

Databasteori 2

Översättning från ER-diagram till tabeller & lite om normalformer

Anders Jensen-Urstad

2017-01-30

Förra föreläsningen

- ☞ ER-modellen: vilka saker som finns och hur de hänger ihop
- ☞ Relationsmodellen: tabeller, kolumner, rader, nycklar

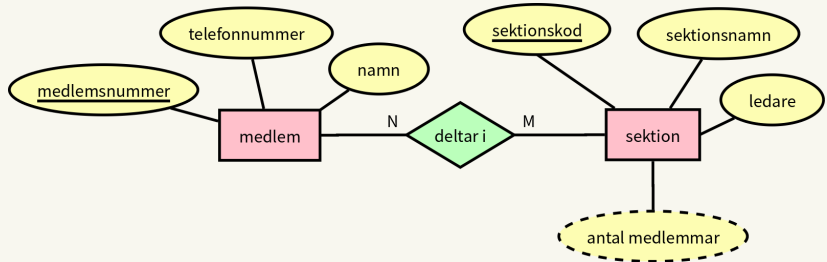
I dag

☞ Repetition

☞ Översättning från ER-modellen till relationsmodellen (tabeller)

☞ Normalisering & normalformer

Repetition: ER-diagram



Repetition: relationsmodellen

Teori	Vanligen	FileMaker
Relation	Tabell	Tabell
Attribut	Kolumn	Fält
Tupel	Rad	Post

Primärnyckel identifierar en specifik rad.

Referensattribut pekar på en primärnyckel i en annan tabell.

(Filemaker använder termen relation för att prata om primärnyckel/referensattribut. Inte samma sak som databasteorins relation!)

Sportklubben, medlemstabell

Tabell: Medlem

Medlemsnummer	Namn	Telefonnummer
17	Tyrion	098424
18	Arya	078445
19	Daenerys	082347

Medlemsnummer är *primärnyckel* i tabellen Medlem.

Sportklubben: sektionstabell

Tabell: Sektion

Sektionskod	Namn	Ledare
A	Schack	18
B	Fäktning	18
C	Eldslukning	19

Sektionskod är primärnyckel i tabellen Sektion.
Ledare är ett *referensattribut* som hänvisar till primärnyckeln i tabellen Medlem.

Sportklubben: tabell för många-till-många

Tabell: Deltar

Medlem	Sektion
25	A
25	B
28	C
28	A
23	B

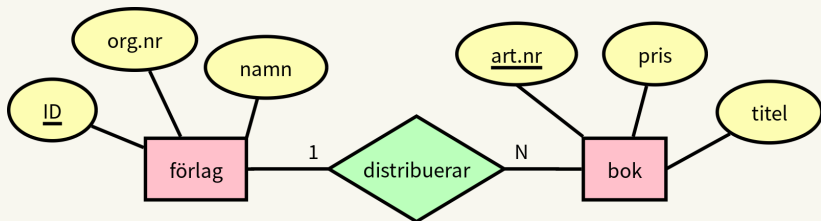
Medlem och Sektion är i tabellen Deltar båda *referensattribut* och hänvisar till *primärnycklarna* i tabellerna Medlem respektive Sektion.

De utgör tillsammans enda *kandidatnyckeln* (minsta möjliga nyckel) och är således primärnyckel i tabellen Deltar.

Översättning från ER-modellen till relationsmodellen

- ☞ Vanliga entitetstyper blir tabeller
- ☞ Vanliga attribut blir kolumner
- ☞ Sambandstyp blir referensattribut i ena tabellen, *eller* egen tabell, beroende på kardinalitet

1:N-samband (ett-till-många)



Ett förlag kan ha många böcker.

En bok kan bara ha ett förlag.

Översättning till tabeller: 1:N (ett-till-många)

Tabell: Förlag

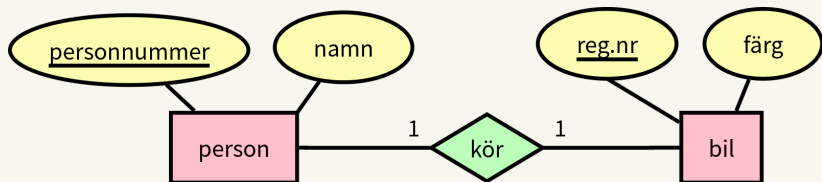
Namn	Org.nr	ID
Harper	143534-0142	1
Penguin Books	783534-0984	2
Hellman and Rogers	228539-5943	3

Tabell: Bok

Art.nr	Titel	Pris	Förlag
1	The Dispossessed	99,00	1
2	Cats and the Law	79,00	3
3	The Third Policeman	79,00	2

Sambandet blir referensattribut i “många”-entitetstypens tabell.

1:1-samband (ett-till-ett)



Ett-till-ett-samband (1:1).

Varje person kan bara köra *en* bil.

Varje bil kan bara ha en förare.

Översättning till tabeller: 1:1

Tabell: Person

Namn	Personnummer	Kör bilen
Dana	640223-7361	ABC123
Fox	611013-9213	FAR451

Tabell: Bil

Bil	Färg
ABC123	svart
FAH451	silver

Referensattribut i tabellen Person. *Eller* tvärtom! Se nästa sida...

Översättning till tabeller: 1:1

Alternativt:

Tabell: Person

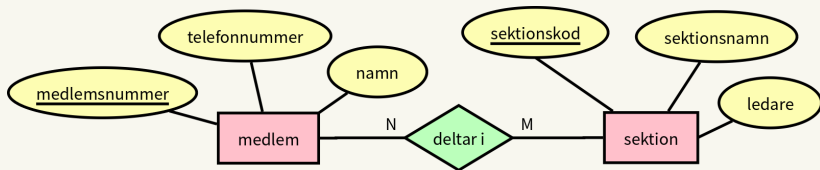
Namn	Personnummer
Dana	640223-7361
Fox	611013-9213

Tabell: Bil

Bil	Färg	Körs av
ABC123	svart	640223-7361
FAH451	silver	611013-9213

...referensattribut i tabellen Bil.

Översättning till tabeller: N:M



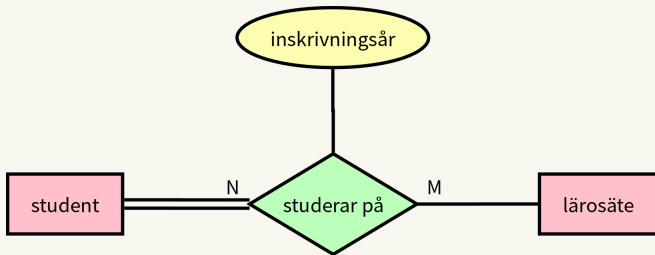
N:M-sambandstypen blir en egen tabell, *deltar*, med två referensattribut som pekar på primärnycklarna (medlemsnummer och sektionskod) i de andra tabellerna.
(Se sportsklubbstabellerna ovan.)

Översättning till tabeller: attribut på sambandstyper

Sambandstypernas attribut blir kolumner:

- 🦄 1:1- och 1:N-sambandstyper: kolumn i den tabell som har referensattributet.
- 🦄 N:M-sambandstyp: kolumn i den särskilda sambandstabellen.

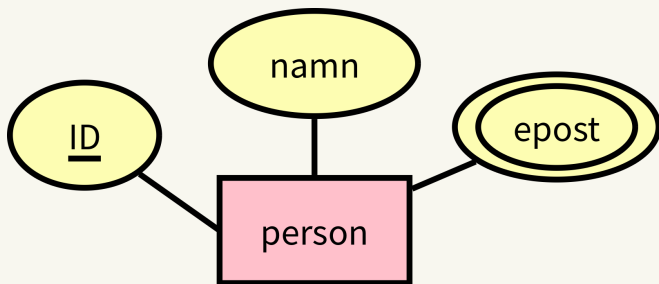
Attribut på sambandstyp i N:M



Tabell: Studerar på

Personnummer	Lärosäte	Inskrivningsår
850523-4666	3	2015
790901-4777	4	2012

Översättning till tabeller: flervärda attribut



Varje flervärd attribut blir en egen tabell ("attributtabell") med personens ID som referensattribut. (Se nästa slide!)

Översättning till tabeller: flervärda attribut

Tabell: Person

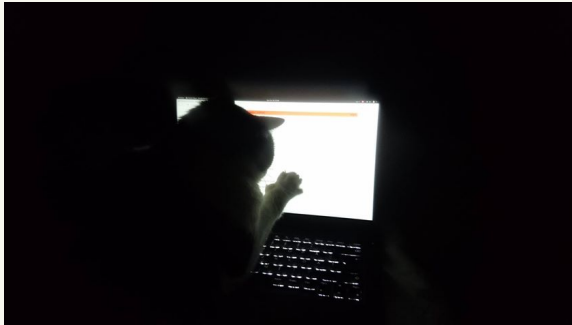
ID	Namn
45	Brienne
57	Sandor

Tabell: Epost

ID	Epost
45	brienne@westeros.org
45	oathkeeper@gmail.com
57	hound@gmail.com

Översättning till tabeller: härledda attribut

Ska *inte* finnas i själva tabellerna/schemat!



Utfyllnadsbild

Paus



Normalisering

- ☞ En *process* för att designa bra relationsdatabaser
- ☞ ...eller omvänt: för att undvika “dum design”
- ☞ Dumt: redundans (att lagra saker i onödan)
- ☞ Dumt: om vissa saker inte *går* att lagra!
- ☞ ”En typ av sak per tabell, och en sån sak per rad”
- ☞ Normalformer: olika regler att följa för att det ska bli bra

Första normalformen (1NF)

Måste finnas primärnyckel, och varje ruta får bara innehålla *ett* värde.

Tabell: uppfyller inte 1NF:

Medlemsnummer	Namn	Telefonnummer
17	Tyrion	098424
18	Arya	078445, 497535
19	Daenerys	082347

Första normalformen (1NF)

Tabell: uppfyller 1NF:

Medlemsnummer	Namn	Telefonnummer
17	Tyrion	098424
18	Arya	078445,
18	Arya	497535
19	Daenerys	082347

Dock fortfarande inte bra. 2NF to the rescue! →

Andra normalformen (2NF)

1NF + varje kolumn som inte är en nyckel måste vara beroende av varje *hel* kandidatnyckel.

Kom ihåg:

Nyckel = kolumn, eller kombination av kolumner, som unikt identifierar en rad.

Kandidatnyckel = nyckel utan onödiga kolumner (dvs alla nyckelns kolumner krävs för att garantera unikheter).
Kan finnas flera! En väljs till primärnyckel.

Andra normalformen (2NF)

Tabell: uppfyller inte 2NF:

Kurskod	Termin	Typ	Kursnamn
MU-507	VT2017	Campus	Nihilism för nybörjare
MU-507	HT2017	Campus	Nihilism för nybörjare
MU-508	HT2017	Distans	Eskatologi A
MU-508	VT2018	Campus	Eskatologi A
MU-509	VT2017	Distans	Valyriska C

{Kurskod, Termin} är primärnyckel, men kursnamn är bara beroende av kurskod. Alltså inte 2NF. Kursnamn behöver brytas ut!

Andra normalformen (2NF)

Kurskod	Termin	Typ
MU-507	VT2017	Campus
MU-507	HT2017	Campus
MU-508	HT2017	Distans
MU-508	VT2018	Campus
MU-509	VT2017	Distans

Kurskod	Kursnamn
MU-507	Nihilism för nybörjare
MU-508	Eskatologi A
MU-509	Valyriska C

Bättre! 2NF uppfyllt.

Tredje normalformen (3NF)

2NF + varje kolumn som inte är nyckel får *bara* vara beroende av nyckeln.

Kurskod	Termin	Typ	LärareID	LärareNamn
MU-507	VT2017	Campus	145	Daria Morgendorffer
MU-507	HT2017	Campus	145	Daria Morgendorffer
MU-508	HT2017	Distans	207	Malaclypse the Younger
MU-508	VT2018	Campus	207	Malaclypse the Younger
MU-509	VT2017	Distans	382	Maester Aemon

LärareNamn är beroende av LärareID – uppfyller inte 3NF!

Tredje normalformen (3NF)

Kurskod	Termin	Typ	LärareID
MU-507	VT2017	Campus	145
MU-507	HT2017	Campus	145
MU-508	HT2017	Distans	207
MU-508	VT2018	Campus	207
MU-509	VT2017	Distans	382

LärareID	LärareNamn
145	Daria Morgendorffer
207	Malaclypse the Younger
382	Maester Aemon

3NF uppfyllt!

Normalisering: sammanfattning

- ☞ 1NF: Primärnyckel ska finnas + bara ett värde per kolumn
- ☞ 2NF: 1NF + varje kolumn som inte är nyckel får bara vara beroende av *hela* nyckeln
- ☞ 3NF: 2NF + varje kolumn som inte är nyckel får *bara* vara beroende av nyckeln

“Data should depend on the key [1NF], the whole key [2NF], and nothing but the key [3NF].”

Alltid normalisera?

☞ Ibland inte: t.ex. adressregister

☞ Påverkar prestandan

☞ Använd sunt förnuft; men motivera om du *inte* normaliserar

ID	Namn	Gatuadress	Postnummer	Ort
1	Sniff	Storgatan 52	881 30	Sollefteå
2	Too-ticki	Kungsvägen 24	914 32	Nordmaling
3	Dronten Edward	Strandgatan 8	881 30	Sollefteå

Läsanvisningar

- ☞ Padron-McCarthy & Risch (2005): kap. 6, 11
- ☞ Webbkursen: “Översättning från ER-modellen till relationsmodellen”, “Normalformer och normalisering”
- ☞ Bonus! Se <https://anders.unix.se/db17> för tips om FileMaker och skapande av ER-diagram.

Laboration 2

- 🦄 Vidareutveckling av sportklubben
- 🦄 Manus och skript
- 🦄 Portaler och relaterade fält

