

Databaser, modellering och implementering 6/7,5 hp

Lektion nr. 7:

Funktionella Beroenden och Normalisering

I Fundamentals of Database Systems: kapitel 15

I Databasteknik: kapitel 11

Steg i Databas-design

1. Samla fakta om verksamheten, S.K. "fact-finding phase"
 - Textuell-beskrivning över det som databasen skall innefatta
 - [Exempel](#)
2. Omvandla punkt 1. till ett ER-diagram
3. Omvandla punkt 2. till relationer i enlighet med Relationsmodellen
4. Skapa en test-databas, implementeringsmodellen, SQL
5. **Normalisera där det behövs**
6. Gå till punkt 2. om du har normaliserat
7. Skapa slutgiltiga databasen

Design ledtrådar

- Tydlig semantik för attribut

Det bör vara enkelt att förklara meningen med relationsschemat.

- Reducera redundanta värden i tuppler

Basrelationsschemat bör inte orsaka insättnings-, borttagnings- eller uppdateringsavvikelser(anomalier).

- Reducera null värden i tuppler – helst undvik dem

Undvik placering av attribut vilkas värden kan vara NULL i basrelationer.

- Inga falska tupler

Loss-less join egenskap

Funktionellt beroende

Funktionellt beroende definierar sambanden kring attributen i en relation.

Den definierar den effekt som en tillståndsförändring i en uppsättning av attribut orsakar hos en annan uppsättning av attribut.

Vi använder \rightarrow för att markera ett funktionellt beroende mellan en uppsättning av attribut och FD för "Functional Dependency".

Med $X \rightarrow Y$ menas att X funktionellt bestämmer Y.

Funktionellt beroende

- Fullt funktionellt beroende:

$X \rightarrow Y$ är ett fullständigt FD om alla medlemmar av X måste vara närvarande för att FD:et skall hålla.

Y är fullt funktionellt beroende av X

- Delvist funktionellt beroende:

Med ett FD där $X = \{A, B, C\}$ och $Y = \{D\}$ om $\{A, B\} \rightarrow \{D\}$

då Y är partiellt beroende av X .

Funktionellt beroende

<u>Anstnr</u>	<u>projnr</u>	Fnamn	Timmar
1	2	Fred	200
1	20	Fred	100
1	40	Fred	50
2	10	Ben	100
2	20	Ben	80

5 rows in set (0.000 sec)

Giltiga FD:

Anstnr → Fnamn

Anstnr , Projnr → Fnamn, Timmar

NORMALISERING

- "Organisering av data så att dupliceringar minimeras".

Anstnr	Fnamn
1	Fred
2	Ben

2 rows in set (0.000 sec)

- **Målet** är att isolera data så att insättning, borttagning och modifiering av ett fält skall göras endast i en tabell för att hela databasen med de definierade kopplingar skall påverkas.

NORMALISERING

- Elimineras modifikationsavvikelser:
 - Borttagning:
Om en rad tas bort, info. om ett objekt går förlorat.
 - Insättning:
Ett inlägg kan inte göras, innan något annat händer först. (kund & beställning)
 - Uppdatering:
Om ett nytt värde anges till en kolumn, måste detta upprepas för flera rader.

NORMALFORMER

En NF är en (uppsättning av regler) regel som garanterar att vissa anomalier (avvikelser) inte förekommer. Varje NF inkluderar även den föregående normalformen.

- Första NormalFormen (1NF)
- Andra NormalFormen (2NF)
- Tredje NormalFormen (3NF)
- Boyce-Codd NormalFormen (BCNF)
- Fjärde NormalFormen (4NF)
- Domän/Nyckel NormalFormen (DK/NF)
- Femte NormalFormen (5NF)

NORMALFORMER



Mittuniversitetet
MID SWEDEN UNIVERSITY

3:e NF

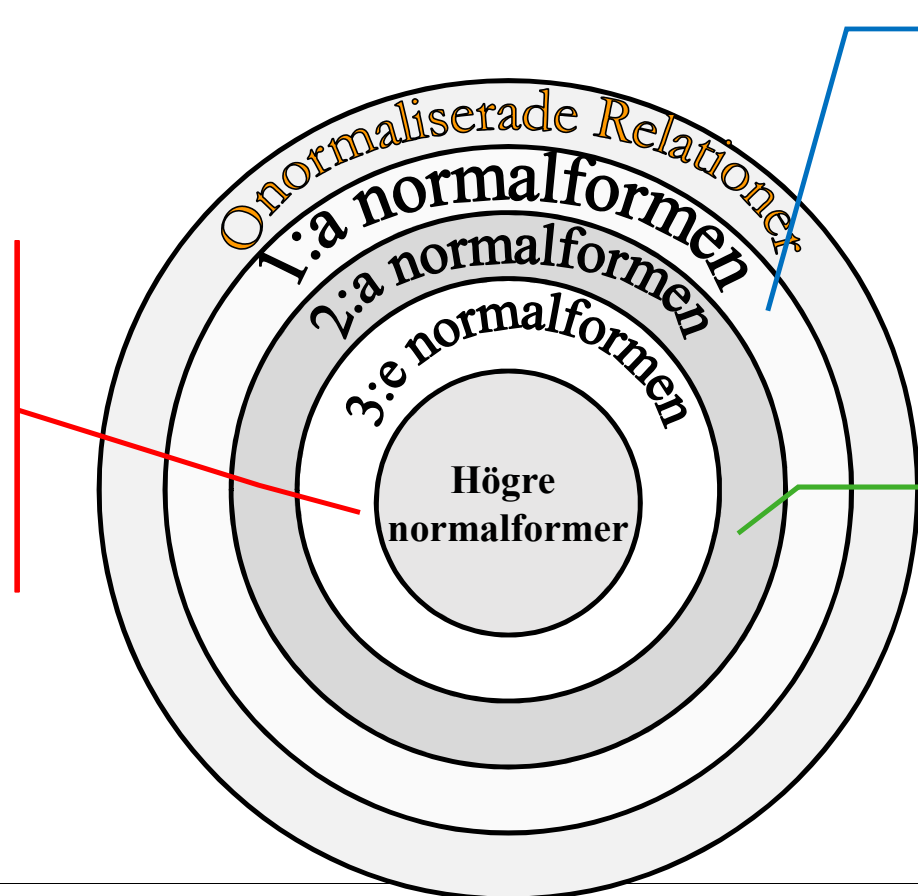
Inga transitiva beroenden
mellan icke-nyckel attribut.

$A \rightarrow B$

$B \rightarrow C$

därför:

$A \rightarrow C$



1:a NF

–Attributen innehåller atomära värden.

Inga upprepningar i innehåll eller struktur

2:a NF

Icke-nyckel attributens fulla funktionella beroende av primärnyckeln.

1NF

Purchase

ReceiptNo	Ordinal	PDate	Customer	Item	Category	Price
1234	1,2	2008-01-21	Fred Klos	Chip Cookie,Sweet Bread	Cookie,Bread	0.95,2.50
3042	1,2	2008-01-23	Susie Glad	Sweet Bread,Opera Cake	Bread,Cookie	15.0,2.50
3403	1	2008-01-26	Susie Glad	Gana Cookie	Cookie	1.30
5612	1	2008-01-26	Dean Dews	Chip Cookie	Cookie	0.95
1564	1,2	2008-01-21	Fred Klos	Gana Cookie,Coffee	Cookie,Drink	1.50,1.30

5 rows in set (0.000 sec)

- Primärnyckel :
ReceiptNo
- Ett funktionellt beroende:

ReceiptNO→Ordinal, PDate, Customer, Item, Category, Price

1NF

Purchase

ReceiptNo	Ordinal_1	Ordinal_2	PDate	Customer	Item_1	Item_2	Category1	Category2	Price_1	Price_2
1234	1	2	2008-01-21	Fred Klos	Chip Cookie	Sweet Bread	Cookie	Bread	0.95	2.5
3042	1	2	2008-01-23	Susie Glad	Sweet Bread	Opera Cake	Bread	NULL	15	2.5
3403	1	NULL	2008-01-26	Susie Glad	Gana Cookie	NULL	Cookie	NULL	1.3	NULL
5612	1	NULL	2008-01-26	Dean Dews	Chip Cookie	NULL	Cookie	NULL	0.95	NULL
1564	1	2	2008-01-21	Fred Klos	Gana Cookie	Coffee	Cookie	Drink	1.5	1.3

5 rows in set (0.000 sec)

- Primärnyckel :
ReceiptNo

1NF = atomära värden



Mittuniversitetet
MID SWEDEN UNIVERSITY

Purchase

ReceiptNo	Ordinal	PDate	Customer_fname	Customer_lname	Item	Category	Price
1234	1	2008-01-21	Fred	Klos	Chip Cookie	Cookie	0.95
1234	2	2008-01-21	Fred	Klos	Sweet Bread	Bread	2.5
1564	1	2008-01-21	Fred	Klos	Gana Cookie	Cookie	1.3
1564	2	2008-01-21	Fred	Klos	Coffee	Drink	1.5
3042	1	2008-01-23	Susie	Glad	Sweet Bread	Bread	2.5
3042	2	2008-01-23	Susie	Glad	Opera Cake	Cookie	15
3403	1	2008-01-26	Susie	Glad	Gana Cookie	Cookie	1.3
5612	1	2008-01-26	Dean	Dews	Chip Cookie	Cookie	0.95

8 rows in set (0.000 sec)

- Primärnyckel :

ReceiptNo, Ordinal

- Några funktionella beroenden:

ReceiptNO, Ordinal → PDate, Customer_fname, Customer_lname, Item, Category, Price

Item → Category, Price

ReceiptNo → Customer_fname, Customer_lname, PDate

1NF eliminera synliga upprepningar



Mittuniversitetet
MID SWEDEN UNIVERSITY

Purchase

ReceiptNo	Ordinal	PDate	CustomerID	Item	CategoryID	Price
1234	1	2008-01-21	1	Chip Cookie	1	0.95
1234	2	2008-01-21	1	Sweet Bread	2	2.5
1564	1	2008-01-21	1	Gana Cookie	1	1.3
1564	2	2008-01-21	1	Coffee	3	1.5
3042	1	2008-01-23	2	Sweet Bread	2	2.5
3042	2	2008-01-23	2	Opera Cake	1	15
3403	1	2008-01-26	2	Gana Cookie	1	1.3
5612	1	2008-01-26	3	Chip Cookie	1	0.95

8 rows in set (0.000 sec)

Category

CategoryID	CategoryName
1	Cookie
2	Bread
3	Drink

3 rows in set (0.001 sec)

CategoryID → CategoryName

- Några giltiga funktionella beroenden för Purchase:

ReceiptNO, Ordinal → PDate, CustomerID, Item, CategoryID, Price

Item → CategoryID, Price

ReceiptNo → CustomerID, PDate

Customer

CustomerID	Customer_fname	Customer_lname
1	Fred	Klos
2	Susie	Glad
3	Dean	Dews

3 rows in set (0.001 sec)

CustomerID → Customer_fname, Customer_lname

OBS: Elimineringen kunde göras senare. Förfarandet tillhör inte 1NF

2NF

Purchase

ReceiptNo	Ordinal	PDate	CustomerID	Item	CategoryID	Price
1234	1	2008-01-21	1	Chip Cookie	1	0.95
1234	2	2008-01-21	1	Sweet Bread	2	2.5
1564	1	2008-01-21	1	Gana Cookie	1	1.3
1564	2	2008-01-21	1	Coffee	3	1.5
3042	1	2008-01-23	2	Sweet Bread	2	2.5
3042	2	2008-01-23	2	Opera Cake	1	15
3403	1	2008-01-26	2	Gana Cookie	1	1.3
5612	1	2008-01-26	3	Chip Cookie	1	0.95

8 rows in set (0.000 sec)

- Några giltiga funktionella beroenden för Purchase:

ReceiptNO, Ordinal → PDate, CustomerID, Item, CategoryID, Price

Item → CategoryID, Price

ReceiptNo → CustomerID, PDate

2NF = Inga partiella beroenden



Mittuniversitetet
MID SWEDEN UNIVERSITY

Purchase

ReceiptNo	Ordinal	Item	CategoryID	Price
1234	1	Chip Cookie	1	0.95
1234	2	Sweet Bread	2	2.5
1564	1	Gana Cookie	1	1.3
1564	2	Coffee	3	1.5
3042	1	Sweet Bread	2	2.5
3042	2	Opera Cake	1	1.5
3403	1	Gana Cookie	1	1.3
5612	1	Chip Cookie	1	0.95

8 rows in set (0.000 sec)

Receipt

ReceiptNo	PDate	CustomerID
1234	2008-01-21	1
1564	2008-01-21	1
3042	2008-01-23	2
3403	2008-01-26	2
5612	2008-01-26	3

5 rows in set (0.000 sec)

ReceiptNo → CustomerID, PDate

- Några giltiga funktionella beroenden för Purchase:

ReceiptNO, Ordinal → Item, CategoryID, Price

Item → CategoryID, Price

3NF

Purchase

ReceiptNo	Ordinal	Item	CategoryID	Price
1234	1	Chip Cookie	1	0.95
1234	2	Sweet Bread	2	2.5
1564	1	Gana Cookie	1	1.3
1564	2	Coffee	3	1.5
3042	1	Sweet Bread	2	2.5
3042	2	Opera Cake	1	15
3403	1	Gana Cookie	1	1.3
5612	1	Chip Cookie	1	0.95

8 rows in set (0.000 sec)

- Några giltiga funktionella beroenden för Purchase:

ReceiptNO, Ordinal → Item, CategoryID, Price

Item → CategoryID, Price

3NF = Inga transitiva beroenden

Purchase

ReceiptNo	Ordinal	Item
1234	1	Chip Cookie
1234	2	Sweet Bread
1564	1	Gana Cookie
1564	2	Coffee
3042	1	Sweet Bread
3042	2	Opera Cake
3403	1	Gana Cookie
5612	1	Chip Cookie

3 rows in set (0.000 sec)

Item

Item	CategoryID	price
Chip cookie	1	0.95
Sweet Bread	2	2.5
Gana Cookie	1	1.3
Coffee	3	1.5
Opera Cake	1	15

5 rows in set (0.000 sec)

Item → CategoryID, Price

- Några giltiga funktionella beroenden för Purchase:

ReceiptNO, Ordinal → Item

Item (namnet till en vara) som primärnyckel är inte ett bra val.
Varför?

3NF. Bättre än förra bilden!!

Purchase

ReceiptNo	Ordinal	ItemID
1234	1	01CH
5612	1	01CH
1234	2	02RB
3042	1	02RB
3042	2	03OC
1564	1	04GC
3403	1	04GC
1564	2	05CO

3 rows in set (0.000 sec)

Item

ItemID	ItemName	CategoryID	Price
01CH	Chip Cookie	1	0.95
02RB	Sweet Bread	2	2.5
03OC	Opera Cake	1	15
04GC	Gana Cookie	1	1.3
05CO	Coffee	3	1.5

5 rows in set (0.001 sec)

Item → ItemName, CategoryID, Price

ReceiptNO, Ordinal → ItemID

ALL IN ONE

Purchase

ReceiptNo	Ordinal	ItemID
1234	1	01CH
5612	1	01CH
1234	2	02RB
3042	1	02RB
3042	2	03OC
1564	1	04GC
3403	1	04GC
1564	2	05CO

3 rows in set (0.000 sec)

Item

ItemID	ItemName	CategoryID	Price
01CH	Chip Cookie	1	0.95
02RB	Sweet Bread	2	2.5
03OC	Opera Cake	1	15
04GC	Gana Cookie	1	1.3
05CO	Coffee	3	1.5

5 rows in set (0.001 sec)

Receipt

ReceiptNo	PDate	CustomerID
1234	2008-01-21	1
1564	2008-01-21	1
3042	2008-01-23	2
3403	2008-01-26	2
5612	2008-01-26	3

5 rows in set (0.000 sec)

Customer

CustomerID	Customer_fname	Customer_lname
1	Fred	Klos
2	Susie	Glad
3	Dean	Dews

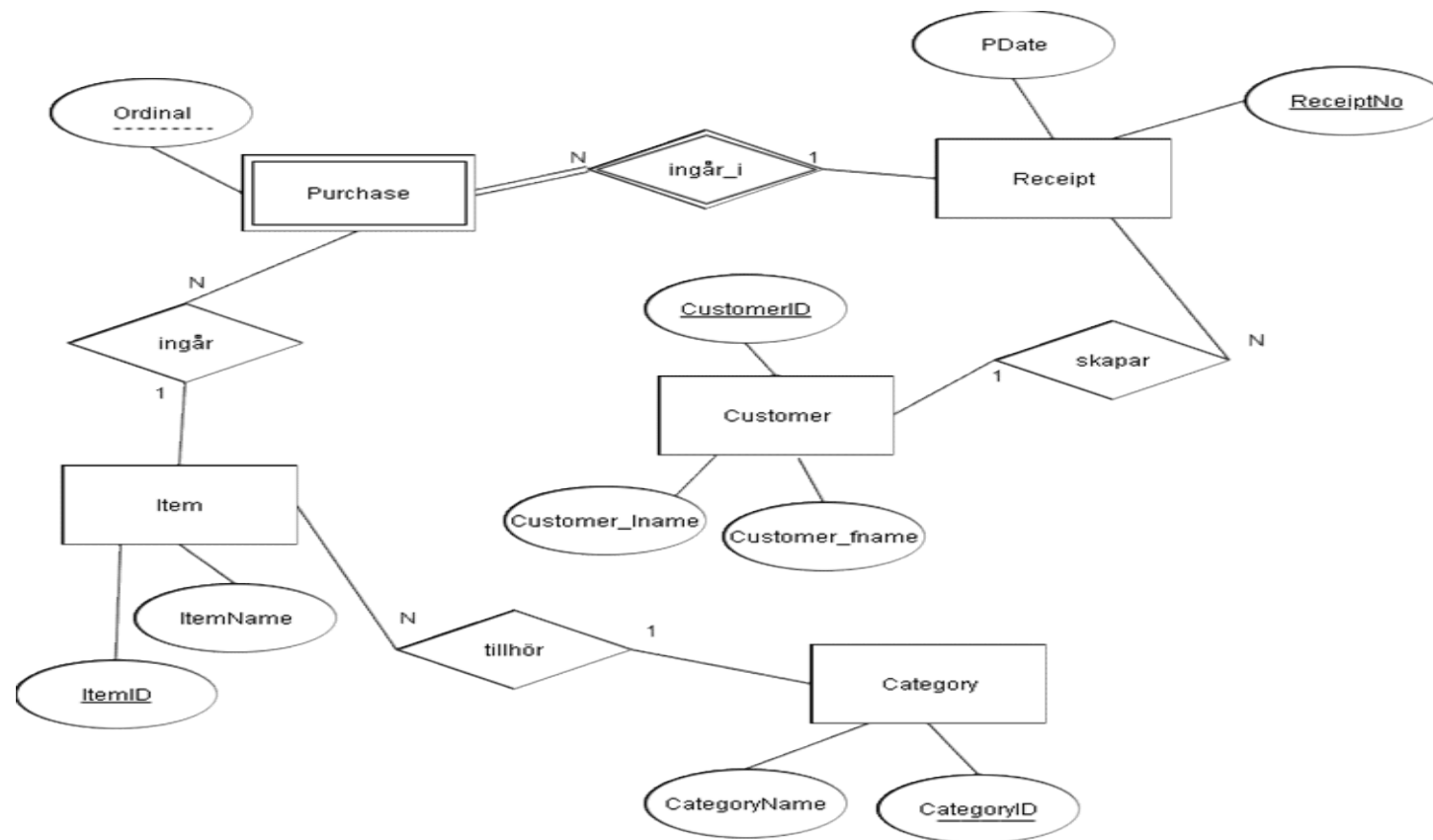
3 rows in set (0.001 sec)

Category

CategoryID	CategoryName
1	Cookie
2	Bread
3	Drink

3 rows in set (0.001 sec)

ER efter normalisering



HEM UPPGIFT

Bilden nedan visar en onormaliserad tabell Teacher.

Din uppgift är:

- 1) Beskriv vilka anomalier finns i tabellen
- 2) Ange vilka NF tabellen inte följer och med minst ett FD, normalisera tabellen upp till 3NF.
- 3) Rita ett ER-diagram som visar kardinalitetsgrader

```
MariaDB [sql_lecture]>
MariaDB [sql_lecture]> select * from Teacher;
```

TeacherNr	TeacherName	DeptCode	RoomNr	DeptManager	DeptFaxNr
B17	B. Brown	ECO	263	D. Davis	523-7441
J12	J. Jones	ACC	243	W. Smith	523-7318
L22	L. Lewis	ACC	248	W. Smith	523-7318
L29	L. Lawless	CIS	222	C. Coles	523-7319
M22	M. Morris	ECO	257	D. Davis	523-7441
W26	W. West	ECO	257	D. Davis	523-7441

5 rows in set (0.000 sec)

Tack