Relasjonsdatabasedesign

Normalformer

Hvordan dekomponere tapsfritt

Fagins teorem

Gitt et relasjonsskjema R(XYZ) med FDer F.
En dekomposisjon D={XY, YZ} er tapsfri mhp. F hvis og bare hvis minst en av følgende holder:

1) $Y \rightarrow X \in F^+$
2) $Y \rightarrow Z \in F^+$

Her er F^+ mengden av alle FDer som kan avledes av F

<u>Eksempel</u>: R(A,B,C,D), F={$C \rightarrow AD$}
Dekomposisjon: R_1(A,C,D), R_2(B,C)

Normalformer

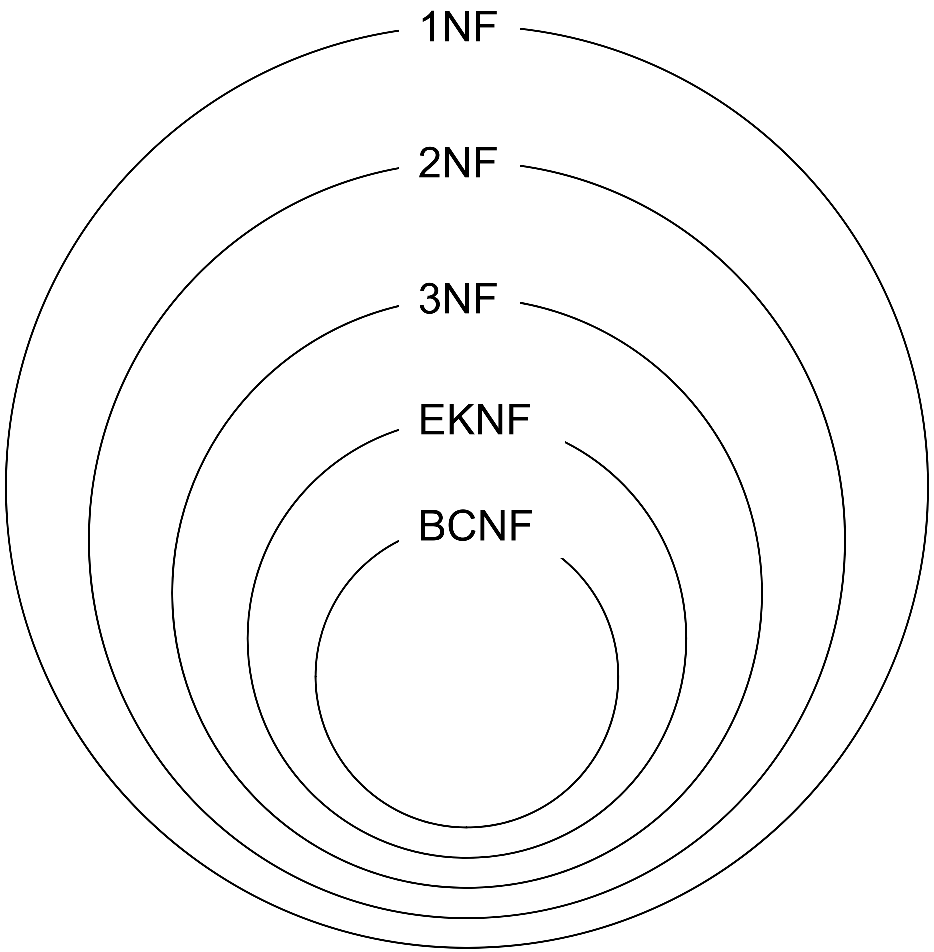
- Normalformer er et uttrykk for hvor godt vi har lykkes i en dekomposisjon
- Jo høyere normalform, jo færre oppdateringsanomalier
- Det fins algoritmer for å omforme fra lavere til høyere normalformer

Utgangspunkt for normalformene 1NF-BCNF

- Alle integritetsregler er i form av FDer

  (i tillegg til domeneskranker og fremmednøkler)

Normalformer, oversikt



Første normalform

- **Definisjon 1NF** (Codd 1972):
  - Alle domener består av atomære verdier
  - Funksjonsverdien til et tuppel for et gitt attributt skal være en slik atomær verdi (eller nil)
- Alle relasjoner er automatisk på 1NF

Andre normalform

- Repetisjon: En FD $X \rightarrow Y$ er ikketriviell hvis Y ikke er inneholdt i X, dvs. hvis $Y - X \neq \emptyset$
- **Definisjon 2NF** (Codd 1972):
  En relasjon R er på **andre normalform** hvis alle ikketrivielle FDer i R på formen $X \rightarrow A$, der X er en mengde attributter og A et attributt i R, tilfredsstiller minst ett av følgende:
  i. X er en supernøkkel i R
  ii. A er et nøkkelattributt i R
  iii. $X \not\subset K$ for noen kandidatnøkkel K i R

Egenskaper ved 2NF

- 2NF$\subseteq$1NF siden alle relasjoner er 1NF
- Når er en relasjon 1NF, men ikke 2NF?

  Svar: Når det fins en ikketriviell FD X$\rightarrow$A og en kandidatnøkkel K hvor X$\subset$K, X $\neq$ K og A ikke er et nøkkelattributt

  Eksempel: R(A,B,C,D), F={BC$\rightarrow$D, C$\rightarrow$A}
  Dekomposisjon: R_1(A,<u>C</u>), R_2(<u>B,C</u>,D)
- 2NF er mest av historisk interesse;
  vi gjør sjelden feil som bryter 2NF

Tredje normalform

- **Definisjon 3NF** (Codd 1972):
  En relasjon R er på **tredje normalform** hvis alle ikketrivielle FDer i R på formen $X \rightarrow A$ tilfredsstiller minst ett av følgende:
  i. X er en supernøkkel i R
  ii. A er et nøkkelattributt i R

Egenskaper ved 3NF

- 3NF$\subseteq$2NF fordi kravene til 3NF er en skjerping av kravene til 2NF
- Når er en relasjon 2NF, men ikke 3NF?

  <u>Svar</u>: Når det fins en ikketriviell FD $X \rightarrow A$ hvor X ikke er en supernøkkel og A ikke er et nøkkel-attributt og $X \not\subseteq K$ for noen kandidatnøkkel K

  <u>Eksempel</u>: R(A,B,C), F={$A \rightarrow BC$, $B \rightarrow C$}
  Dekomposisjon: $R_1(\underline{A},B)$, $R_2(\underline{B},C)$
- Også 3NF er lett å oppnå

Elementære FDer og nøkler

- En FD $X \rightarrow A$ kalles *elementær* dersom
  - A er ett attributt
  - $X \rightarrow A$ er ikketriviell
  - X er minimal
    (dvs. at hvis $Y \subseteq X$ og $Y \rightarrow A$, så er $Y=X$)
- En kandidatnøkkel K kalles *elementær* hvis det finnes en elementær FD $K \rightarrow B$ i R

Elementary Key Normal Form

- **EKNF** (Zaniolo 1982) (kursorisk pensum)
- **Definisjon EKNF**: En relasjon R er på **elementary key normal form** hvis alle ikketrivielle FDer i R på formen $X \rightarrow A$ tilfredsstiller minst ett av følgende:
  i. X er en supernøkkel i R
  ii. A er et attributt i en elementær kandidatnøkkel i R

Egenskaper ved EKNF

- EKNF$\subseteq$3NF fordi kravene til EKNF er en skjerping av kravene til 3NF
- Når er en relasjon 3NF, men ikke EKNF?

  Svar: Når det fins en ikketriviell FD $X \rightarrow A$ hvor X ikke er en supernøkkel og A er et attributt i en ikke-elementær kandidatnøkkel

  Eksempel: R(A,B,C), F={$C \rightarrow A, A \rightarrow C$}
  (Her kan A=epostadresse, B=emnekode og C=fødselsnummer på studenter)

  Mulig dekomposisjon: $R_1(\underline{A,B})$, $R_2(\underline{A},\underline{C})$

Boyce-Codd Normalform

- **Definisjon BCNF** (Boyce & Codd 1974):
  En relasjon R er på **Boyce-Codd normalform** hvis alle ikketrivielle FDer i R på formen $X \rightarrow A$ tilfredsstiller følgende:
  i. X er en supernøkkel i R

Egenskaper ved BCNF

- BCNF$\subseteq$EKNF fordi kravene til BCNF er en skjerping av kravene til EKNF

- Når er en relasjon EKNF, men ikke BCNF?
- Svar: Når det fins en ikketriviell FD $X \rightarrow A$ hvor X ikke er en supernøkkel og A er et attributt i en elementær kandidatnøkkel

Egenskaper ved BCNF (forts.)

- Eksempel: $R(A,B,C)$, $F=\{AB \rightarrow C, C \rightarrow A\}$
  Dekomposisjon til BCNF: $R_1(\underline{B,C})$, $R_2(A,\underline{C})$
- Merk: FDen $AB \rightarrow C$ kan ikke sjekkes i R_1 eller R_2 alene: Vi må ta en join av R_1 og R_2 før vi kan teste om $AB \rightarrow C$.
- Hvis alle FDer etter dekomposisjonen kan sjekkes lokalt i en av de nye relasjonene, sier vi at dekomposisjonen er **FD-bevarende**.
  - I eksempelet over er dekomposisjonen **ikke FD-bevarende**.
- Man kan alltid dekomponere til BCNF, men ikke alltid FD-bevarende.
- Man kan alltid dekomponere FD-bevarende til EKNF.

Eksempel: Studenter

Student(Stud#, Brukernavn, Emnekode, Karakter)

FDer: Stud# $\rightarrow$ Brukernavn

Brukernavn $\rightarrow$ Stud#

Stud# Emnekode $\rightarrow$ Karakter

Hvilke av normalformene 1NF-BCNF tilfredsstiller denne?

Hvorfor BCNF ikke alltid er gunstig

- Dersom vi har en dekomposisjon som ikke er FD-bevarende, må vi foreta en join mellom relasjonsskjemaene i forbindelse med oppdateringer for å sjekke at integritetsreglene overholdes
- Vi har valget mellom

  Enten: Gå tilbake til EKNF (3NF) og bryte BCNF og godta dobbeltlagring som vi må ta hensyn til ved oppdateringer slik at FDene overholdes

  Eller: Beholde BCNF og foreta join mellom relasjoner for å sjekke at FDer overholdes ved oppdateringer

Oppsummering normalisering

- Brudd på normalformer gir opphav til dobbeltlagring
  – Jo høyere normalform, desto mindre dobbeltlagring
- BCNF: Alle FDer er i form av kandidatnøkler.
  Dvs. ingen dobbeltlagring *innen relasjonsskjemaet*.
  Men: Noen FDer kan gå på tvers av relasjonsskjemaer
- EKNF (3NF): Ingen FDer går på tvers av relasjonsskjemaer, men det kan være noen FDer innen et relasjonsskjema som gir opphav til dobbeltlagring
- Krav til dekomposisjon: Den må være ***tapsfri***, dvs. aldri kunne gi opphav til falske tupler
  – Det er en fordel om dekomposisjonen også er ***FD-bevarende*** (dvs. at ingen FDer går på tvers av relasjonsskjemaer)

Det fins algoritmer som…

- Tester om en dekomposisjon er tapsfri (Chasealgoritmen)
- Tester om en dekomposisjon er FD-bevarende
- Dekomponerer relasjonsskjemaer tapsfritt og FD-bevarende
  - Garanterer EKNF, men gir ikke alltid BCNF
- Dekomponerer relasjonsskjemaer tapsfritt til BCNF
  - Er ikke alltid FD-bevarende

Tapsfri dekomposisjon til BCNF

Gitt en relasjon R med et sett FDer F

1. Hvis $X \rightarrow Y$ er et brudd på BCNF:
   1. Beregn X^+
   2. Dekomponer R i R_1 og R_2 der R_1 er lik X^+ og R_2 er lik X samt alle attributtene i R som ikke er i X^+
2. Fortsett på samme måte med R_1 og R_2 inntil alle relasjonene i dekomposisjonen tilfredsstiller BCNF.

Eksempel: Studenter

Student(Stud#, Brukernavn, Emnekode, Karakter)

FDer: Stud# $\rightarrow$ Brukernavn

Brukernavn $\rightarrow$ Stud#

Stud# Emnekode $\rightarrow$ Karakter

Hvordan dekomponere denne til BCNF?

Minimale overdekninger

La F være en mengde FDer over relasjon R.
En **minimal overdekning** for F er en mengde FDer G som er ekvivalent med F og som tilfredsstiller følgende krav:

- Alle høyresidene i G er atomære
- Venstresidene er minst mulige
- Ingen av FDene i G er overflødige

Algoritme for å finne minimale overdekninger

1. Initialiser G til F.
2. Lag atomære høyresider ved hjelp av splitting.
3. Gjør venstresidene minimale:
   For hver $X \rightarrow A$ i G og hver $B \in X$:
   1. Beregn $(X\text{-}B)^+$ med hensyn på G.
   2. Hvis $A \in (X\text{-}B)^+$, erstatt $X \rightarrow A$ med $(X\text{-}B) \rightarrow A$ i G.
4. Fjern overflødige FDer:
   For hver $X \rightarrow A$ i G:
   1. Beregn X^+ med hensyn på G uten å bruke $X \rightarrow A$.
   2. Hvis $A \in X^+$, fjern $X \rightarrow A$ fra G.

Tapsfri FD-bevarende dekomposisjon til EKNF

Gitt en relasjon R med et sett FDer F.

1. Finn en minimal overdekning G for F.
2. For hver X som opptrer som venstreside i G, finn alle $X \rightarrow A_i$ i G og lag en relasjon R_X som inneholder X samt alle høyresidene A_i.
3. Hvis det finnes attributter som ikke er med i noen R_X, lag en egen relasjon R_0 med disse.
4. Hvis hverken R_0 eller noen R_X er en supernøkkel, utvid R_0 til en kandidatnøkkel.

Eksempel

- R(A,B,C,D,E,F,G,H)
- FDer: $AB \rightarrow CH$, $DE \rightarrow FG$, $D \rightarrow H$, $H \rightarrow G$
- Hva er høyeste normalform som R tilfredsstiller?
- Dekomponer R tapsfritt og FD-bevarende til EKNF.

Høyere normalformer

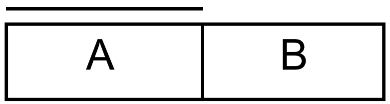
- Det finnes relasjoner som tilfredsstiller BCNF, men likevel kan gi oppdateringsanomalier.
- Dette skyldes blant annet flerverdiavhengigheter (MVDer), som er en generalisering av FDer.
- Høyere normalformer kan eliminere også noen av disse oppdateringsanomaliene.

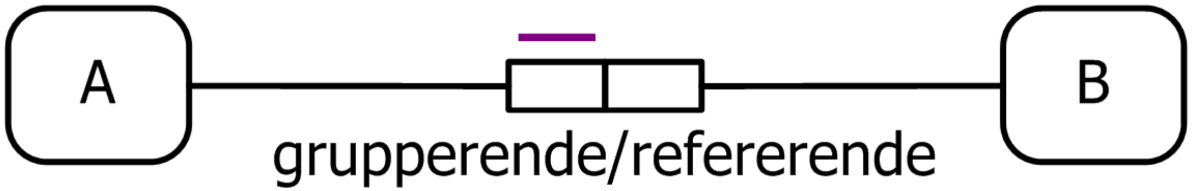
Ekstramateriale og repetisjon fra INF1300 (kursorisk pensum)

ORM og normalisering

- Gruppering av ”korrekte” ORM-diagrammer gir **alltid** 3NF
- Gruppering av ”korrekte” ORM-diagrammer gir **som regel** BCNF. Unntak skyldes alltid problemer knyttet til ekvivalente stier.
  – Det samme gjelder EKNF: Gruppering der EKNF er brutt, skyldes alltid problemer knyttet til ekvivalente stier.

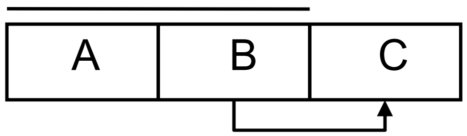
ORM og 1NF

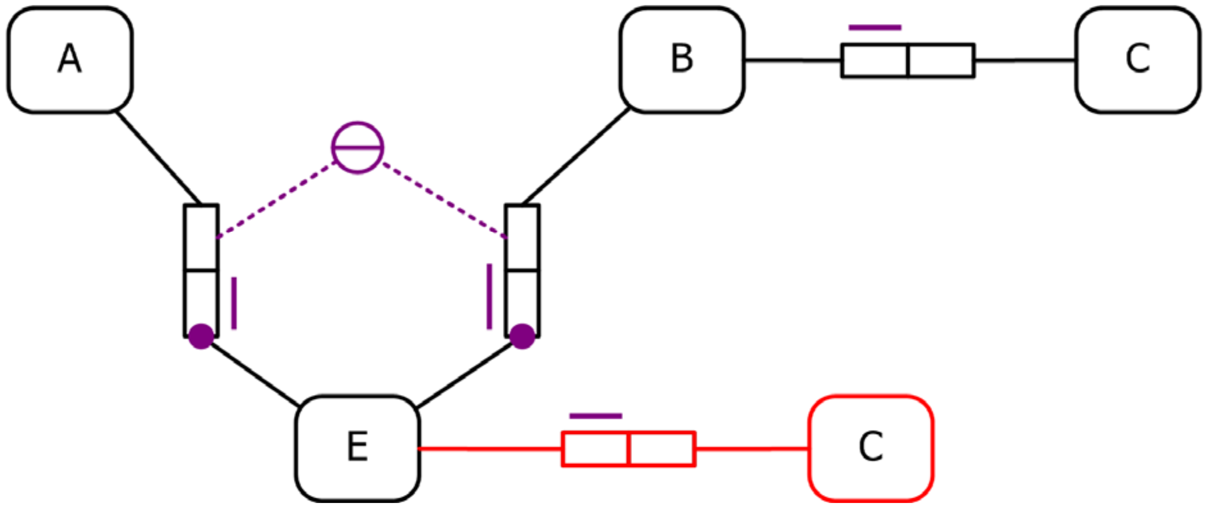




Den grupperende rollen er alltid entydig.
Datamodellen blir alltid 1NF.

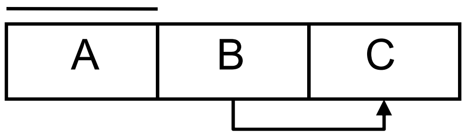
ORM og 2NF

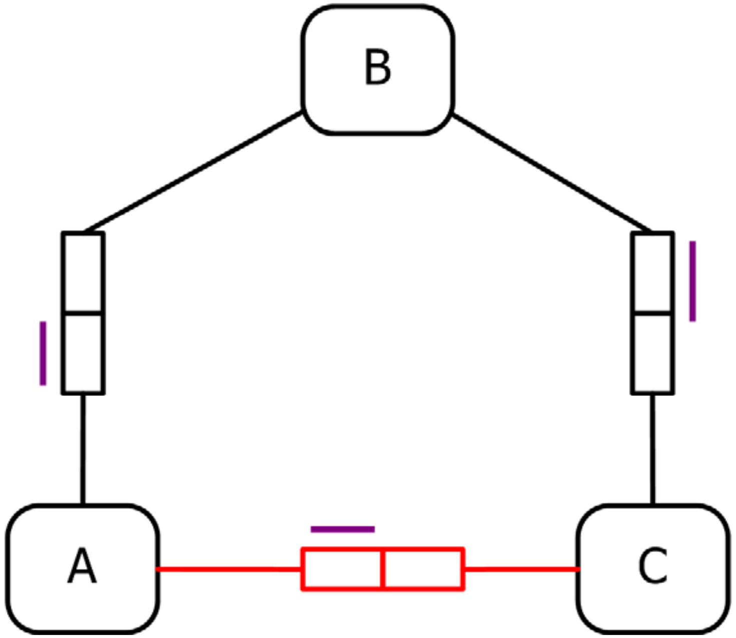




Hvis B→C plasseres på E også, vil 2NF bli brutt.

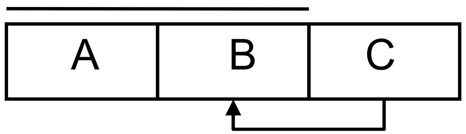
ORM og 3NF

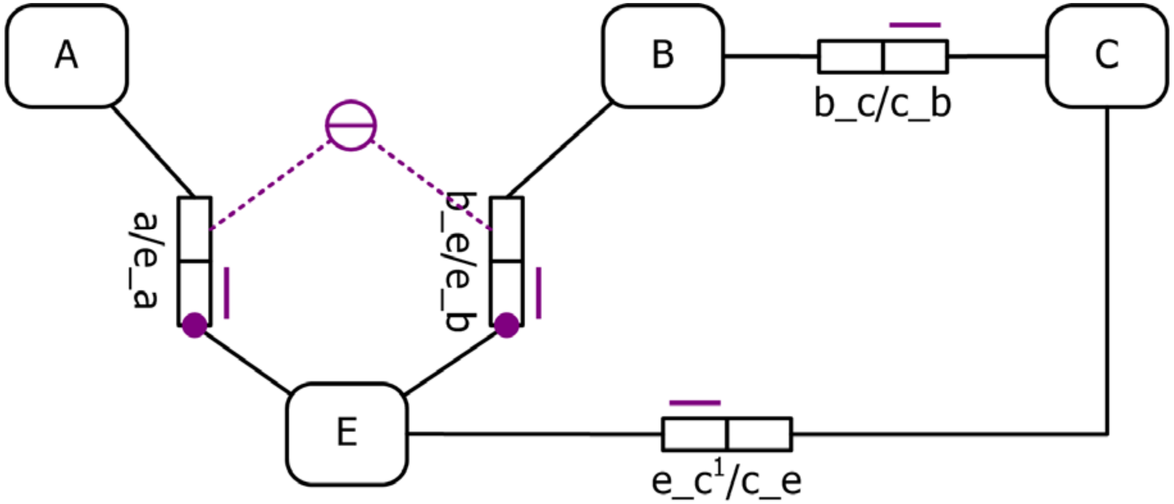




Hvis A→C representeres i tillegg til
A→B, B→C, vil 3NF bli brutt

ORM og BCNF

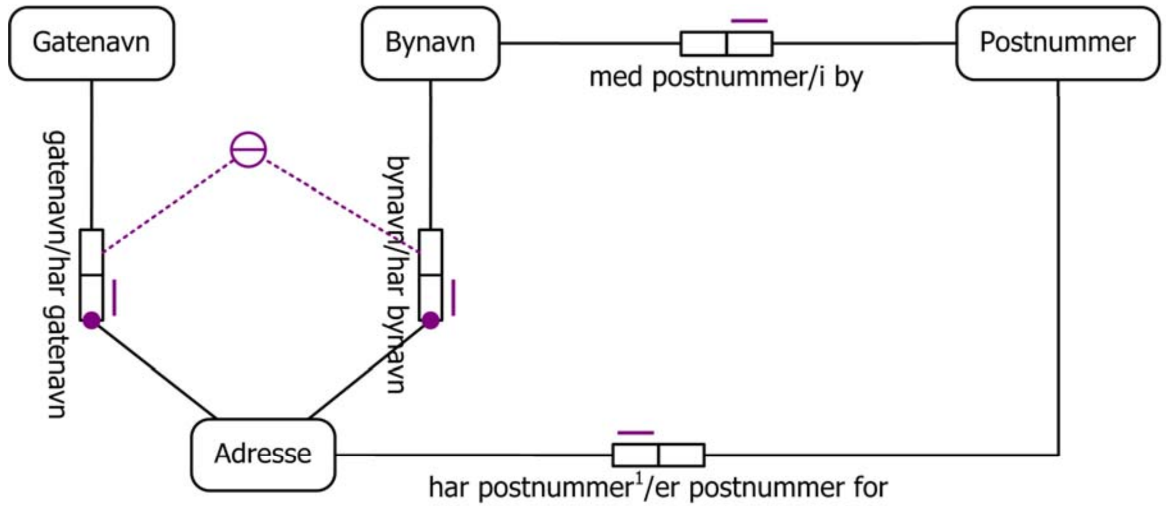




[1]**if an** E e_c C c_b **a** B
then that E e_b **that** B

ORM og BCNF – eksempel

gatenavn	bynavn	postnummer



[1]**hvis en** Adresse *har postnummer* **et** Postnummer *i by* **et** Bynavn
så skal det være slik at samme Adresse *har bynavn* **samme** Bynavn

Gruppert ORM-modell:

gatenavn	bynavn	postnummer

postnummer	bynavn

pluss en ”ekkel” integritetsregel som ivaretar sjekk av de ekvivalente stiene