Normalisation (Functional Dependencies)

G51DBI – Databases and Interfaces

Yorgos Tzimiropoulos
yorgos.tzimiropoulos@nottingham.ac.uk

This Lecture

- Normalisation
  - Functional Dependencies

Redundancy and Normalisation

- Redundant data
  - Can be determined from other data in the database
  - Leads to various problems
    - INSERT Anomalies
    - UPDATE Anomalies
    - DELETE Anomalies
- Normalisation
  - Aims to reduce data redundancy
  - Redundancy is found/expressed in functional dependencies
  - Normal forms are defined so that they don't contain specific types of functional dependency

Functional Dependencies

- Redundancy is often caused by a functional dependency
- A functional dependency (FD) is a link between two sets of attributes in a relation
- We can normalise a relation by removing undesirable FDs
- A set of attributes, A, functionally determines (fd) another set, B, if whenever two rows of the relation have the same values for all the attributes in A, then they also have the same values for all the attributes in B.
- In this case, we can say there exists a functional dependency between A and B ($A \to B$).

Functional Dependencies

A	B
a1	b1
a2	b2

A->B

If a1 = a2, then

b1 = b2

- A set of attributes, A, functionally determines (fd) another set, B, if whenever two rows of the relation have the same values for all the attributes in A, then they also have the same values for all the attributes in B.
- In this case, we can say there exists a functional dependency between A and B ($A \to B$).

Properties of FDs

In any relation, the primary key fd any set of attributes in that relation

$$K \to X$$

- K is the primary key, X is a set of attributes
- Same for candidate keys

K	X
k1	x1
k2	x2

Proof:
Assume the opposite. Then there is pair of rows such that:
k1 = k2 and x1 != x2

Properties of FDs

In any relation, the primary key fd any set of attributes in that relation

$K \to X$

- K is the primary key, X is a set of attributes
- Same for candidate keys

K	X
k1	x1
k2	x2

Proof:
Assume the opposite. Then there is pair of rows such that:
k1 = k2 and x1 != x2
This implies that the 2 rows are different and that for these rows k1 = k2. This violates **uniqueness property** for primary key

Properties of FDs

In any relation, any set of attributes is FD on itself

$X \to X$

X	X
x1	x1
x2	x2

Proof:
Obviously, if x1 = x2 from left side, then x1 = x2 from the right side

Properties of FDs

Rules for FDs

- Reflexivity: If B is a subset of A then

$A \to B$

A1	A2	A3	A4	A1	A3
a1,1	a2,1	a3,1	a4,1	a1,1	a3,1
a1,2	a2,2	a3,2	a4,2	a1,2	a3,2

Proof:
Without loss of generality assume A = {A1, A2, A3, A4} and B={A1, A3}
Obviously, if the 2 rows of A are equal, then
a1,1 = a1,2 and a2,1 = a2,2 and a3,1 = a3,2 and a4,1 = a4,2
which implies that the 2 rows of B are also equal

Properties of FDs

Rules for FDs

- Augmentation: If $A \to B$

then

$A \cup C \to B \cup C$

A	C	B	C
a1	c1	b1	c1
a2	c2	b2	c2

Proof:
It is given that:
A->B : if a1 = a2 then b1 = b2
C->C : if c1 = c2 then c1 = c2
So if a1 = a2 and c1 = c2, then b1 = b2 and c1 = c2
i.e. AUC -> BUC

Properties of FDs

Rules for FDs

- Transitivity: If $A \to B$ and $B \to C$ then

$A \to C$

A	B	C
a1	b1	c1
a2	b2	c2

Proof:
It is given that A->B so if a1 = a2 then b1 = b2
but because B->C if b1 = b2 then c1 = c2
So a1 = a2 implies c1 = c2
i.e. A -> C

Summary of Properties of FDs

- In any relation
  - The primary key fd any set of attributes in that relation

    $K \to X$

    - K is the primary key, X is a set of attributes
  - Same for candidate keys
  - Any set of attributes is FD on itself

    $X \to X$
- Rules for FDs
  - Reflexivity: If B is a subset of A then

    $A \to B$
  - Augmentation: If $A \to B$ then

    $A \cup C \to B \cup C$
  - Transitivity: If $A \to B$ and $B \to C$ then

    $A \to C$