

NORMALISERING

Peter Bellström

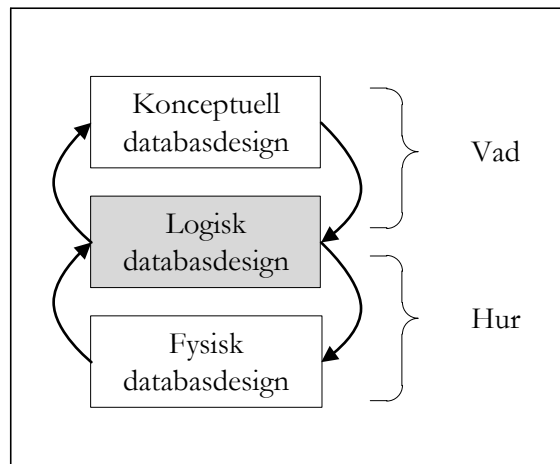
Logisk databasdesign

2

- Arbetssteget vars syfte är att konstruera en modell (diagram, schema), baserad på en specifik datamodell, över verksamhetens begrepp och samband. Modellen skall vara helt oberoende av DBHS och andra fysiska överväganden.

Fokus för dagens föreläsning är...

3



Peter Bellström - Databasdesign

Normalisering

4

- Vad?
 - ▣ Dela upp de "stora" tabellerna vi erhöill efter översättningen från den konceptuella modellen till mindre tabeller.
- Varför?
 - ▣ Undvika redundans
 - ▣ Upprätthålla integritet
 - ▣ Undvika andra fallgropar
 - ▣ Tumregler för att ta fram en korrekt databas
- Reversibel process
- Nycklar

Peter Bellström - Databasdesign

Funktionella beroenden

5

- Integritetsregel som i praktiken tillämpas som primärnyckeln = önskvärda.
- Om värdet i attribut A, i tabellen T, bestämmer värdet på attribut B, i tabellen T, så är attribut B funktionellt beroende av attribut A.
- $A \rightarrow B$
- Personnummer \rightarrow Efternamn
- Determinant \rightarrow Dependent

Peter Bellström - Databasdesign

Transitiva beroenden

6

- Uppstår när det finns funktionella beroenden inne i en tabell = icke önskvärda.
- Om $A \rightarrow B$ och $B \rightarrow C$ så $A \rightarrow C$
- Om Personnummer \rightarrow Postnummer och Postnummer \rightarrow Ort så Personnummer \rightarrow Ort

Peter Bellström - Databasdesign

Relation vs. Tabell

7

- En relation är alltid i 1 NF vilket en tabell inte alltid är.
- En relation har inga dubletter.
- En relation har en unik identifierare!
- I en relation har ordningen ingen betydelse varken vad det gäller tupler (rader) eller attribut (kolumner).
- I en relation är alla attributvärden odelbara.

Peter Bellström - Databasdesign

Första normalformen (1NF)

8

- Definition
 - ▣ En relationsvariabel (tabell) är i första normalformen om och endast om varje attribut innehåller **ett** värde.
- Tillämpning
 - ▣ Ett värde i varje attribut
 - ▣ Fungerande primärnyckel
- Minnesregel
 - ▣ Nyckeln!

Peter Bellström - Databasdesign

Första normalformen (1NF)

- Relationsvariabel (tabell) som **inte** uppfyller första normalformen.
- Konsekvenser?

<u>Knr</u>	<u>Telnr</u>	<u>Namn</u>	<u>Adress</u>
K1	1234, 1010	Anna Ek	Stigen 3 117 65 Stockholm
K2	2425	Bertil Bo	Byg. 18 653 51 Karlstad

Peter Bellström - Databasdesign

Första normalformen (1NF)

- Relationsvariabel (tabell) som **uppfyller** första normalformen.
- Lösningen?

<u>Knr</u>	<u>Telnr</u>	<u>ENamn</u>	<u>FNAmn</u>	<u>Gadr</u>	<u>Postnr</u>	<u>Padr</u>
K1	1234	Ek	Anna	Stigen 3	117 65	Stockholm
K1	1010	Ek	Anna	Stigen 3	117 65	Stockholm
K2	2425	Bo	Bertil	Byg. 18	653 51	Karlstad

Peter Bellström - Databasdesign

Andra normalformen (2NF)

11

- Definition
 - ▣ En relationsvariabel (tabell) är i andra normalformen om och endast om den är i första normalformen och varje icke-nyckelattribut är beroende av primärnyckeln.
- Tillämpning
 - ▣ Endast vid sammansatta primärnycklar.
- Minnesregel
 - ▣ Hela Nyckeln!

Peter Bellström - Databasdesign

Andra normalformen (2NF)

- Relationsvariabel (tabell) som *inte* uppfyller andra normalformen.
- Konsekvenser?

<u>Knr</u>	<u>Telnr</u>	ENamn	FNAmn	Gadr	Postnr	Padr
K1	1234	Ek	Anna	Stigen 3	117 65	Stockholm
K1	1010	Ek	Anna	Stigen 3	117 65	Stockholm
K2	2425	Bo	Bertil	Byg. 18	653 51	Karlstad

Peter Bellström - Databasdesign

Andra normalformen (2NF)

13

- Relationsvariabel (tabell) som **uppfyller** andra normalformen.
- Lösningen?

<u>Knr</u>	<u>ENamn</u>	<u>FNamn</u>	<u>Gadr</u>	<u>Postnr</u>	<u>Padr</u>
K1	Ek	Anna	Stigen 3	117 65	Stockholm
K2	Bo	Bertil	Byg. 18	653 51	Karlstad

<u>Knr*</u>	<u>Telnr</u>
K1	1234
K1	1010
K2	2425

Peter Bellström - Databasdesign

Tredje normalformen (3NF)

14

- Definition
 - En relationsvariabel (tabell) är i tredje normalformen om och endast om den är i andra normalformen och varje icke-nyckelattribut inte är transitivt beroende av primärnyckeln.
- Tillämpning
 - Eliminera inre beroenden
- Minnesregel
 - Ingenting annat än nyckeln!

Peter Bellström - Databasdesign

Tredje normalformen (3NF)

- Relationsvariabel (tabell) som **inte** uppfyller tredje normalformen.
- Konsekvenser?

<u>Knr</u>	ENamn	FNamn	Gadr	Postnr	Padr
K1	Ek	Anna	Stigen 3	117 65	Stockholm
K2	Bo	Bertil	Byg. 18	653 51	Karlstad
K3	Al	Sven	Gösv. 3	653 51	Karlstad

Peter Bellström - Databasdesign

Tredje normalformen (3NF)

- Relationsvariabel (tabell) som **uppfyller** tredje normalformen.
- Lösningen?

<u>Knr</u>	ENamn	FNamn	Gadr	Postnr*
K1	Ek	Anna	Stigen 3	117 65
K2	Bo	Bertil	Byg. 18	653 51
K3	Al	Sven	Gösv. 3	653 51

<u>Postnr</u>	Padr
117 65	Stockholm
653 51	Karlstad

Peter Bellström - Databasdesign

Minnesregel 1NF – 3NF

17

- Nyckeln (1NF), hela nyckeln (2NF) och ingenting annat än nyckeln (3NF).

Peter Bellström - Databasdesign

Boyce/Codd normalform (BCNF)

- Definition
 - En relationsvariabel (tabell) är i BCNF om och endast om varje *determinant* är en *kandidatnyckel*.
- Tillämpning
 - Vid två eller flera kandidatnycklar.
 - Sammansatta kandidatnycklar.
 - Överlappande kandidatnycklar.

Boyce/Codd normalform (BCNF)

19

- Relationsvariabel som **inte** uppfyller Boyce/Codd normalform.
 - Tabell som lagrar längden på svenska gator.
 - Gatunamn är unika i varje stad, men inte i hela Sverige.
 - Varje gata ligger inom ett och samma postnummerområde.
 - Det kan dock finnas flera postnummerområden i en och samma ort.
 - Ursprungstabellen består av: Gatunamn, Postnummer, Ortsnamn, Längd
 - Kandidatnycklar:
 - Postnummer & Gatunamn, Gatunamn & Ortsnamn
 - Men vad händer med Postnummer som bestämmer Ortsnamn?
 - Postnummer är med andra ord en determinant men inte en kandidatnyckel (jmf. definitionen av BCNF)

Peter Bellström - Databasdesign

Boyce/Codd normalform (BCNF)

- Relationsvariabel som **inte** uppfyller Boyce/Codd normalform.
- Konsekvenser?

Gatunamn	Postnummer	Ortsnamn	Längd
Päronvägen	65463	Karlstad	20 km
Päronvägen	12345	Kil	5 km
Tomatvägen	65468	Karlstad	25 km
Bananvägen	65463	Karlstad	20 km

Boyce/Codd normalform (BCNF)

- Relationsvariabel som **uppfyller** Boyce/Codd normalform.
- Lösningen?

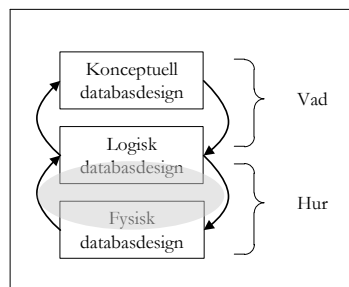
<u>Gatunamn</u>	<u>Postnummer</u>	Längd
Päronvägen	65463	20 km
Päronvägen	12345	5 km
Tomatvägen	65468	25 km
Bananvägen	65463	20 km

<u>Postnummer</u>	<u>Ortsnamn</u>
65463	Karlstad
12345	Kil
65468	Karlstad

Fysisk databasdesign

22

- När tabellerna är normaliserade är vi klara med logisk databasdesign.
- Nästa uppgift i databasdesignprocessen är att översätta (mappa) de normaliserade tabellerna till SQL.



Peter Bellström - Databasdesign