

# IN2090 – Databaser og datamodellering

## 07 – Funksjonelle avhengigheter

Leif Harald Karlsen (Evgenij Thorstensen)

leifhka@ifi.uio.no



Universitetet i Oslo

# Funksjonell avhengighet

- ◆ Et attributt  $A$  er **funksjonelt avhengig** av en mengde attributter  $X$  hvis det bare kan finnes en verdi av  $A$  for hver mengde verdier av attributtene i  $X$ .
- ◆ Det skrives  $X \rightarrow A$ , og en slik formel kalles en funksjonell avhengighet (FD).
- ◆ For eksempel er Karakter funksjonelt avhengig av  $\{\text{Brnavn}, \text{Kurskode}\}$  i Karakter-tabellen:

Brnavn	Kurskode	Kara
evgenit	IN2090	B
peternl	IN2090	A
evgenit	IN2080	B
leifhka	IN2090	B
leifhka	IN3110	C

- ◆ Og både Navn, Etternavn og Adresse er funksjonelt avhengig av Brnavn i Student-tabellen:

Brnavn	Navn	Etternavn	Adresse
evgenit	Evgenij	Thorstensen	Addr1
peternl	Petter	Nilsen	Addr2
leifhka	Leif H.	Karlsen	Addr3

# FDer, data og virkeligheten

---

- ◆ FDer uttrykker det vi mener er sant i virkeligheten som dataene våre beskriver
- ◆ F.eks. er brukernavnet til en student faktisk unikt for hver student, mens adressen kanskje ikke trenger å være det
- ◆ Kan fort bli et komplisert spørsmål om verdens tilstand
- ◆ FDer forteller oss hvilke data hører sammen, og hva de hører til

# Syntaks for FDer

---

- ◆ Jeg leser ofte pilen som "bestemmer", så

$$X \rightarrow Y$$

leses enten "X bestemmer Y" eller "Y er funksjonelt avhengig av X"

- ◆ Vi dropper ofte mengde-tegnene i FDer, så skriver f.eks. i stedet for

$$\{\text{Brnavn, Kurskode}\} \rightarrow \{\text{Karakter}\}$$

skriver vi ofte

$$\text{Brnavn, Kurskode} \rightarrow \text{Karakter}$$

- ◆ Dersom attributtene er enkle bokstaver (A, B, osv.) dropper vi ofte også komma og skriver f.eks. i stedet for:

$$A, B \rightarrow X, Y, Z$$

skriver vi ofte

$$AB \rightarrow XYZ$$

# FDers oppførsel

---

- ◆ Vi kan samle opp høyresider i FDer, og skrive

$$X \rightarrow A, B$$

dersom vi både har  $X \rightarrow A$  og  $X \rightarrow B$ .

- ◆ FDer er transitive: Hvis  $X \rightarrow Y$  og  $Y \rightarrow Z$ , så har vi at  $X \rightarrow Z$ .
- ◆ En FD  $X \rightarrow Y$  hvor  $Y \subseteq X$  kalles triviell, f.eks.:

$$\text{Brnavn, navn} \rightarrow \text{navn}$$

- ◆ Vi ignorerer slike trivielle FDer, fordi de alltid er sanne og dermed ikke gir oss noe informasjon

# Eksempel, FDer

---

R(Brnavn, Navn, Etternavn, Adresse, Kurskode, Tittel, Beskrivelse, AntSP, Karakter)

Jeg foreslår følgende FDer:

- ◆ Brnavn  $\rightarrow$  Navn, Etternavn, Adresse
- ◆ Kurskode  $\rightarrow$  Tittel, Beskrivelse, AntSP
- ◆ Brnavn, Kurskode  $\rightarrow$  Karakter

- ◆ En **supernøkkel** for en relasjon er jo enhver mengde attributter som sammen er unike for relasjonen
- ◆ En **kandidatnøkkel** er en  $\subseteq$ -minimal supernøkkel
- ◆ Dersom en mengde attributter er unike forekommer jo hver kombinasjon av disse kun i et tuppel, og bestemmer derfor de andre verdiene i tuplet
- ◆ Med andre ord, en nøkkel (enten super eller kandidat) er en mengde attributter som bestemmer de andre attributtene i relasjonen
- ◆ FDer sier jo hvilke attributter som bestemmer hvilke andre attributter
- ◆ Altså, FDene sier hvilke supernøkler og kandidatnøkler vi har!

# FDer og nøkler

---

- ◆ Dersom  $R$  er en relasjon med attributter  $X$ , så vil:
  - ◆  $Y \subseteq X$  være en supernøkkel for  $R$  hvis  $Y \rightarrow X \setminus Y$ , som er equivalent med  $Y \rightarrow X$
  - ◆  $Y \subseteq X$  er en kandidatnøkkel for  $R$  hvis  $Y$  er en minimal supernøkkel
- ◆ For å sjekke om  $X$  er en supernøkkel, sjekk om alt er avhengig av  $X$
- ◆ Altså, bruk FDene og finn alle attributter som er avhengige av  $X$ , de som er avhengige av disse igjen, osv.



- ◆ **Tillukningen**  $X^+$  av  $X$  på en mengde FDer er mengden attributter som er funksjonelt avhengige av  $X$
- ◆ Hvis  $X \rightarrow A$ , så er  $A \in X^+$  sant
- ◆ Hvis  $A \notin X^+$ , så er ikke  $X \rightarrow A$  sant
- ◆ Tillukningen kan regnes ut ved å bruke FDene om og om igjen:
  - ◆ sett  $X^+ = X$
  - ◆ sålenge  $X^+$  forandres:
    - ◆ finn en FD  $Y \rightarrow Z$  med  $Y \subseteq X^+$
    - ◆ sett  $X^+ = X^+ \cup Z$

# Eksempel tillukning

---

Gitt følgende FDer:

- ◆  $\text{Brnavn} \rightarrow \text{Navn, Adresse}$
- ◆  $\text{Kurskode} \rightarrow \text{Grad}$
- ◆  $\text{Brnavn, Kurskode} \rightarrow \text{Karakter}$
- ◆  $\text{Grad, Karakter} \rightarrow \text{Bestått}$

Så har vi følgende tillukninger:

- ◆  $\text{Brnavn}^+ = \text{Brnavn, Navn, Adresse}$
- ◆  $\{\text{Brnavn, Kurskode}\}^+ = \text{Brnavn, Kurskode, Navn, Adresse, Grad, Karakter, Bestått}$
- ◆  $\text{Navn}^+ = \text{Navn}$
- ◆  $\text{Grad}^+ = \text{Grad}$

# Finne kandidatnøkler

---

- ◆ Vi må sjekke alle delmengder av attributter, nedenfra. Men, følgende to regler hjelper oss:
  - ◆ Hvis  $A$  ikke forekommer i noen høyreside, er  $A$  med i **alle** kandidatnøkler.
  - ◆ Hvis  $A$  forekommer i minst en høyreside, men ingen venstresider, er  $A$  **ikke del** av noen kandidatnøkkel.
- ◆ Så begynn med alle attributter som ikke forekommer på høyre side. Beregn tillukningen.
- ◆ Hvis alle attributter er med, sjekk minimalitet. Hvis ikke, utvid i tur og orden med ett og ett nytt attributt.

## Eksempel (lett)

---

R(Brnavn, Navn, Etternavn, Adresse, Kurskode, Tittel, Beskrivelse, AntSP, Karakter)

- ◆ Brnavn → Navn, Etternavn, Adresse
- ◆ Kurskode → Tittel, Beskrivelse, AntSP
- ◆ Brnavn, Kurskode → Karakter

Attributter som ikke er på høyresider: Brnavn, Kurskode

Attributter som er i høyresider, men ikke venstre: Alle andre!

Ergo er {Brnavn, Kurskode} eneste kandidatnøkkel.

## Eksempel (litt mer komplisert)

---

$R(\text{Brnavn, Navn, Adresse, Kurskode, Tittel, Beskrivelse, AntSP, Karakter, Bestått})$

- ◆  $\text{Brnavn} \rightarrow \text{Navn, Adresse}$
- ◆  $\text{Kurskode} \rightarrow \text{Tittel, Beskrivelse, AntSP}$
- ◆  $\text{Tittel} \rightarrow \text{Kurskode, Beskrivelse, AntSP}$
- ◆  $\text{Brnavn, Kurskode} \rightarrow \text{Karakter}$
- ◆  $\text{Karakter} \rightarrow \text{Bestått}$

Ikke på høyresider:  $\text{Brnavn}$

Kun på høyresider:  $\text{Navn, Adresse, Beskrivelse, AntSP, Bestått}$

Forsøke å utvide med:  $\text{Kurskode, Tittel, Karakter}$

$X = \text{Brnavn}$

$X^+ = \text{Brnavn, Navn, Adresse}$

## Eksempel (litt mer komplisert)

---

$R(\text{Brnavn, Navn, Adresse, Kurskode, Tittel, Beskrivelse, AntSP, Karakter, Bestått})$

- ◆  $\text{Brnavn} \rightarrow \text{Navn, Adresse}$
- ◆  $\text{Kurskode} \rightarrow \text{Tittel, Beskrivelse, AntSP}$
- ◆  $\text{Tittel} \rightarrow \text{Kurskode, Beskrivelse, AntSP}$
- ◆  $\text{Brnavn, Kurskode} \rightarrow \text{Karakter}$
- ◆  $\text{Karakter} \rightarrow \text{Bestått}$

Ikke på høyresider:  $\text{Brnavn}$

Kun på høyresider:  $\text{Navn, Adresse, Beskrivelse, AntSP, Bestått}$

Forsøke å utvide med:  $\text{Kurskode, Tittel, Karakter}$

$X = \text{Brnavn, Kurskode}$      $X^+ = \text{Brnavn, Kurskode, Navn, Adresse, Tittel, Beskrivelse, AntSP, Karakter, Bestått}$

Kandidatnøkler:  $\{\text{Brnavn, Kurskode}\}$

## Eksempel (litt mer komplisert)

---

$R(\text{Brnavn, Navn, Adresse, Kurskode, Tittel, Beskrivelse, AntSP, Karakter, Bestått})$

- ◆  $\text{Brnavn} \rightarrow \text{Navn, Adresse}$
- ◆  $\text{Kurskode} \rightarrow \text{Tittel, Beskrivelse, AntSP}$
- ◆  $\text{Tittel} \rightarrow \text{Kurskode, Beskrivelse, AntSP}$
- ◆  $\text{Brnavn, Kurskode} \rightarrow \text{Karakter}$
- ◆  $\text{Karakter} \rightarrow \text{Bestått}$

Ikke på høyresider:  $\text{Brnavn}$

Kun på høyresider:  $\text{Navn, Adresse, Beskrivelse, AntSP, Bestått}$

Forsøke å utvide med:  $\text{Kurskode, Tittel, Karakter}$

$X = \text{Brnavn, Tittel} \quad X^+ = \text{Brnavn, Tittel, Navn, Adresse, Kurskode, Beskrivelse, AntSP, Karakter, Bestått}$

Kandidatnøkler:  $\{\text{Brnavn, Kurskode}\}, \{\text{Brnavn, Tittel}\}$

## Eksempel (litt mer komplisert)

---

$R(\text{Brnavn, Navn, Adresse, Kurskode, Tittel, Beskrivelse, AntSP, Karakter, Bestått})$

- ◆  $\text{Brnavn} \rightarrow \text{Navn, Adresse}$
- ◆  $\text{Kurskode} \rightarrow \text{Tittel, Beskrivelse, AntSP}$
- ◆  $\text{Tittel} \rightarrow \text{Kurskode, Beskrivelse, AntSP}$
- ◆  $\text{Brnavn, Kurskode} \rightarrow \text{Karakter}$
- ◆  $\text{Karakter} \rightarrow \text{Bestått}$

Ikke på høyresider:  $\text{Brnavn}$

Kun på høyresider:  $\text{Navn, Adresse, Beskrivelse, AntSP, Bestått}$

Forsøke å utvide med:  $\text{Kurskode, Tittel, Karakter}$

$X = \text{Brnavn, Karakter}$      $X^+ = \text{Brnavn, Karakter, Navn, Adresse, Bestått}$

**Kandidatnøkler:**  $\{\text{Brnavn, Kurskode}\}, \{\text{Brnavn, Tittel}\} \leftarrow$  alle kandidatnøklerne for  $R$



## Oppsummering så langt

---

- ◆ Skjemaer som er dårlig designet inneholder anomalier
- ◆ Som regel skyldes dette at ikke-relatert informasjon er i samme tabell
- ◆ FDer sier hvilken informasjon som henger sammen, samt hvilke nøkler tabeller har
- ◆ FDer og nøkler gir oss dermed det vi trenger for å spesifisere kriterier for når vi får anomalier og ikke
- ◆ Disse kriteriene definerer ulike normalformer

Takk for nå!

---

Neste video vil handle om normalformer.