

## IN2090 – Oblig 3

### Oppgave 1

- a) Det står at født bestemmer alder og der med er alder avhengig av når man er født så derfor kom jeg fram til dette:  
**født  $\rightarrow$  alder**
- b) Står at født og personnr bestemmer fødselsnummer så tenker at det vil være sånn her, men vet ikke hvorfor det er skrevet om et attributt som ikke står i relasjonen. Men min teori er at man da kan lage et unikt attributt for tabellen:  
**født, personnr  $\rightarrow$  fødselsnummer**
- c) Sånn jeg tolker det som står er at stillingen bestemmer lønnen sammen med masse verdier som gir et tillegg til lønnen:  
**stilling, fareT, hemmeligT, skummeltT  $\rightarrow$  lønn**
- d) Denne FDen sier at hvis du vet stillingen og om det er farlig vet vi da at den boolske verdien til skummelt også:  
**stilling, fareT  $\rightarrow$  skummeltT**

### Oppgave 2

- a) Tillukningen til A er alltid seg selv og hva som avhenger av A, så tillukningen til A er:  
 **$A^+ = A, L$**
- b) Tillukningen til NKA er:  **$\{N, K, A\}^+ = N, K, A, L, F, G, P, B$**
- c) Siden A og N ikke forekommer i noen høyresider, vil A og N være med i alle kandidatnøkler. Ser at P og B kun forekommer på en høyreside, så P og B vil aldri være i en kandidatnøkkel. Alle kandidatnøklerne er:  **$\{A, N, F, G\}, \{A, N, K\}$**

### Oppgave 3

- a) Normalformer er en måte å designe relasjonelle databaser på slik at man minimerer duplisering av data. Det å duplisere data kan skape anomalier, som er at data som skulle ha vært det samme to steder er forskjellig. Vi har 4 normaliseringsgrader og for høyere normaliseringsgrad må de lavere være godkjent. Desto høyere normaliseringsgrad vil gi oss flere tabeller og flere relasjoner, som igjen vil gjøre spørringene våre tregere og ineffektive, pga flere joins.

b) Må starte med å skrive om redundante FDer. De nye FDene er:

agentId → navn

agentId → født

navn, født → agentId

navn → initialer

oppdragsNavn → varighet

oppdragsNavn → lokasjon

Også har vi kandidatnøkklene:

{agentId, oppdragsNavn}, {navn, født, oppdragsNavn}

Det vi må gjøre nå er å gå gjennom diverse spørsmål for hver FD, men hvis en FD feiler trenger man ikke sjekke for resten da kan man bare gå videre til neste spørsmål.

Spørsmål 1 er, Er X en supernøkkel? Hvis ja, BCNF så langt, gå til neste FD. Hvis nei, brudd på BCN. Gå til neste spørsmål. Her er ikke agentId en supernøkkel så da er det brudd på BCNF (høyeste normaliseringsgraden) og vi går til neste spørsmål.

Neste er, Er A et nøkkelattributt? Hvis ja, 3NF så langt, gå til neste FD. Hvis nei, brudd på 3NF, gå til neste spørsmål. Det stemmer hele veien ned til FDen, navn → initialer, fordi initialer er i ingen av kandidatnøkklene. Dermed er det brudd på 3NF og vi går til neste spørsmål.

Neste er, Er X del av en kandidatnøkkel? Hvis nei, 2NF så langt, gå til neste FD. Hvis ja, brudd på 2NF og skjemaet er på 1NF. Stopp. Starter på toppen igjen og ser at i første FD så er agentId en del av en kandidatnøkkel og bryter dermed med 2NF. Når har vi nådd bunnen og kan konkludere med at skjemaet er på 1NF.

## Oppgave 4

Det første vi må gjøre når vi skal dekomponere noe tapsfritt til BCNF er å finne alle kandidatnøkklene. De har vi som **{A,D}** og **{B,C,D}**. Neste er å sørge for at alle FDer kun har et attributt navn på høyresiden. Så de nye FDene er:

$A \rightarrow B$

$A \rightarrow C$

$BC \rightarrow A$

$D \rightarrow E$

$AD \rightarrow F$

$E \rightarrow F$

Deretter må vi sjekke om R bryter med BCNF, som den gjør. Hvis R ikke hadde brutt med BCNF så skulle man bare ha returnert R. Men siden den bryter med BCNF må vi gjennom en del steg før R er på BCNF.

Nå må vi finne en av FDene som bryter med BCNF også finne tillukningen til den. Jeg velger den første FDen,  $A \rightarrow B$ , den bryter med BCNF, fordi A er ingen supernøkkel.

Tillukningen til A er  $A^+ = \{A,B,C\}$ .

Så er det og dekomponerer R til  $S_1(Y^+)$  og  $S_2(Y, X/Y^+)$ . Y i dette tilfellet er A og X er resten hele mengden med attributter. Da får vi  $S_1(A,B,C)$  og  $S_2(A,D,E,F)$ .

Nå skal vi gjøre dette rekursivt over først  $S_1$  og med  $S_2$ , men kun med de FDene som inneholder attributter som er i  $S_1$  og  $S_2$  separat.

La oss gjøre  $S_1$  først. Sjekk om den bryter med BCNF.  $S_1$  bryter ikke med BCNF og er på formen BCNF nå. Da kan vi gå videre til  $S_2$ .

Sjekk om  $S_2$  bryter med BCNF. Vi har at FDen,  $D \rightarrow E$  bryter med BCNF. Dekomponerer og ender opp med  $S_{21}(D,E,F)$  og  $S_{22}(D,A)$ . Gjør det samme igjen på  $S_{21}$  og  $S_{22}$ . Nå ser vi at  $S_{22}$  ikke bryter med BCNF så den er vi ferdig med, men  $S_{21}$  bryter med BCNF med FDen,  $D \rightarrow E$ . Beregner tillukningen til  $D \rightarrow D^+ = (D,E)$ . Da får vi  $S_{211}(D,E)$  og  $S_{212}(D,F)$ . Når vi så sjekker igjen ser vi at ingen av de bryter med BCNF og vi er da ferdige med dekomponeringen.

Så da står vi igjen med relasjonene  $S_1(A,B,C)$ ,  $S_{22}(A,D)$ ,  $S_{211}(D,E)$  og  $S_{212}(D,F)$  etter dekomponeringen.