Databaser 1: Relasjonsmodellen

19.1 Oppsummering

Logikk i et nøtteskall:

- Signatur: Symboler, syntaks
- Tolkning: Sannhet, gyldighet, konsekvens
- Kalkyle: Deduksjon, bevis

Overført til databaser:

- Signatur = databaseskjema
  - åpen formel = spørring
  - $\forall x$(spørring $\to$ spørring) = "constraint"
- Tolkning:
  - Instans (dvs data) = modell
  - Sannhet av åpen formel = svar på spørring
- Kalkyle: Ingen i det generelle tilfellet!

19.2 Relasjonsmodellen

Skjema

Vi antar følgende som gitt:

- En tellbar uendelig mengde konstanter **dom**
- En tellbar uendelig mengde relasjons-symboler
- En tellbar uendelig mengde attributt-symboler
- En funksjon **sort** som mapper hvert relasjonssymbol til en mengde attributter

Hvis **sort**$(R) = U = \{A_1, \ldots, A_n\}$ kan vi skrive

- $R[U]$
- $R[A_1, \ldots, A_n]$

Definisjon 19.1
Et *skjema* er en endelig mengde $R_1[U_1], \ldots, R_n[U_n]$

Instanser

La R ha aritet n:

- n-tuppel: $\langle t_1, \ldots, t_n \rangle$, hver $t_i \in$ **dom**
- Relasjon over R: mengde n-tupler

Definisjon 19.2
En *instans* $\mathbf{I}(R)$ over R er relasjon over R. En instans over et skjema er gitt ved en instans over hvert av skjemaets relasjonssymboler

Konjunktive spørringer

Gitt en tellbar mengde av variable Var:

- Åpent tuppel: tuppel som kan inneholde variable i tillegg til konstanter
- Atom: $R(u)$, u åpent tuppel
- $\mathsf{FV}(R(u))$: mengden av variable i u

Formler over et relasjonsskjema:

- Et atom er en formel
- $\varphi \wedge \psi$ formel hvis φ, ψ er formler
  - $\mathsf{FV}(\varphi \wedge \psi) = \mathsf{FV}(\varphi) \cup \mathsf{FV}(\psi)$
- $\exists x \varphi$ formel hvis $varphi$ er formel
  - $\mathsf{FV}(\exists x \varphi) = \mathsf{FV}(\varphi) \setminus \{x\}$

Definisjon 19.3
En *spørring* i den konjunktive kalkylen er et uttrykk på formen

$$\{u \mid \varphi\}$$

der

- $u = \langle e_1, \ldots, e_n \rangle$ er fritt tuppel
- $\mathsf{FV}(u) = \mathsf{FV}(\varphi)$

Valuasjoner

Definisjon 19.4
En *valuasjon* over $V \subseteq$ Var er funksjon

$$\nu : V \to \mathrm{Var}$$

I oppfyller φ under ν, $\mathbf{I} \models \varphi[\nu]$, hvis

- $\varphi = R(u)$ og $\nu(u) \in \mathbf{I}(R)$
- $\varphi = \psi \wedge \xi$ og $\mathbf{I} \models \psi[\nu|_{\mathsf{FV}(\psi)}]$ og $\mathbf{I} \models \xi[\nu|_{\mathsf{FV}(\xi)}]$
- $\varphi = \exists x \psi$ og for en $c \in$ **dom**, $\mathbf{I} \models \psi[\nu \cup \{x/c\}]$

der ν er valuasjon over $\mathsf{FV}(\varphi)$.

Bilde og ekvivalens

La $q = \{u \mid \varphi\}$ og $q' = \{u' \mid \varphi'\}$.

Definisjon 19.5
$q(\mathbf{I}) = \{\nu(u) \mid \mathbf{I} \models \varphi[\nu] \text{ og } \nu \text{ valuasjon over } FV(\varphi)\}$

Definisjon 19.6
q er *inneholdt i* q', $q \subseteq q'$, hvis

- $FV(q) = FV(q')$
- $q(\mathbf{I}) \subseteq q'(\mathbf{I})$ for hver $\mathbf{I}$.

q og q' er *ekvivalente* hvis $q \subseteq q'$ og $q' \subseteq q$

Merk: Kvantifiserer over endelige modeller

Enkle resultater

Aktivt domene:

- $adom(\varphi)$: Mengde konstanter i φ
- $adom(\varphi, \mathbf{I}) = adom(\varphi) \cup \mathbf{I}$

Teorem 19.7
La q være konjunktiv spørring:

- $adom(q(\mathbf{I})) \subseteq adom(q, \mathbf{I})$
- $q(\mathbf{I}) = q(adom(q(\mathbf{I})))$
- q er monoton: Hvis $\mathbf{I} \subseteq \mathbf{J}$, så $q(\mathbf{I}) \subseteq q(\mathbf{J})$
- q er oppfyllbar: Det finnes en $\mathbf{I}$ slik at $q(\mathbf{I}) \neq \emptyset$ $\dashv$

Teorem 19.8
En konjunktiv spørring er ekvivalent med en spørring på normalform:

$$\exists x_1, \ldots, x_m (R(u_1), \ldots, R(u_n))$$

Regelnotsjon for spørring på normalform:

$$ans(u) \leftarrow R(u_1), \ldots, R(u_n)$$

der alle variable i u forekommer i minst en u_i.

Sammensatte spørringer

Et *view* er en presentasjon av et utvalg fra databasen

- Et view kan betraktes som en spørring
- View kan materialiseres, men trenger ikke

Spørringer mot view genererer sammensatte spørringer:

$$S_1(u_1) \leftarrow B_1$$

$$\vdots$$

$$S_n(u_n) \leftarrow B_n$$

der S_i kan forekomme i B_j hvis $i < j$.

Teorem 19.9

For enhver sammensatt spørring finnes det en ekvivalent konjunktiv spørring (dvs som gir u_n som svar). $\dashv$

Identitet gir mer uttrykkskraft:

- Merk: Unique Name Assumption! $a = b$ er alltid usann
- Vi kan få inkonsistente spørringer:
  - $\text{ans}(x) \leftarrow R(x), x = a, x = b$
  - Ved transitiv tillukking kan vi lett teste konsistens
  - Hvis konsisten kan likhet elimineres
- Vi kan få resultater som ikke er instanser:
  - $\text{ans}(x, y) \leftarrow R(x), y = z$
  - Krav: En term $y = t$ krever at y er lik en term i en $R_i(u_i)$

Tablåer

Tablåer er instanser som kan ha forekomster av variable:

- fritt n-tuppel: $\langle t_1, \ldots, t_n \rangle$, hver $t_i \in \mathbf{dom} \cup \mathrm{Var}$
- *Tablå* $\mathbf{T}(R)$ over R: mengde frie n-tupler over R
- Tablå over et skjema = tablå over hver av skjemaets relasjonssymboler

Definisjon 19.10

En *tablåspørring* er et par $(\mathbf{T}, u)$ der hver variabel i u forekommer i $\mathbf{T}$.

- Tablåer er en hendig syntaktisk variant av regel-notasjonen
- I tablåer betraktes spørringer som instanser
- En spørring anvendt på et tablå gir en homomorfi!

Homomorfiteoremet

Gitt $q = (\mathbf{T}, u)$ og $q' = (\mathbf{T}', u')$ over samme skjema.

Definisjon 19.11
En substitutsjon Θ er en *homomorfi* fra q' til q hvis $\theta(\mathbf{T}') \subseteq \mathbf{T}$ og $\theta(u') = u$.

Teorem 19.12
$q \subseteq q'$ hviss det finnes en homomorfi fra q' til q. ⊣

Teorem 19.13
$q \subseteq q'$ hviss $u \in q'(\mathbf{T})$. ⊣