Functional Dependencies

Session Objectives

- Need for functional Dependencies (FD’s)
- FD and Examples
- Constraints on FD
- Trivial and Nontrivial FD
- Closure of a set of dependencies

Session Outcomes

- At the end of this session, participants will be able to
  - Understand need for Functional Dependencies.
  - Understand FD’s and Example
  - Understand constraint on FD’s
  - Understand Trivial and Nontrivial FD
  - Closure of a set of dependencies

Need for Functional Dependencies

- Entity relational model is a conceptual model meant for communication with end user's.

- Relational schema is internal or physical schema meant for storing data.

- Relational schema can be optimized, for efficient storage and retrieval of data.

- **Functional dependencies are a frame work for systematic design and optimization of relational schemas.**

Functional Dependencies

- **Functional dependencies (FDs)**
  - Are used to specify *formal measures* of the "**goodness**" of relational designs
  - And keys are used to define **normal forms** for relations
  - Are **constraints** that are derived from the ***meaning*** **and** ***interrelationships*** of the data attributes
- **A set of attributes X *functionally determines* a set of attributes Y if the value of X determines a unique value for Y**

Functional Dependencies

- X ->Y holds if whenever two tuples have the same value for X, they must have the same value for Y
- Then **X functionally determines Y or X->Y**
- For any two tuples t1 and t2 in any relation instance r(R):
- If t1[X]=t2[X], then t1[Y]=t2[Y]
- **Left and right sides** of an FD are called the **determinant and the dependent** respectively.
- The determinant and dependent are both sets of attributes.

Functional Dependencies

SUPPLIER_ NUMBER	CITY	PART_ NUMBER	QUANTITY
S1	Bombay	P1	100
S1	Bombay	P2	100
S2	Chennai	P1	200
S2	Chennai	P2	200
S3	Chennai	P2	300
S4	Bombay	P2	400
S4	Bombay	P4	400
S4	Bombay	P5	400

Supplier_number ->city

Functional Dependencies

A functional dependency is a property of the **semantics or meaning** of the attributes – that is how **they relate to one another.**

• Whenever two sets of attributes indicate that a FD should hold, **dependency is specified as a constraint.**

• **Main use of FD:** to describe the relation schema by specifying constraints on its attributes that must **hold at all times.**

For example:
1) SSN --> ename
2) Pnumber --> {Pname, Plocation}
3) {SSN, Pnumber} --> Hours

Functional Dependencies- Constraints

- X determines Y if, and only if, whenever two tuples agree on their X-value, they must necessarily agree on their Y-value.
  - If a constraint on R states that, there **can not be more than one tuple** with a given X-value in any instance
    - r(R) – that is, **X is a candidate key of R.**(since we never have two distinct tuples with t1[x]=t2[x])
  - **If X --> Y in R, does not say whether or not Y --> X in R.**

Functional Dependencies- Constraints

TEACH relation:

FD: Text → Course may exist. However, the FDs
Teacher → Course,
Teacher → Text and Couse → Text are ruled out.

TEACH

Teacher	Course	Text
Smith	Data Structures	Bartram
Smith	Data Management	Martin
Hall	Compilers	Hoffman
Brown	Data Structures	Horowitz

Which FDs may exist?

A relation R(A, B, C, D) with its extension.

Which FDs may exist in this relation?

A	B	C	D
a1	b1	c1	d1
a1	b2	c2	d2
a2	b2	c2	d3
a3	b3	c4	d3

What FDs may exist?

The following FDs may hold

B->C, C->B, {A,B}->D,{A,B}->C,{C,D}->B,{C,D}->A, {A,B,C}->D ,{B,C,D}->A, and so on

The following FDs do not hold

A->B, A->C,A->D

B->A,B->D,C->A,C->D

D->C ,D->A,D->B

{B.C}->D,{B,C}->A

A	B	C	D
a1	b1	c1	d1
a1	b2	c2	d2
a2	b2	c2	d3
a3	b3	c4	d3

What FDs may exist?

Two cases:

Case a: the value of a given relvar at a given point in time.

Case b: the set of all possible values that the given relvar might assume at different time

SUPPLIER_ NUMBER	CITY	PART_ NUMBER	QUANTITY
S1	Bombay	P1	100
S1	Bombay	P2	100
S2	Chennai	P1	200
S2	Chennai	P2	200
S3	Chennai	P2	300
S4	Bombay	P2	400
S4	Bombay	P4	400
S4	Bombay	P5	400

Case a:

Satisfying FDs:

- { SUPPLIER_NUMBER } -> { QUANTITY }
- { SUPPLIER_NUMBER }->CITY
- { QUANTITY } -> { SUPPLIER_NUMBER }
- { QUANTITY }->CITY
- CITY-> { QUANTITY }
- { SUPPLIER_NUMBER, PART_NUMBER } -> { CITY }
- { SUPPLIER_NUMBER, PART_NUMBER } -> {QUANTITY}
- {SUPPLIER_NUMBER, PART_NUMBER }→{CITY, QUANTITY}
- { SUPPLIER_NUMBER, PART_NUMBER } -> { SUPPLIER_NUMBER }

Case a:

- Not interested in FDs that hold for the particular value that the relvar have at some particular time.
- Rather those FDs that hold for ***all possible values* of that relvar.**
- For example, any two tuples appearing in SCP at same time with the same supplier number must necessarily have the same city as well.
- { SUPPLIER_NUMBER, PART_NUMBER } -> { CITY }

Case B:

- Some (time-independent) FDs that apply to SCP:

- { SUPPLIER_NUMBER, PART_NUMBER } -> CITY
- { SUPPLIER_NUMBER, PART_NUMBER }-> QUANTITY
- { SUPPLIER_NUMBER, PART_NUMBER }->{CITY,QUANTITY}
- { SUPPLIER_NUMBER }->CITY
- { SUPPLIER_NUMBER, PART_NUMBER } -> { SUPPLIER_NUMBER }

Following FD do not hold *for all time*:

- SUPPLIER_NUMBER -> QUANTITY
- QUANTITY -> SUPPLIER_NUMBER
- QUANTITY->CITY
- CITY->QUANTITY

Trivial and Nontrivial FD

- An FD is trivial **if and only if the right side is a subset of the left side.**
- { SUPPLIER_NUMBER, PART_NUMBER } -> { SUPPLIER_NUMBER }
- Trivial dependencies are not very interesting in practice.

Closure of a set of dependencies

- Some FDs might imply others:
- {SUPPLIER_NUMBER, PART_NUMBER }->{CITY,QUANTITY}

  implies
- { SUPPLIER_NUMBER, PART_NUMBER } -> CITY
- { SUPPLIER_NUMBER, PART_NUMBER } -> QUANTITY
- **The set of all FDs that are implied by a given set S of FDs is called the closure of S, written S+**
- **Armstrong's axioms** provides set of *inference rules* by which new FDs can be inferred from given ones.

Armstrong Axioms

- Let A, B, and C be arbitrary subsets of the set of attributes of the given relvar R, then

- Some FDs might imply others:

1. Reflexivity: If B is a subset of A, then $A \rightarrow B$.
2. Augmentation: If $A \rightarrow B$, then $AC \rightarrow BC$.
3. Transitivity: If $A \rightarrow B$ and $B \rightarrow C$, then $A \rightarrow C$.

- Several FDs can be derived from three

4. Self-determination: $A \rightarrow A$.
5. Decomposition: If $A \rightarrow BC$, then $A \rightarrow B$ and $A \rightarrow C$.
6. Union: If $A \rightarrow B$ and $A \rightarrow C$, then $A \rightarrow BC$.
7. Composition: If $A \rightarrow B$ and $C \rightarrow D$, then $AC \rightarrow BD$.

Closure of F

- Suppose given a relvar R with attributes A, B, C, D, E, F and the FDs:

  A -> BC, B -> E, CD -> EF

- Show that the FD AD -> F holds for R, and is thus a member of the closure of the given set:

Closure of F

- Suppose given a relvar R with attributes A, B, C, D, E, F and the FDs:

  A -> BC, B -> E, CD -> EF

- Show that the FD AD -> F holds for R, and is thus a member of the closure of the given set:

1.	$A \rightarrow BC$	(given)
2.	$A \rightarrow C$	(decomposition)
3.	$AD \rightarrow CD$	(augmentation)
4.	$CD \rightarrow EF$	(given)
5.	$AD \rightarrow EF$	(transitivity from 3 and 4)
6.	$AD \rightarrow F$	(decomposition)

Closure set of Attributes

- Given a relvar R, a set Z of attributes of R, and a set S of FDs that hold for R

- We can determine the set of all attributes of R that are functionally dependent on Z – the closure Z^+ of Z under S.

- For each such set of attributes Z, determine the set Z^+ of attributes that are functionally determined by Z based on S; Z^+ is called the closure of Z under S

Closure of a set of Attributes

- Suppose given a relvar R with attributes A, B, C, D, E, F and FDs:

  A -> BC, E -> CF, B -> E, CD -> EF

- Compute the closure $\{A,B\}^+$ of the set of attributes {A,B} under the set of FDs :

```
CLOSURE [Z,S] := Z;
do forever ;
      for each FD X → Y in S
            do ;
                  if X ⊆ CLOSURE [Z,S]
                  then CLOSURE [Z,S] := closure [Z,S] ∪ Y;
            end
      if CLOSURE[Z,S] did not change on this iteration then
      leave the loop;
end;
```

Closure of a set of Attributes

- Suppose given a relvar R with attributes A, B, C, D, E, F and FDs:

  A -> BC, E -> CF, B -> E, CD -> EF

- Compute the closure {A,B}+ of the set of attributes {A,B} under the set of FDs :

CLOSURE [Z,S] := {A,B} ;

1. A $\subseteq$ {A, B} then CLOSURE [Z,S] := {A, B, C} ;
2. E $\not\subseteq$ {A, B, C} --- no change
3. B $\subseteq$ {A, B, C} then CLOSURE [Z,S] := {A, B, C, E} ;
4. CD $\not\subseteq$ {A, B, C, E} --- no change
5. A $\subseteq$ {A, B, C, E} ---- right side attributes already in closure
6. E $\subseteq$ {A, B, C, E} then CLOSURE [Z,S] := {A, B, C, E, F} ;
7. B $\subseteq$ {A, B, C, E, F} --- no change
8. CD $\not\subseteq$ {A, B, C, E, F} --- no change

no change in CLOSURE[Z,S] for next iteration; terminate;

therefore Closure of {A,B} is {A,B}+ = {A, B, C, E, F}

Closure of a set of Attributes

- Suppose given a relvar R with attributes A, B, C, D, E, F and FDs:
- A -> BC, E -> CF, B -> E, CD -> EF
- Compute the closure $\{A,D\}^+$ of the set of attributes {A,D} under the set of FDs :
- **closure[z,s]= {A,B,C,D}**
- **E $\not\subset$ {A,B,C,D} ---No change**
- **B $\subseteq$ {A,B,C,D} then closure[z,s]={A,B,C,D,E}**
- **CD $\subseteq$ {A,B,C,D} then closure[z,s]={A,B,C,D,E,F}**
- Since $\{A,D\}^+$ contains all the attributes of relation R then {AD} is called as **Superkey.**
- **K is a *superkey* if and only if the closure K^+ of K – under the given set of FDs – is precisely the set of all attributes of R.**
- Remove the attributes and check for the closure individually if the closure of the attributes does not contain all the attributes then superkey be the **candidate key or key**

Corollary

- Superkey for a given relvar R are subsets K of the attributes of R such that the FD:

  K -> A

  holds true for every attribute A of R

- K is a *superkey* if and only if the closure **K^+ of K** – under the given set of FDs – is precisely the set of all attributes of R.

- K is a candidate key if and only if it is an irreducible superkey

Inference Rules for FDs

- **Closure** of a set F of FDs is the set F^+ of all FDs that can be inferred from F (F+ can be obtained by applying axioms inference rules)
- **Closure of a set of attributes X** with respect to F is the set X^+ of all attributes that are functionally determined by X(X^+ can be calculated by using the FDs in F)

Equivalence of Sets of FD

- Two sets of FDs F and G are equivalent if:
  - Every FD in F can be inferred from G, and
  - Every FD in G can be inferred from F
- **Hence, F and G are equivalent if $F^+ = G^+$**
- Definition (**Covers**):
  - F covers G if every FD in G can be inferred from F (i.e., if $\mathbf{G}^+$ subset-of $\mathbf{F}^+$)
  - F and G are equivalent if F covers G and G covers F

Irreducible Sets of Dependencies

• A set S of FDs to be **irreducible if and only if** it satisfies the following three properties:

1. The right side of every FD in S involves just one attribute.
2. No FD in S can be discarded from S without changing the closure S+
3. The left side of every FD in S is irreducible – no attribute can be discarded from the left side without changing the closure S+
3. **Also called as Minimal cover F for a set of FDs S.**

Irreducible Sets of Dependencies

- A set S of FDs to be **irreducible if and only if** it satisfies the following three properties:

    1. Replace each FD in the canonical form

    Replace each FD X {A1, A2, → …, An} by the n functional dependencies X → A1, X → A2, …. , X → An.

    2. if { {F – {X→A}} is equivalent to F then remove X → A from F

    *removal of redundant FD from F

    3. removal of extraneous attribute B in X.

Minimal Sets of FDs

Given a R with attributes A,B,C,D, and FDs:

1. A -> BC
2. B -> C
3. A -> B
4. AB -> C
5. AC -> D

(1) can be rewritten as A -> B and A -> C

•Now the set of FDs are:

1. A -> B
2. A -> C
3. B -> C
4. AB -> C
5. AC -> D

Minimal Sets of FDs

Given a R with attributes A,B,C,D, and FDs:
A -> C (by 2)
AB->BC (aug B)
AB->C (decomp)

AB->C can be inferred from A -> C so AB->C can be removed

- Now the set of FDs are:

1. A -> B
2. A -> C
3. B -> C
4. AC -> D

Minimal Sets of FDs

Given a R with attributes A,B,C,D, and FDs:
A -> C (by 2)
A->AC (Aug A)
AC->D (given)
A->D(by transitivity)

- Now the set of FDs are:

1. A -> B
2. A -> C
3. B -> C
4. A -> D

Minimal Sets of FDs

Given a R with attributes A,B,C,D, and FDs:

A -> B (by 2)

B->C (given)

A->C(by transitivity)

- Now the set of FDs are:

1. A -> B
2. B -> C
3. A -> D

Summary

- Need for functional Dependencies (FD’s)
- What is FD?
- Constraints on FD
- Examples-FD
- Trivial and Nontrivial FD
- Closure of a set of dependencies
- Closure of a set of fds