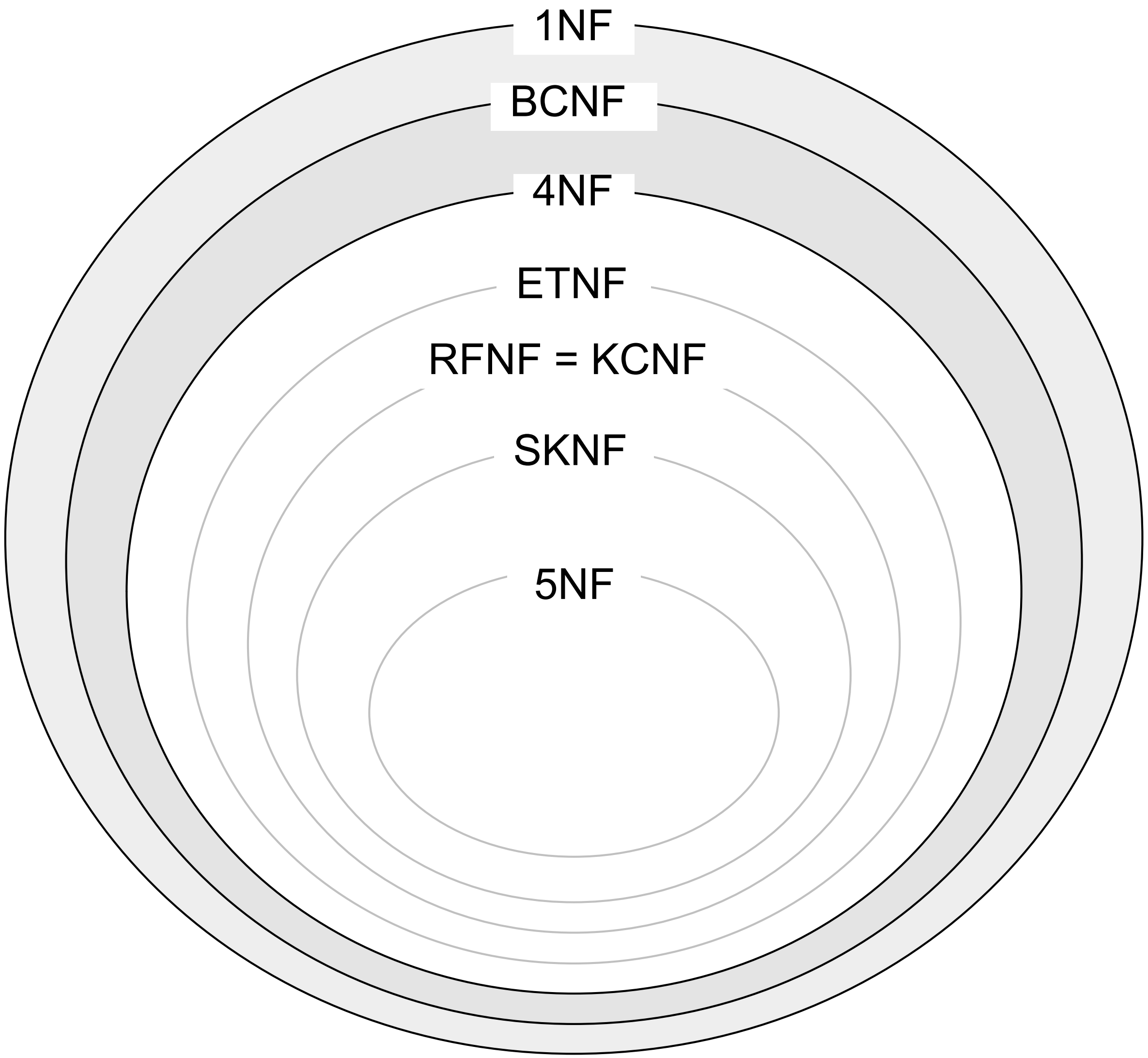
UNIVERSITETET
I OSLO

Normalformer utover 4NF (ikke pensum)

Høyere normalformer, oversikt



Normalformer utover 4NF

- **ETNF** – Essential Tuple Normal Form (Darwen, Date, Fagin 2012)
- **KCNF** – Key-Complete Normal Form (Maier 1983)
- **RFNF** – Redundancy Free Normal Form (Vincent 1995)
- **SKNF** – Superkey Normal Form (Maier 1983)
- **5NF** – Femte normalform (Fagin 1979)

Joinavhengigheter

- La $\mathcal{R}$ være en relasjon med attributtene $R=\{A_1,\ldots,A_k\}$, og la $\{C_1,\ldots,C_m\}$ være en samling av **komponenter** over R (dvs. $C_i \subseteq R$ for alle i og $C_1 \cup \ldots \cup C_m = R$). Vanligvis overlapper hver C_i med minst én annen komponent C_j (der $j \neq i$).
- Vi sier at $\mathcal{R}$ tilfredsstiller **joinavhengigheten** eller **JD**en (**Join Dependency**) $J = \bowtie\{C_1,\ldots,C_m\}$ hvis vi for enhver gyldig instans av $\mathcal{R}$ har at dekomposisjonen av instansen i henhold til komponentene i J er tapsfri.
  – Enhver gyldig instans i $\mathcal{R}$ har altså den egenskapen at hvis vi projiserer den på $C_1,\ldots,C_m$, og deretter beregner $C_1 \bowtie \cdots \bowtie C_m$, så får vi tilbake instansen.

JDer og MVDer

- Joinavhengigheter gir opphav til en større klasse integritetsregler enn de som kan uttrykkes ved bare FDer og MVDer.

- Teorem. Hvis m=2, dvs. $J= \bowtie\{C_1,C_2\}$, så tilfredsstiller $\mathcal{R}$ JDen J hvis og bare hvis vi har en MVD $X \twoheadrightarrow Y$ der $X=C_1 \cap C_2$ og $Y=C_2-C_1$
- Dette betyr at en MVD er det samme som en JD med 2 komponenter

Armstrongs slutningsregler og JDer

- Armstrongs slutningsregler kan ikke utvides til å omfatte JDer: Det finnes intet sunt, komplett og endelig regelsett for JDer. (Det finnes sunne regelsett, men til ethvert slikt regelsett går det alltid an å konstruere eksempelrelasjoner der regelsettet ikke er tilstrekkelig til å vise alt som holder.)

Utgangspunkt for normalformene utover 4NF

- Alle integritetsregler er i form av FDer og JDer
- Dette omfatter også MVDer siden alle MVDer kan uttrykkes som JDer
- Generelt er altså utgangspunktet for alle normalformene en relasjon $\mathcal{R}$ der de gyldige instansene er karakterisert ved at de tilfredsstiller en mengde FDer $\mathcal{F}$ og en mengde JDer $\mathcal{J}$

Trivielle JDer

- En JD kalles **triviell** dersom en av komponentene omfatter samtlige attributter.

  Trivielle JDer holder for alle instanser uansett hvilke andre FDer og JDer som gjelder.

Irredusible JDer

- En JD J= $\bowtie\{C_1, \ldots, C_m\}$ er **reduserbar** hvis det finnes en komponent C_i der også
  J' = $\bowtie\{C_1, \ldots, C_{i-1}, C_{i+1}, \ldots, C_m\}$ holder
  (dvs. der J' følger logisk fra de integritetsreglene $\mathcal{F}$ og $\mathcal{J}$ som gjelder).
- Hvis det finnes en slik C_i, vil C_i alltid være overflødig fordi den ikke gir noen begrensninger på mengden av mulige instanser utover det de øvrige komponentene i $\mathcal{J}$ gir.
- En JD der det ikke fins noen slik komponent, kalles **irredusibel**.

Utgangspunkt for normalformene utover 4NF (forts.)

- Vi kan uten tap av generalitet begrense oss til relasjoner $\mathcal{R}$ der de gyldige instansene er karakterisert ved at de tilfredsstiller en mengde **ikketrivielle** FDer $\mathcal{F}$ og en mengde **ikketrivielle og irredusible** JDer $\mathcal{J}$.
- Så på resten av lysarkene har vi underforstått en relasjon $\mathcal{R}$ der mengden av gyldige instanser er karakterisert ved at de oppfyller reglene i to mengder $\mathcal{F}$ og $\mathcal{J}$ der FDene i $\mathcal{F}$ er ikketrivielle og JDene i $\mathcal{J}$ er ikketrivielle og irredusible.

Eksplisitte og implisitte integritetsregler

- Vi skiller mellom **eksplisitte** og **implisitte** FDer og JDer: De **eksplisitte** er de integritetsreglene som er listet opp i henholdsvis $\mathcal{F}$ og $\mathcal{J}$. De **implisitte** er de som følger logisk fra (kan utledes fra) $\mathcal{F}$ og $\mathcal{J}$

Partielt redundante tupler

- La r være en instans av $\mathcal{R}$. Et tuppel t i r er **partielt redundant** hvis det finnes et annet tuppel t' ≠ t i r og en ikketriviell (eksplisitt eller implisitt) FD $X \rightarrow A$ slik at t'[X] = t[X].
- Intuitivt betyr dette at informasjonen i t'[A] er representert to ganger (i t[A] og i t'[A]).

Fullt redundante tupler

- Et tuppel i en instans r er **fullt redundant** hvis det finnes en mengde av tupler S i r der $t \notin S$ og det er slik at t følger logisk fra S.
- Intuitivt betyr dette at informasjonen i tuppelet både finnes indirekte (kan utledes fra andre tupler i instansen) og direkte (i tuppelet selv). Tuppelet kan imidlertid ikke uten videre fjernes, for da risikerer vi at den resulterende mengden av tupler ikke lenger oppfyller alle integritets-reglene.

Essensielle tupler

- Et tuppel er **essensielt** hvis det hverken er partielt eller fullt redundant.

Verdiredundans

- La $\mathcal{R}$ være en relasjon, r en instans av $\mathcal{R}$, t et tuppel i r og A et attributt. Da er **verdien** t[A] **redundant** i r hvis det er slik at hvis vi erstatter t med et annet tuppel t' der t'[A] ≠ t[A] og t'[B] = t[B] for alle attributter B ≠ A, så får vi noe som ikke er en gyldig instans i $\mathcal{R}$.
- Intuitivt betyr dette at verdien i t[A] er entydig bestemt av de andre attributtverdiene i r.
- Merk at verdiredundans og partielt redundante tupler ikke er det samme.

BCNF – Boyce-Codd Normal Form

- Alternativ definisjon av BCNF:

  En relasjon $\mathcal{R}$ er på BCNF hvis alle eksplisitte og implisitte FDer følger logisk fra kandidatnøklene i $\mathcal{R}$.

- <u>Teorem.</u> En relasjon er på BCNF hvis og bare hvis ingen instans har noe partielt redundant tuppel.

4NF

- Alternativ definisjon av 4NF:

  En relasjon $\mathcal{R}$ er på 4NF hvis alle eksplisitte og implisitte MVDer følger logisk fra kandidatnøklene i $\mathcal{R}$.

ETNF – Essential Tuple Normal Form

- En relasjon $\mathcal{R}$ er på ETNF hvis alle tupler i alle instanser er essensielle.

- <u>Teorem.</u> En relasjon $\mathcal{R}$ er på ETNF hvis og bare hvis den er på BCNF og hver eksplisitt JD har minst én komponent som er en supernøkkel.

KCNF – Key-Complete Normal Form

- En relasjon $\mathcal{R}$ er på KCNF hvis den oppfyller de to kravene under:
  1. Relasjonen er på BCNF.
  2. For hver eksplisitt eller implisitt JD J er det slik at hvis vi tar unionen av de komponentene i J som er supernøkler, så får vi samtlige attributter i $\mathcal{R}$.

RFNF – Redundancy-Free Normal Form

- En relasjon $\mathcal{R}$ er på RFNF hvis det ikke finnes noen instans med et tuppel t der verdien av t er redundant i ett av attributtene.

- <u>Teorem.</u> RFNF = KCNF

SKNF – Superkey Normal Form

- En relasjon $\mathcal{R}$ er på SKNF hvis $\mathcal{R}$ er på BCNF og hver komponent i hver irredusibel eksplisitt eller implisitt JD er en supernøkkel.

5NF –
Femte normalform

- En relasjon $\mathcal{R}$ er på 5NF hvis alle eksplisitte og implisitte JDer følger logisk fra kandidatnøklene i $\mathcal{R}$.

Nøkkelsammenhengende JDer

- 5NF forklart på en annen måte:
  En relasjon $\mathcal{R}$ er på 5NF hvis alle eksplisitte og implisitte JDer J er **nøkkelsammenhengende** i henhold til følgende algoritme:
  - Så lenge det finnes to komponenter i J hvor snittet er en supernøkkel eller hvor den ene komponenten er inneholdt i den andre, så bytt ut de to komponentene med unionen av de to
  - Hvis prosessen terminerer med R (dvs. samtlige attributter) som en av komponentene, er J nøkkelsammenhengende, ellers ikke

BCNF, 4NF og 5NF

Sammenlikn definisjonene under:

- En relasjon $\mathcal{R}$ er på BCNF hvis alle eksplisitte og implisitte FDer følger logisk fra kandidatnøklene i $\mathcal{R}$.
- En relasjon $\mathcal{R}$ er på 4NF hvis alle eksplisitte og implisitte MVDer følger logisk fra kandidatnøklene i $\mathcal{R}$.
- En relasjon $\mathcal{R}$ er på 5NF hvis alle eksplisitte og implisitte JDer følger logisk fra kandidatnøklene i $\mathcal{R}$.

5NF, PJNF og SKNF

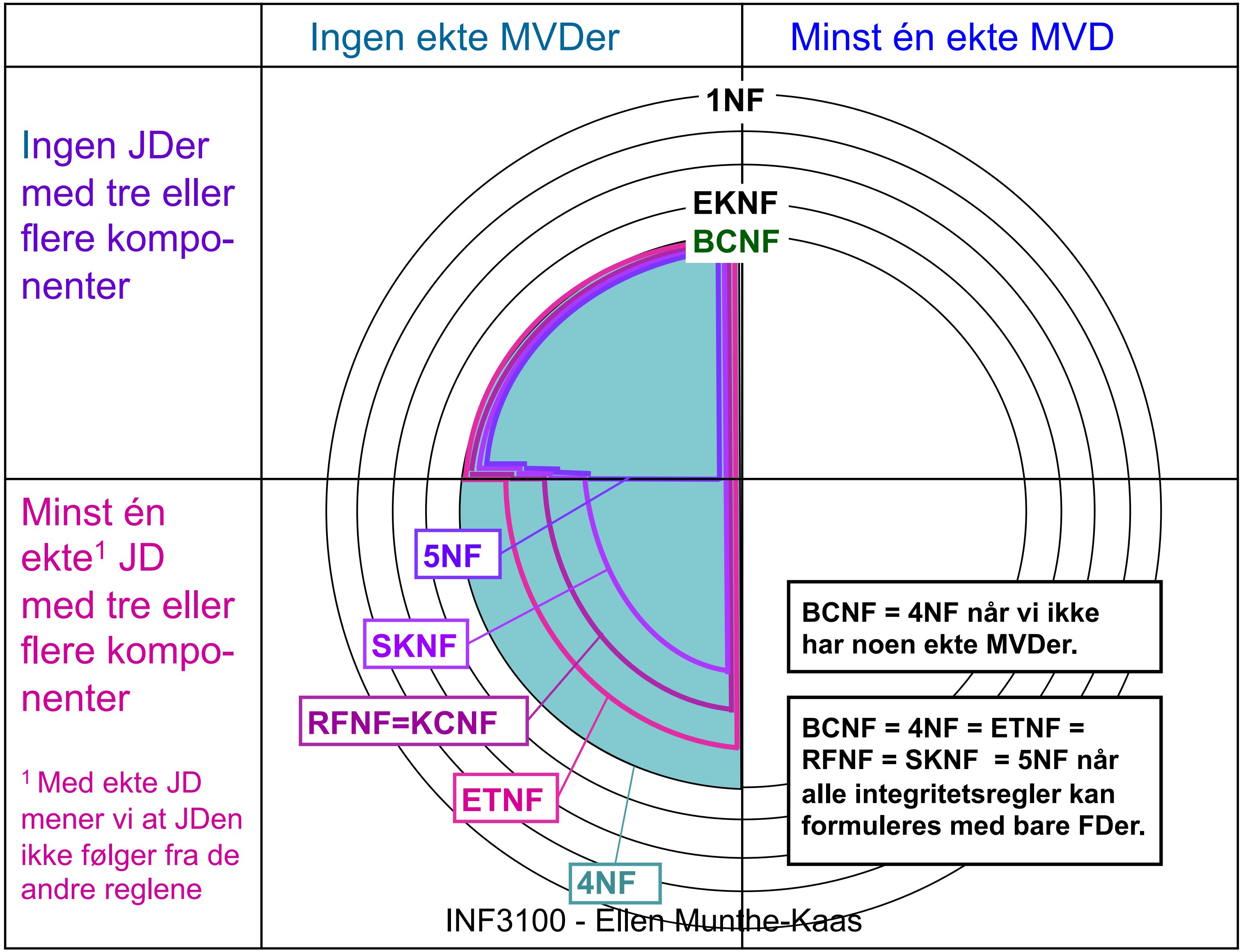
- 5NF kalles også PJNF – Project-Join Normal Form
- Vær oppmerksom på at mange lærebøker definerer 5NF feil – det de kaller 5NF, er i realiteten SKNF.

  Dette skyldes en tidlig misforståelse/feil i definisjonen av 5NF i David Maiers lærebok *The Theory of Relational Databases* (Computer Science Press, 1983) og at denne definisjonen ukritisk er blitt kopiert til andre lærebøker

Oppsummering av høye normalformer

- 5NF er (ekte) inneholdt i SKNF, SKNF er (ekte) inneholdt i RFNF, RFNF er (ekte) inneholdt i ETNF og ETNF er (ekte) inneholdt i 4NF
- RFNF = KCNF
- En relasjon $\mathcal{R}$ er på 4NF hvis og bare hvis alle (eksplisitte og implisitte) JDer med to komponenter består av to supernøkler.
  (Dette bevises ved å vise at en join av to supernøkler er tapsfri hvis og bare hvis deres snitt er en supernøkkel)
- Hvis vi fortsetter å tapsfritt dekomponere en relasjon på ETNF, så vil ikke dekomposisjonen bety noen reduksjon i antall tupler eller verdier som må lagres, tvert imot vil lagerbehovet *øke*

Høyere normalformer, oversikt



Chasealgoritmen utvidet for JDer

Gitt en dekomposisjon av $\mathcal{R}$(A,B,...) til relasjonene $\mathcal{S}_1$, ... , $\mathcal{S}_k$ og et sett $\mathcal{F}$ med FDer og et sett $\mathcal{J}$ med JDer for $\mathcal{R}$. Er dekomposisjonen tapsfri?

1. Lag en tabell med en kolonne for hvert attributt i $\mathcal{R}$ og en rad for hver $\mathcal{S}_i$
2. I kolonnen for attributt A, for hver rad i:
   - Skriv a hvis A er et attributt i $\mathcal{S}_i$
   - Skriv a_i hvis A ikke er et attributt i $\mathcal{S}_i$
3. Så lenge det skjer forandringer i tabellen og det ikke finnes en rad uten subskriptverdier:
   - For hver FD X $\rightarrow$ Y, for alle rader i tabellen med lik X-verdi, gjør Y-verdiene like. Hvis en av Y-ene er en verdi uten subskript, skal denne velges
   - For hver JD $\bowtie\{C_1,\ldots,C_m\}$, for hvert utvalg av m rader $t_1,\ldots,t_m$ der $t_i[C_i \cap C_j]=t_j[C_i \cap C_j]$ for alle par (i,j), lag en ny rad t der $t[C_i]=t_i[C_i]$ for alle i.

- Hvis en rad er uten subskriptverdier, er dekomposisjonen tapsfri, ellers ikke.

Annen bruk av chasealgoritmen

- For å vise en JD $\bowtie\{C_1,...,C_m\}$, start tabellen med m rader der rad j har verdier uten subskript for attributtene i C_j og verdier med subskript forøvrig. Målet er å komme frem til en instans der en av disse radene er uten subskriptverdier.

Eksempel 1: JDer og begrepene ikketriviell og irredusibel

- Betrakt $\mathcal{R}$(A,B,C) der A$\rightarrow$BC er eneste integritetsregel. Da holder de implisitte JDene
  $J_1 = \bowtie\{AB,AC,ABC\}$
  $J_2 = \bowtie\{AB,AC,BC\}$
  $J_3 = \bowtie\{AB,AC\}$
  Dette kan vises ved å bruke chasealgoritmen

Eksempel 1 (forts.)

- J_1 er triviell; den inneholder komponenten ABC. J_1 er reduserbar; den kan reduseres til $\bowtie$\{ABC\}.
- Hverken J_2 eller J_3 er trivielle.
- J_2 er reduserbar: komponenten BC er overflødig; når vi tar den bort, reduseres J_2 til J_3.
- J_3 er irredusibel; vi kan ikke fjerne noen av komponentene AB eller AC uten at vi bryter kravet om at alle attributtene skal være representert i minst én komponent.

Eksempel 2: Relasjon som er på ETNF, men ikke RFNF

- Gitt relasjonen $\mathcal{R}$(A,B,C) der A er leverandør, B er reservedel og C er prosjekt. Følgende integritetsregler skal holde:
  - Hvis leverandør A leverer en reservedel B, og B leveres til prosjekt C, og C får leveranser fra A, så skal A levere B til C.
    Dvs. JDen $\bowtie$\{AB,BC,CA\} skal gjelde.
  - Hvert prosjekt C mottar en reservedel B fra maksimalt én leverandør A.
    Dvs. FDen BC$\rightarrow$A skal gjelde.

Eksempel 2 (forts.)

- Hva sier JDen $\bowtie\{AB,BC,CA\}$?

A	B	C
a	b	c’
a’	b	c
a	b’	c
a	b	c

Hvis disse tre tuplene er med,

må dette tuppelet også være med (tuppelet er i såfall fullt redundant)

- Hva sier FDen $BC \rightarrow A$?

a og a’ i instansen over skal være like – så instansen må se slik ut:

A	B	C
a	b	c’
a	b	c
a	b’	c

(ingen redundante tupler; det redundante tuppelet ”forsvinner”)

Eksempel 2 (forts.)

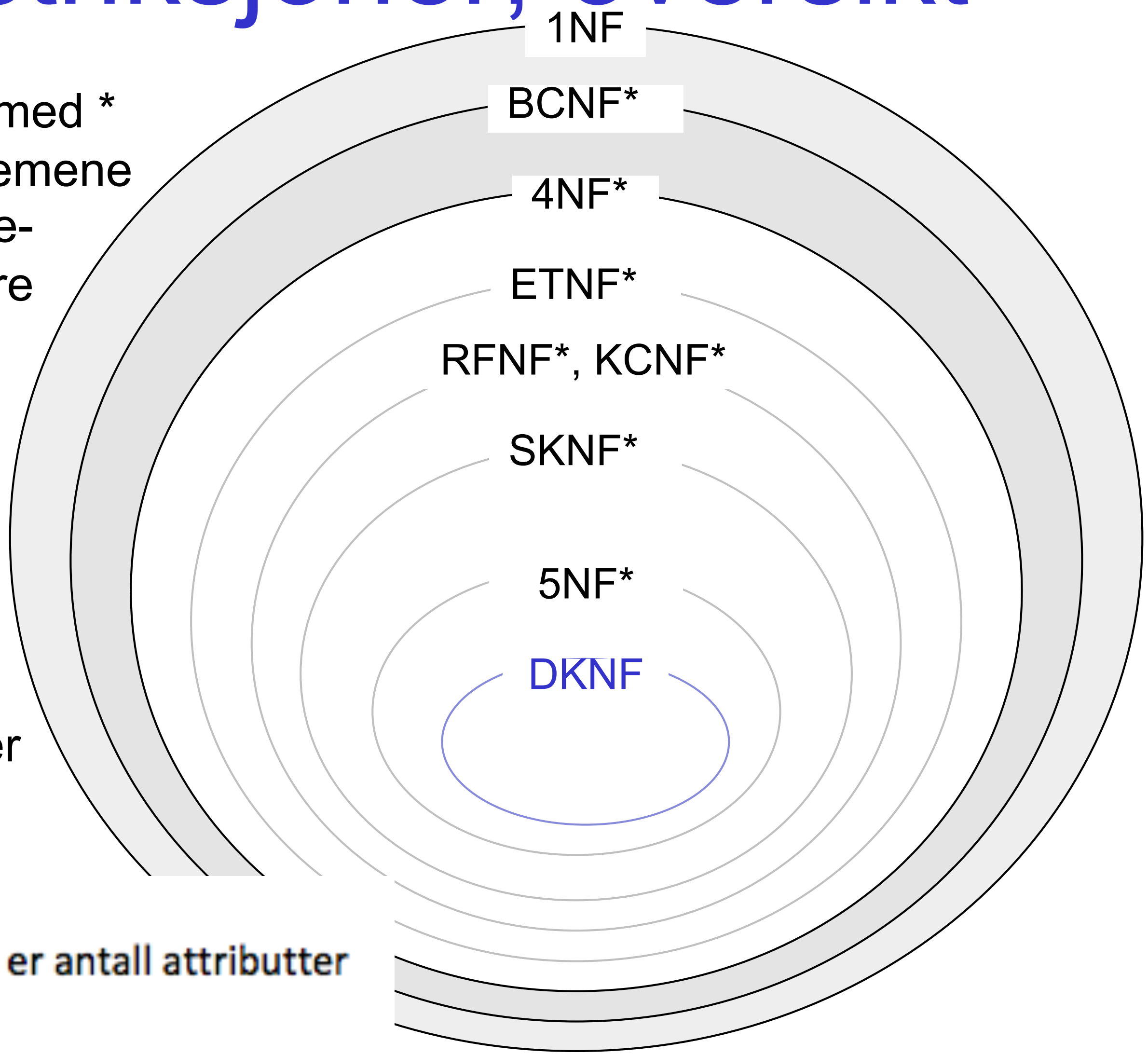
- Relasjonen er på BCNF: Bruk chasealgoritmen til å sjekke at ingen av FDene $AB \rightarrow C$, $B \rightarrow A$, $C \rightarrow A$ eller $C \rightarrow B$ holder.
- Relasjonen er på ETNF: Eneste kandidatnøkkel er BC. Så JDen $\bowtie\{AB,BC,CA\}$ inneholder en komponent som er en supernøkkel (nemlig BC).
- Relasjonen er ikke på RFNF: Siden RFNF=KCNF, er det tilstrekkelig å undersøke om unionen av supernøkkel-komponentene i $\bowtie\{AB,BC,CA\}$ omfatter samtlige attributter. Men BC er eneste supernøkkelkomponent, så A er ikke dekket av noen supernøkkelkomponent.

Domenerestriksjoner

- Det er viktig å være klar over at en av forutsetningene for normalformene slik de er formulert over, er at integritets-reglene **bare** er i form av FDer og JDer
- Hvis vi i tillegg til FDer og JDer kan ha restriksjoner på hvilke verdier et attributt kan inneholde, f.eks. at ett attributt har domenet boolean og et annet har en av verdiene {gull, sølv, bronse}, så holder ikke nødvendigvis teoremene lenger. I tillegg må (noen av) normalform-definisjonene endres litt.

Høyere normalformer med domenerestriksjoner, oversikt

For alle normalformer merket med * blir definisjonene og/eller teoremene påvirket hvis vi kan ha domene-restriksjoner. For 4NF og lavere er det tilstrekkelig å forutsette at alle domener har minst to elementer; da holder teorem-ene. I praksis er dette alltid oppfylt. For *-normalformene høyere enn 4NF holder alle teoremene når vi i tillegg har at nøkkelattributtenes domener er store nok[1]



[1] større enn $\binom{n}{\lfloor \frac{n}{2} \rfloor} \approx \sqrt{\frac{2}{\pi n}} \times 2^n$ der n er antall attributter

DKNF –
Domain-Key Normal Form

- En relasjon $\mathcal{R}$ er på DKNF hvis de eneste integritetsreglene i $\mathcal{R}$ er FDer som oppfyller BCNF (også kalt “nøkkelrestriksjoner”) og domenerestriksjoner (Fagin 1981)
- DKNF er lett å håndheve, men det er for mange integritetsregler som ikke kan uttrykkes på denne måten til at vi kan begrense oss til å ha komponenter på DKNF

Eksempel 3: ETNF-teoremet bryter sammen ved domenerestriksjoner

- Gitt relasjonen $\mathcal{R}$ (A,B,C,D,E,F). Følgende integritetsregler gjelder:
  - dom(A) = dom(B) = {0,1}
    dom(C, D, E, F) = {0,1,2,...}
  - ABC→DEF, DEF→ABD (så ABC og DEF er kandidatnøkler)
  - J = ⋈{ACDE, ACDF, ACEF,
    BCDE, BCDF, BCEF}

Eksempel 3 (forts.)

- Det kan vises at J følger fra FDene og domenerestriksjonene alene, så vi får nøyaktig samme mengde gyldige instanser (nøyaktig samme relasjon $\mathcal{R}$) om vi fjerner J.
- Relasjonen $\mathcal{R}$ oppfyller derfor BCNF og DKNF.
- På lysark 18 stod følgende teorem: En relasjon er på ETNF hvis og bare hvis den er på BCNF og hver eksplisitt JD har minst én komponent som er en supernøkkel.
- J inneholder ingen supernøkkelkomponent, så hvis vi bruker teoremet, får vi at $\mathcal{R}$ *ikke* er i ETNF.
- Samtidig er det slik at om vi fjerner J, og deretter bruker teoremet, får vi at $\mathcal{R}$ *er* i ETNF. Så vi har en inkonsistens (selvmotsigelse): $\mathcal{R}$ kan vises å både tilhøre og ikke tilhøre ETNF.
- Problemet skyldes at teoremet ikke er gyldig i nærvær av ”for små” domener. Derfor må alle definisjoner og beviser revurderes i nærvær av domenerestriksjoner.