $\beta \subseteq \alpha$

Examples:

- A film has a unique title, rental price, and distributor.

  FilmID $\rightarrow$ Title, RentalPrice, Distributor

- The CustomerID uniquely identifies the customer and his/her address.

  CustomerID $\rightarrow$ Name, Street, City, State

- Each video tape has a unique status.

  FilmID, TapeNum $\rightarrow$ Status

- On any particular day, a video tape can be rented to at most one customer.

  Date, FilmID, TapeNum $\rightarrow$ CustomerID

- A performer can have only one role in a particular film.

  PerformerID, FilmID $\rightarrow$ Role

- A role can be only one role.

  Role $\rightarrow$ Role

  This is a trivial dependency!

Closure

Definition:

Let F be a set of functional dependencies. The closure of F, F^+, consists of all dependencies implied by dependencies in F.

Example:

$R = (A, B, C, D)$

$F = \{A \rightarrow B, A \rightarrow C, CD \rightarrow A\}$

Some members of F^+: $A \rightarrow BC, CD \rightarrow B, AD \rightarrow B, AD \rightarrow ABCD$

Superkey

Definition:

If $\alpha \rightarrow R \in F^+$ then α is a superkey (candidate key, primary key) of R

Example:

SIN LASTNAME $\rightarrow$ SIN LASTNAME FIRSTNAME ADDRESS

SIN LASTNAME is a superkey

Key

Definition:

X is a key, if:

1. X is a superkey, i.e. $X \rightarrow R \in F^+$
2. There is no $Y \subset X$, such that $Y \rightarrow R \in F^+$

Example:

STREET NUMBER CITY $\rightarrow$ STREET NUMBER CITY ZIPCODE

STREET NUMBER CITY is a key

Example 1

Let F be the following set of functional dependencies: $\{AB \rightarrow CD, B \rightarrow DE, C \rightarrow F, E \rightarrow G, A \rightarrow B\}$. Use Armstrong's axioms to show that $\{A \rightarrow FG\}$ is logically implied by F.

Solution:

1	$A \rightarrow B$	Given
2	$A \rightarrow AB$	1, Augmentation
3	$AB \rightarrow CD$	Given
4	$A \rightarrow CD$	2, 3, Transitivity
5	$B \rightarrow DE$	Given
6	$A \rightarrow DE$	1, 5, Transitivity
7	$A \rightarrow ACD$	4, Augmentation
8	$ACD \rightarrow CDE$	6, Augmentation twice
9	$A \rightarrow CDE$	7, 8, Transitivity
10	$A \rightarrow CE$	9, Trivial dependency
11	$C \rightarrow F$	Given
12	$CE \rightarrow FE$	11, Augmentation
13	$E \rightarrow G$	Given
14	$FE \rightarrow FG$	13, Augmentation
15	$CE \rightarrow FG$	12, 14, Transitivity
16	$A \rightarrow FG$	10, 15, Transitivity

Example 2

Consider the relation scheme R(A,B,C,D) with functional dependencies $A \to C$ and $B \to D$.

Is $\{A, B\}$ a key?

Fact: $\{A, B\}$ is a superkey.

Indeed from Armstrong's Axioms we can infer:

$A \to C \Rightarrow AB \to$ ABC (augmentation by AB)

$B \to D \Rightarrow ABC \to$ ABCD (augmentation by ABC)

We obtain $AB \to ABCD$ (transitivity)

$\{A, B\}$ is a candidate key (minimal). We must show that neither $\{A\}$ nor $\{B\}$ alone are candidate keys.

Neither $\{A\}$ nor $\{B\}$ are superkeys since $\{A\}^+ = \{A, C\}$, $\{B\}^+ = \{B, D\}$

(But $\{A, B\}^+ = \{A, B, C, D\}$)

Example 3

Consider a database of ship voyages with the following attributes: S (ship name), T (type of ship), V (voyage identifier), C (cargo), P (port), and D (day). We assume that a voyage consists of a sequence of events where one ship picks up a single cargo, and delivers it to a sequence of ports. A ship can visit only one port in a single day. Thus, the following functional dependencies may be assumed: $S \to T, V \to SC, SD \to P$. Find the key.

Solution:

Step 1: We know that V and D should appear in the key since they do not appear on right side of any functional dependency.

Step 2: Using Armstrong axioms we can infer the following:

$V \to VSC$ (augmentation by V)

$VD \to VSCD$ (augmentation by D)

$VSCD \to TVSCD$ (augmentation by VSCD)

$TVSCD \to TVSCDP$ (augmentation by TVSCD)

Thus $VD \to TVSCDP$ (transitivity)

$\{V, D\}^+ = \{TVSCDP\}$

$\{V, D\}$ is a superkey and it is minimal, i.e. it is a key.

Example 4

Consider a relation with schema R(A, B, C, D) and functional dependencies $F = \{AB \to C, C \to D, D \to A\}$.

a) What are all the nontrivial FD's that follow from the given FD's? You should restrict yourself to FD's with single attributes on the right side.

b) What are all the keys of R?

c) What are all the superkeys of R that are not keys?

Solution:

a) $\{C \to A\}, \{\mathrm{AB} \to D\}$

b) B must appear as an attribute of the key.

Then $\{A, B\}^+ = \{A, B, C, D\}$ and $\{A, B\}$ is a key.

$\{B, C\}^+ = \{A, B, C, D\}$ and $\{B, C\}$ is a key.

$\{B, D\}^+ = \{A, B, C, D\}$ and $\{B, D\}$ is a key.

c) $\{A, B, C\}, \{B, C, D\}, \{A, B, D\}$ and $\{A, B, C, D\}$