

INF1300

Introduksjon til databaser

Dagens tema:

- Begrepsdannelse
- Eksterne entydighetsskranker
- Verdiskranker
- Mengdeskranker

Et eksempel fra virkeligheten

- La oss se på setningen
 «**På Blindern klokken 8 målte Jens 9 grader**»
- De tre spørsmålene vi alltid må stille oss, er:
 - Hvilke **begreper** har vi?
 - Hvilken **representasjon** bruker vi for disse begrepene?
 - Er setningen **elementær** i vårt UoD, og hvis den ikke er det, hvordan kan den splittes opp?

Et eksempel fra virkeligheten

- «På Blindern klokken 8 målte Jens 9 grader»
 - **Blindern** er et **sted** med representasjon **stedsnavn**
 - **8** er et **tidspunkt** med representasjon **dato og klokkeslett**
 - **Jens** er en **person** med representasjon **personnavn**
 - **9** er en **temperatur** med representasjon **°C**
- For å avgjøre om setningen er **elementær**, tyr vi til et **forekomstdiagram**

Et eksempel fra virkeligheten

Tid	er tid for måling	er sted for måling	er resultatet av målingen	foretok målingen	Person
	3.8.08 kl.8:00	Blindern	9		Jens
	3.8.08 kl.8:00	Værnes	9		Siri
	3.8.08 kl.8:00	Blindern	11		Lars
	4.8.08 kl.8:00	Blindern	9		Jens

Diagram annotations: 'Sted' points to the 'er sted for måling' column. 'Temperatur' points to the 'er resultatet av målingen' column.

- Er det noen av forekomstene som må strykes?

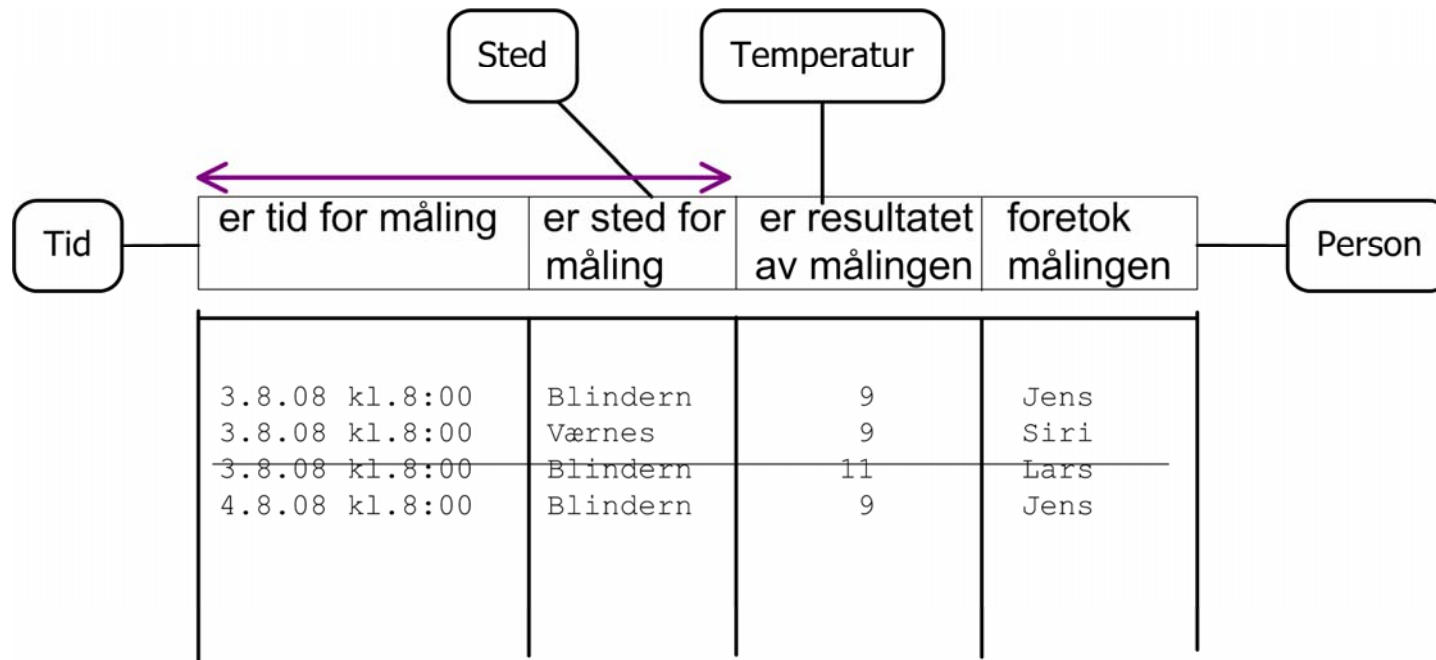
Et eksempel fra virkeligheten

The diagram shows a table with four columns. Above the table, there are four labels in rounded rectangles: 'Tid' (pointing to the first column), 'Sted' (pointing to the second column), 'Temperatur' (pointing to the third column), and 'Person' (pointing to the fourth column). The table has four rows of data. The third row is crossed out with a horizontal line.

er tid for måling	er sted for måling	er resultatet av målingen	foretok målingen
3.8.08 kl.8:00	Blindern	9	Jens
3.8.08 kl.8:00	Værnes	9	Siri
3.8.08 kl.8:00	Blindern	11	Lars
4.8.08 kl.8:00	Blindern	9	Jens

- Hvor skal entydighetsskranken(e) stå?

Et eksempel fra virkeligheten

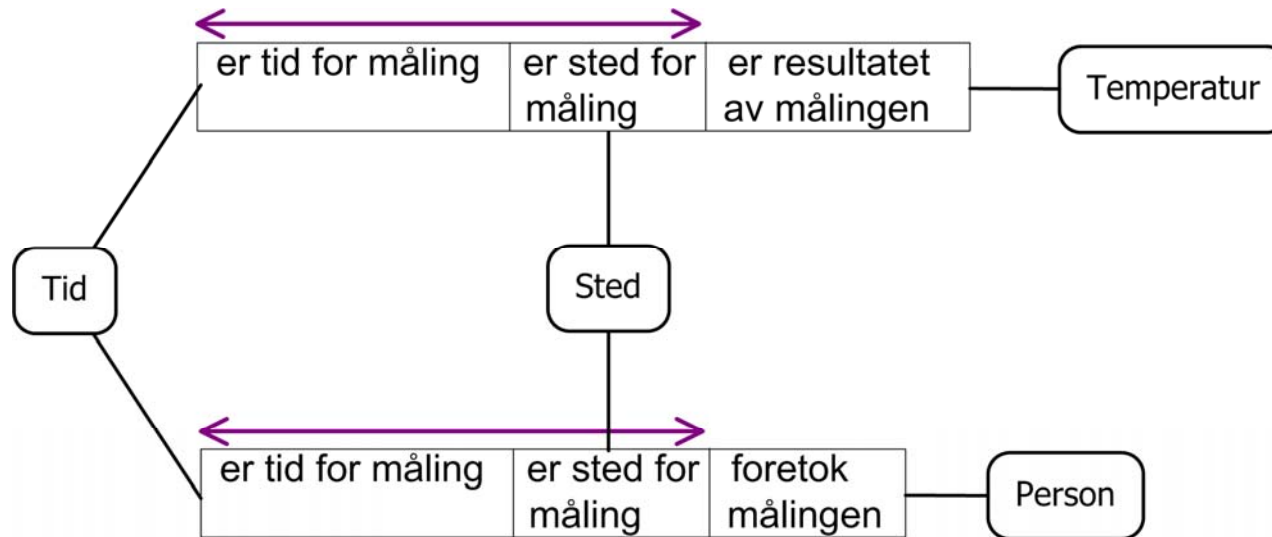


- Dette er en entydighetsskranke over de to begrepene «Tid» og «Sted»
- Det gjør at setningen ikke er elementær (se neste lysark)

n-1-regelen

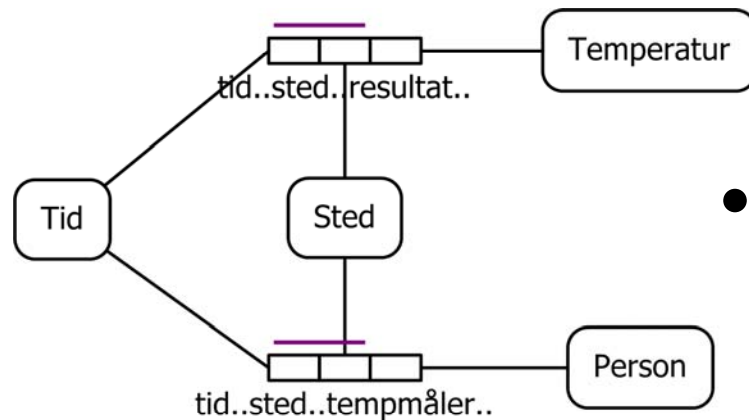
- *En setning med aritet n er aldri elementær hvis det tilhørende forekomstdiagrammet har en entydighetsskranke som er **kortere** enn $n-1$*
- *Hvis korteste entydighetsskranke har lengde $n-1$, er setningen elementær*
- Hvis korteste entydighetsskranke har lengde n , er setningen *nesten* alltid elementær
- Unntakene forekommer svært sjelden i praksis, og de er ikke pensum i INF1300 (men i INF3100)
- *Setninger med aritet 1 eller 2 er alltid elementære*

Et eksempel fra virkeligheten



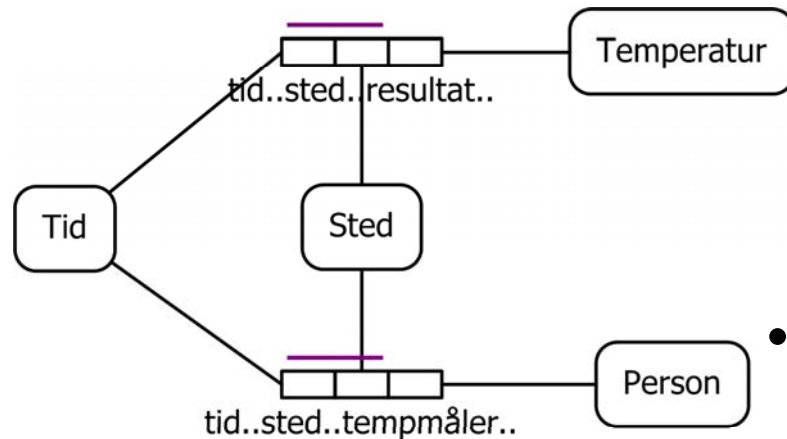
- Vi har nå et skjema for elementære ternære setninger med entydighetsskranker av lengde to
- Entydighetsskrankene sier at Tid og Sted sammen bestemmer Temperatur og sammen bestemmer Person.

Et eksempel fra virkeligheten

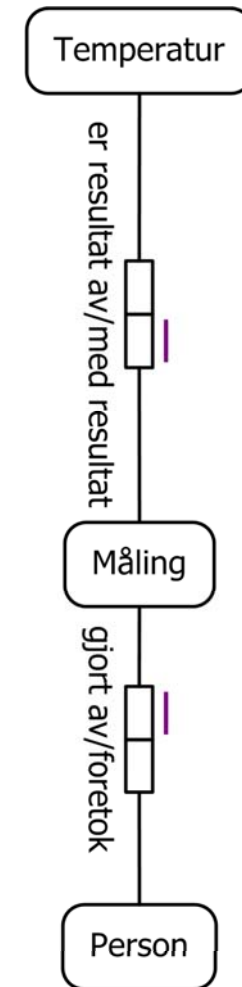


- De to setningene har til felles at Tid og Sted *sammen* bestemmer en tredje verdi
- ORM-diagrammet til venstre er korrekt, men tydeliggjør ikke at for både Temperatur og Person er forekomster bestemt av **en og samme** kombinasjon av Tid og Sted

Et eksempel fra virkeligheten

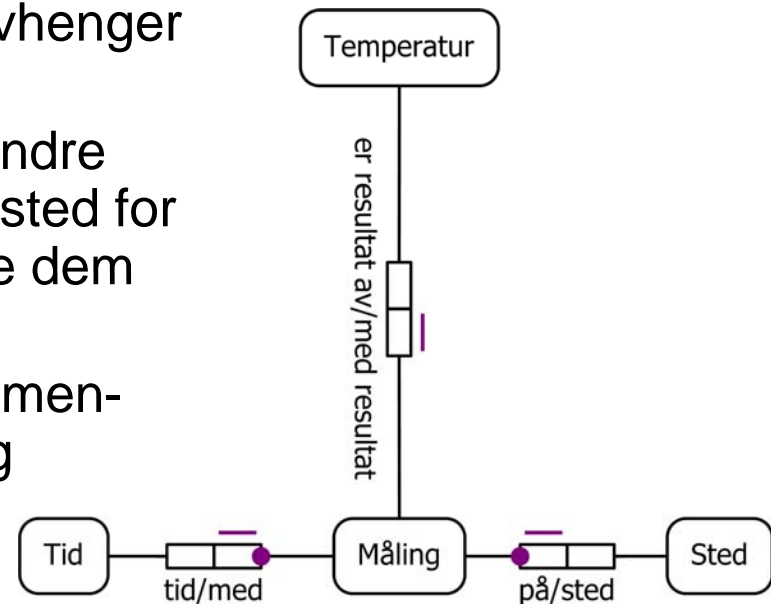
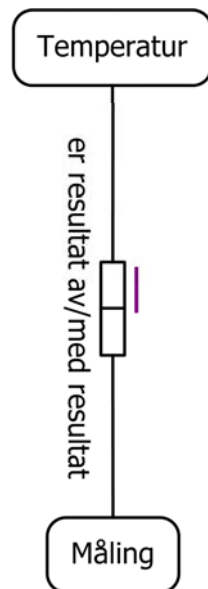


- Vi kan tydeliggjøre dette ved å lage **et nytt begrep** av kombinasjonen av begrepene Tid og Sted. Det nye begrepet kan vi i vårt UoD kalle **Måling**.
- Deretter kan vi knytte forekomster av Temperatur og Person til det nye begrepet.
- Entydighetsskrankene på rollene *med_resultat* og *gjort_av* uttrykker at én forekomst av Måling maksimalt har én forekomst av Temperatur og må være gjort av maksimalt én Person. Dette svarer til entydighetsskrankene over tid+sted i modellen til venstre.

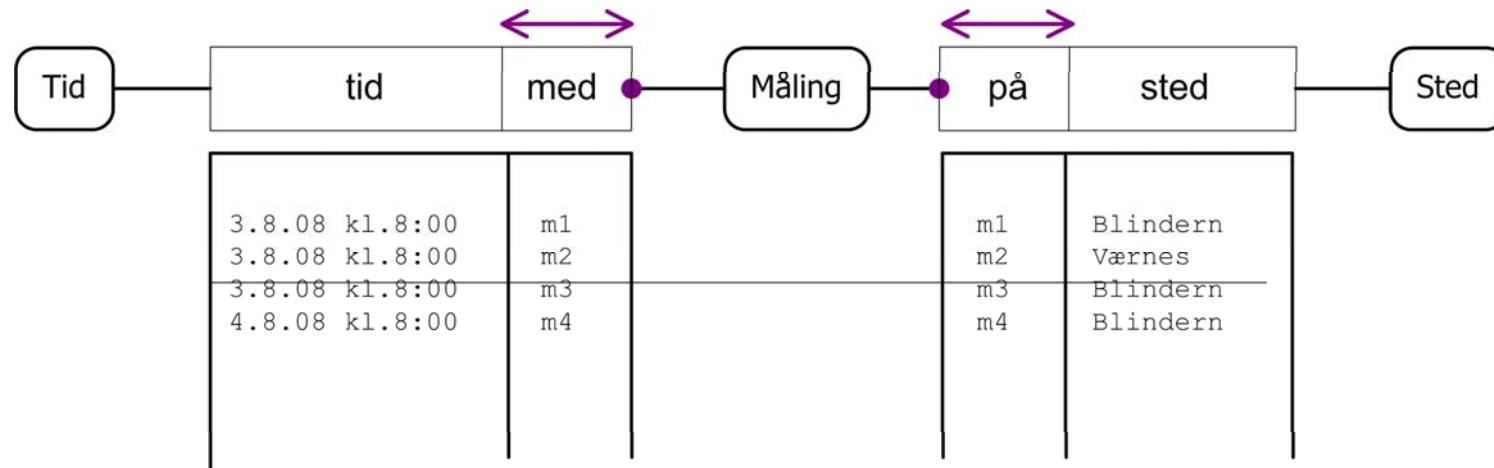


Begrepsdannelse

- Et problem er at begrepene Tid og Sted ble borte fra modellen
- Hvorvidt vi kan tillate det, avhenger av UoD
 - Hvis Tid eller Sted har andre roller enn å være tid og sted for Måling, kan vi ikke fjerne dem fra modellen
- La oss se nærmere på sammenhengen mellom Tid, Sted og Måling:
- Entydighetsskrankene og de totale rollene uttrykker at én forekomst av Måling må være utført på eksakt ett Sted og én Tid

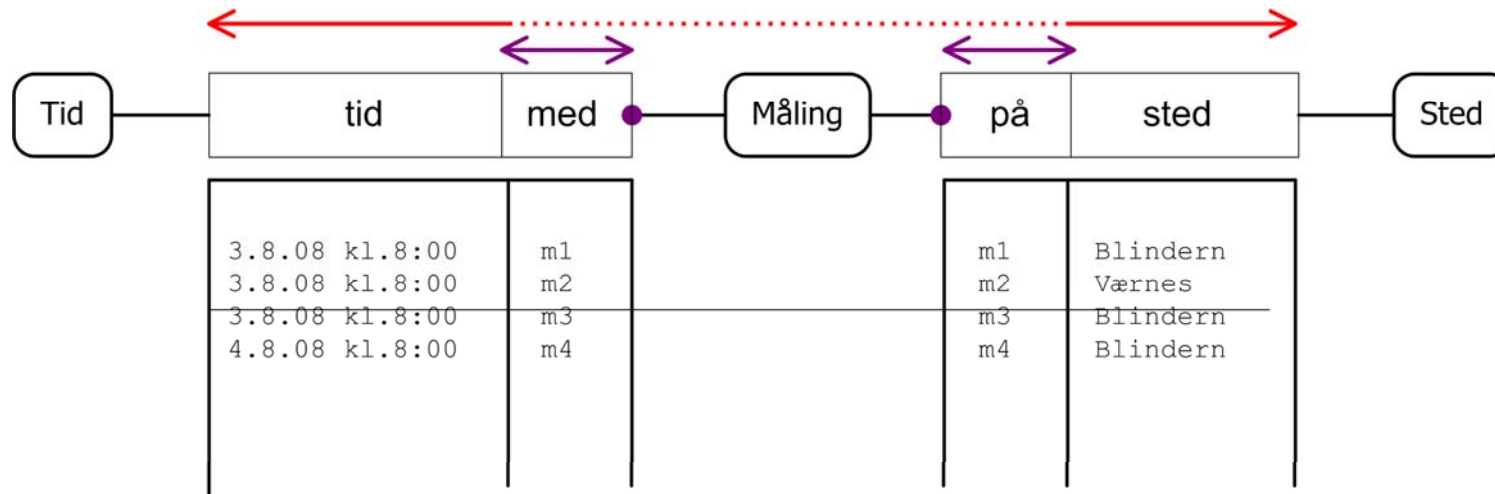


Begrepsdannelse



- Modellen er fortsatt ikke god nok:
 - Den utelukker ikke at to forskjellige målinger (av samme fenomen) kan være foretatt på nøyaktig samme tid og sted

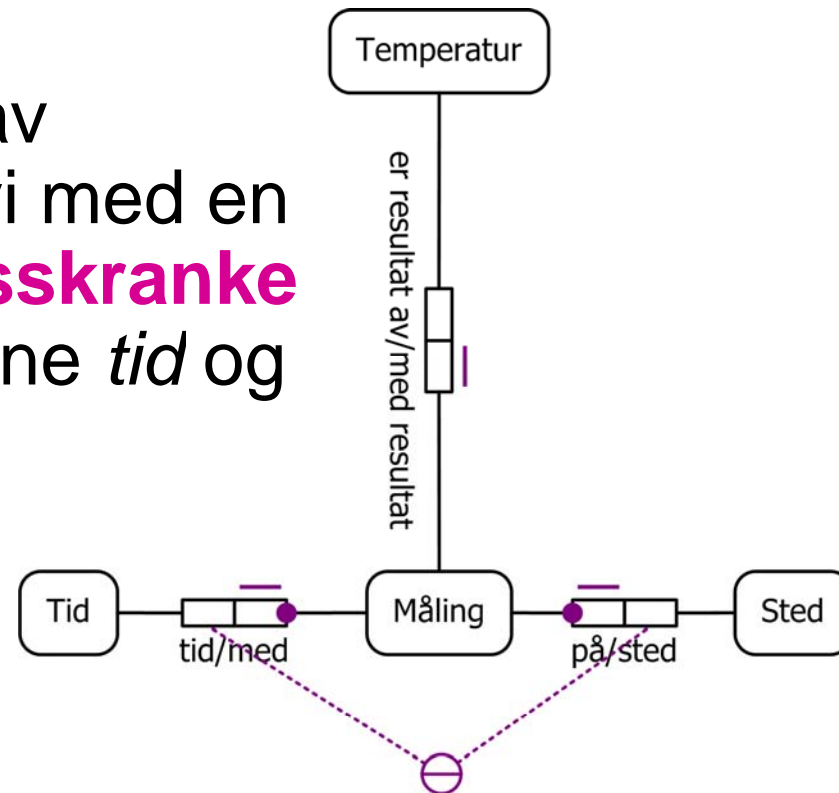
Begrepsdannelse



- Modellen er fortsatt ikke god nok:
 - Den utelukker ikke at to forskjellige målinger (av samme fenomen) kan være foretatt på nøyaktig samme tid og sted
 - Vi trenger en entydighetskranke *på tvers* av faktatypene

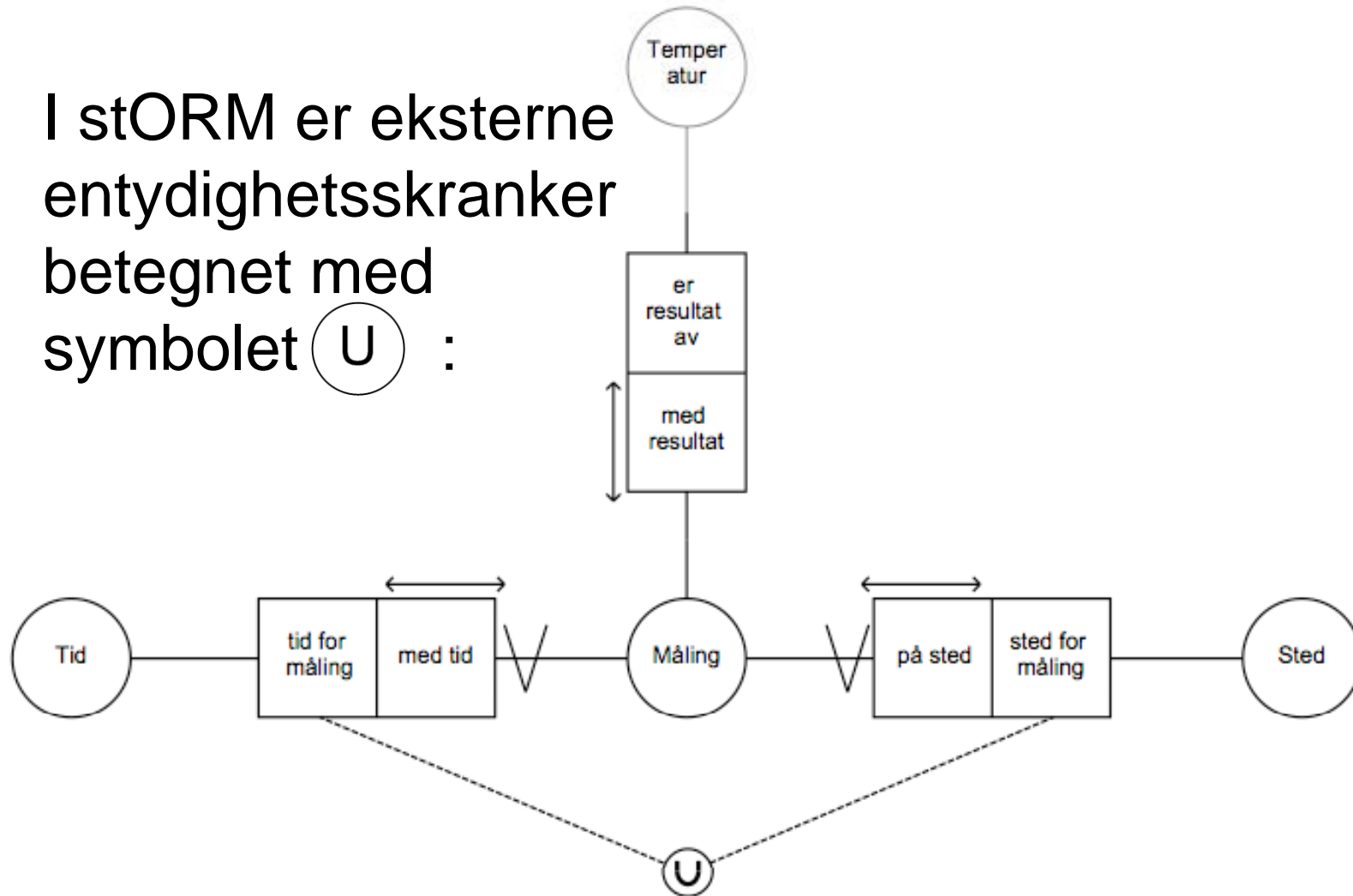
Eksterne entydighetsskranker

- Entydighet på tvers av faktatyper indikerer vi med en **ekstern entydighetsskranke** på de involverte rollene *tid* og *sted*

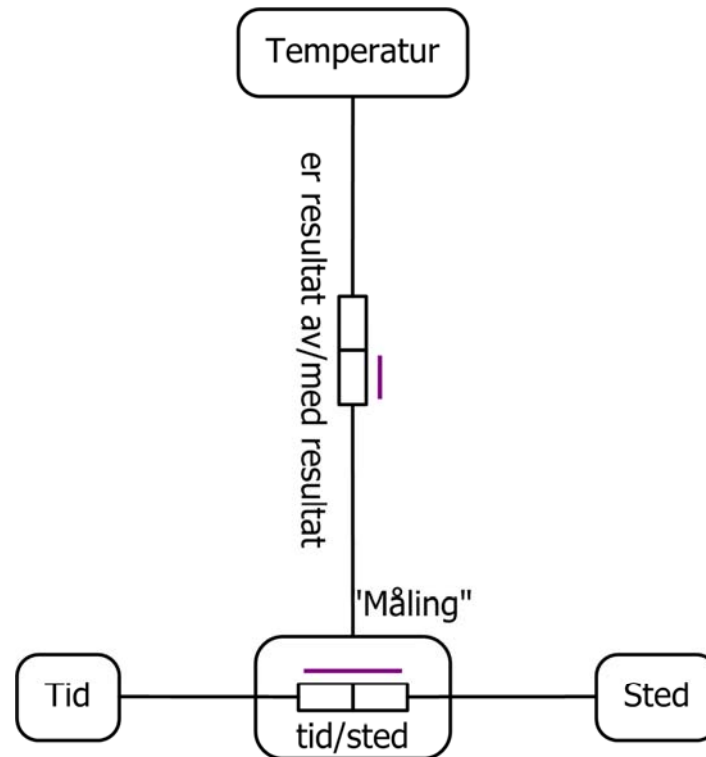


Eksterne entydighetsskranker

I stORM er eksterne entydighetsskranker betegnet med symbolet \textcircled{U} :

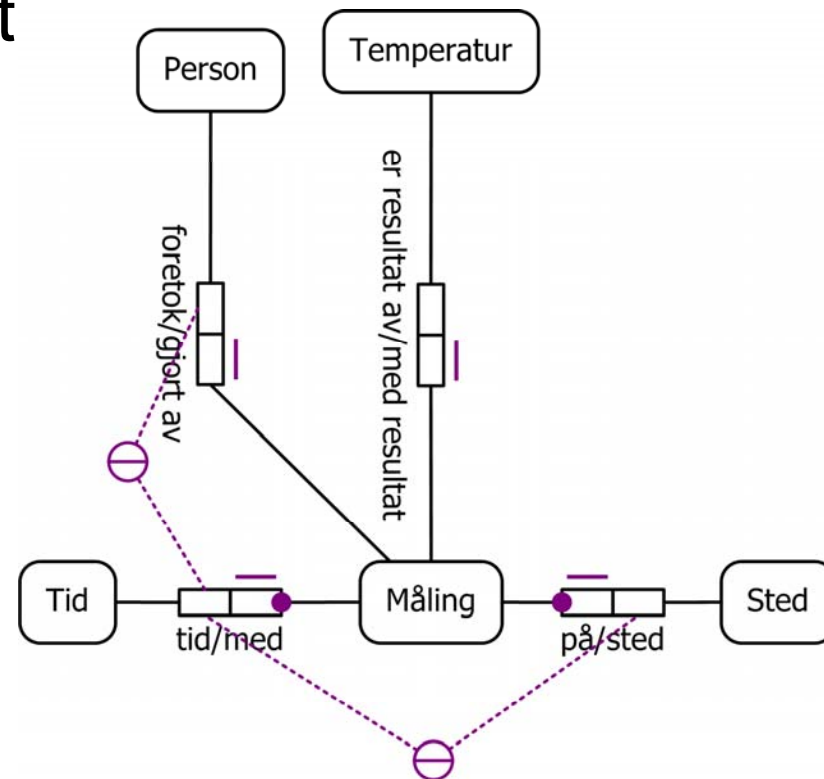


Begrepsdannelse – alternativ notasjon



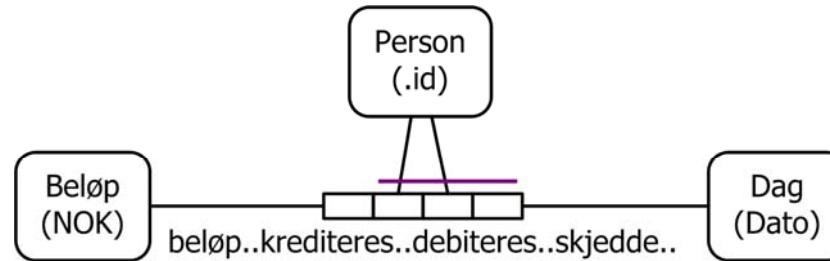
Eksterne entydighetsskranke

- Vi kan også uttrykke at en person ikke kan foreta mer enn én måling av gangen
- Dette gjøres med en ekstern entydighets-skranke på rollene *foretok* og *tid*

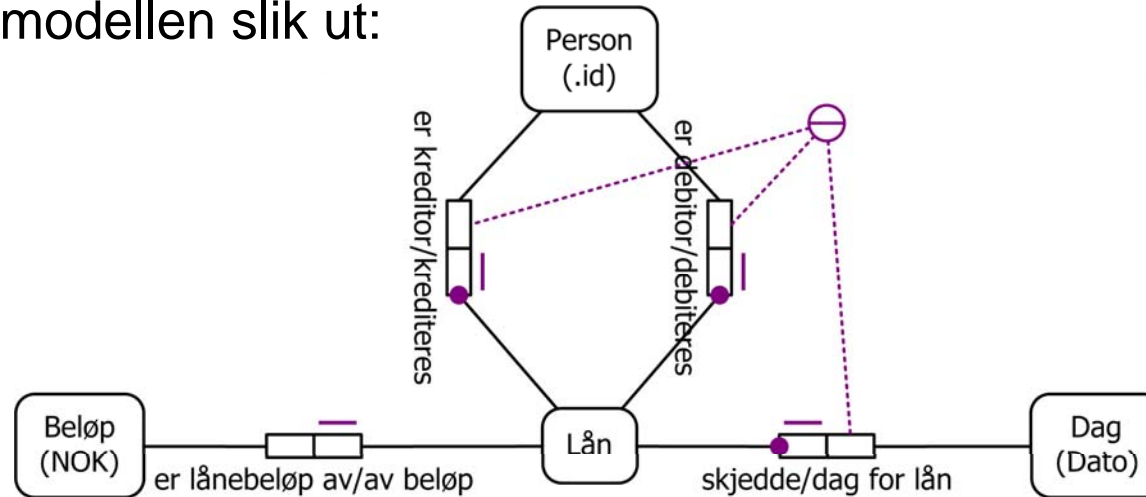


Eksempel med aritet 4

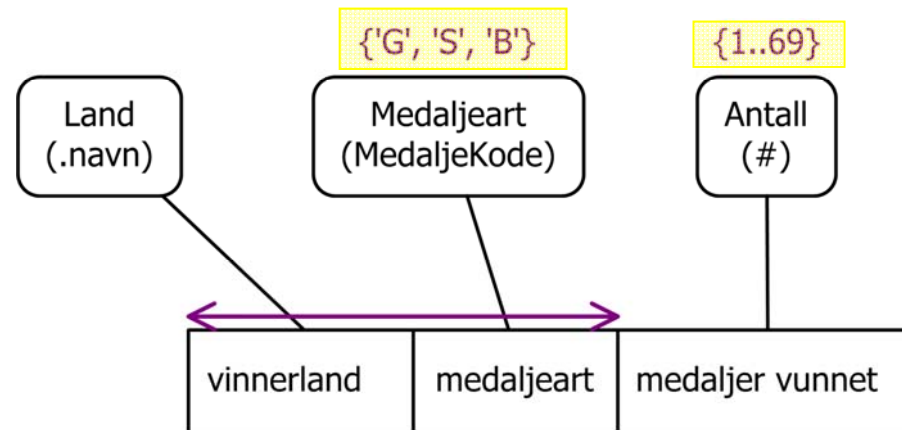
- På en gitt dag låner en person, kalt debitor, et beløp fra en annen person, kalt kreditor



- Det nye begrepet, **Lån**, består av én dag og to personer
- Hvordan vi modellerer «en annen person», kommer vi tilbake til. Forøvrig ser modellen slik ut:



Verdiskranker



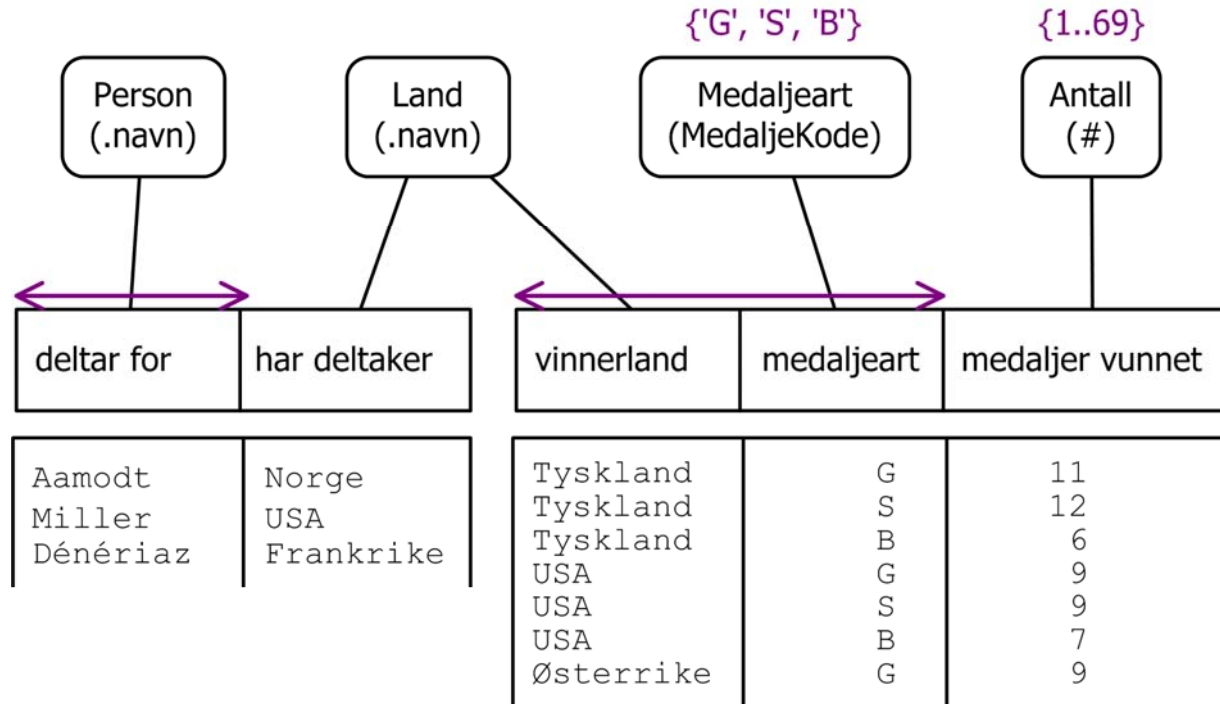
Verdiskranker

- Begrenser mulige forekomster av et begrep
- I praksis: Angir en mengde verdier som er lovlige representasjoner – f.eks. ved direkte opprams av verdiene, angivelse av en nedre og/eller øvre grense, innebygde kontrollcifre (eks. fødselsnumre)

Populasjoner

- **Populasjon for en rolle:** Hvis r er en rolle, betegner $\text{pop}(r)$ mengden av forekomster i kolonnen for r i forekomsttabellen
- **Populasjon for et begrep:** Begreper har egentlig ikke forekomster løsrevet fra roller, men vi definerer likevel populasjonen til et begrep A som har roller r_1, r_2, \dots, r_n ved
$$\text{pop}(A) = \text{pop}(r_1) \cup \text{pop}(r_2) \cup \dots \cup \text{pop}(r_n)$$
- **Merk:** Populasjonen til en rolle/et begrep varierer med tiden

Populasjoner - eksempel



$\text{pop}(\text{Land som vinnerland}) = \{\text{Tyskland, USA, Østerrike}\}$
 $\text{pop}(\text{Land som har_deltaker}) = \{\text{Norge, USA, Frankrike}\}$
 $\text{pop}(\text{Land}) = \{\text{Tyskland, USA, Østerrike, Norge, Frankrike}\}$

Mengdeskranker

- Mengdeskrankene begrenser mengden av forekomster i en eller flere roller i forhold til forekomstene i andre roller
- Finnes i følgende varianter:
 - Likhetskrank
 - Ulikhetskrank
 - Delmengdeskrank

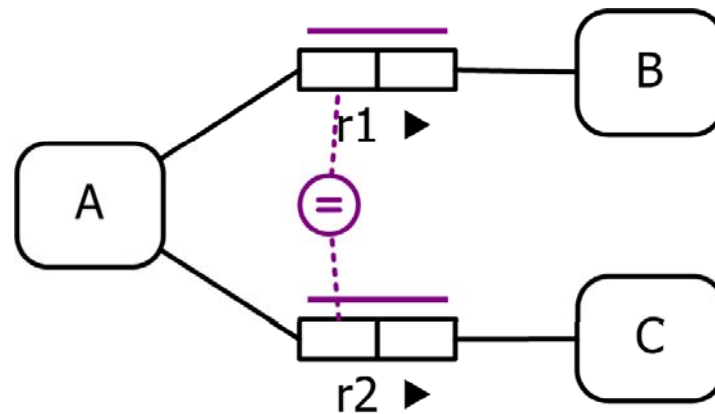
Mengdelikhetsranken

A skal ha rollen r1

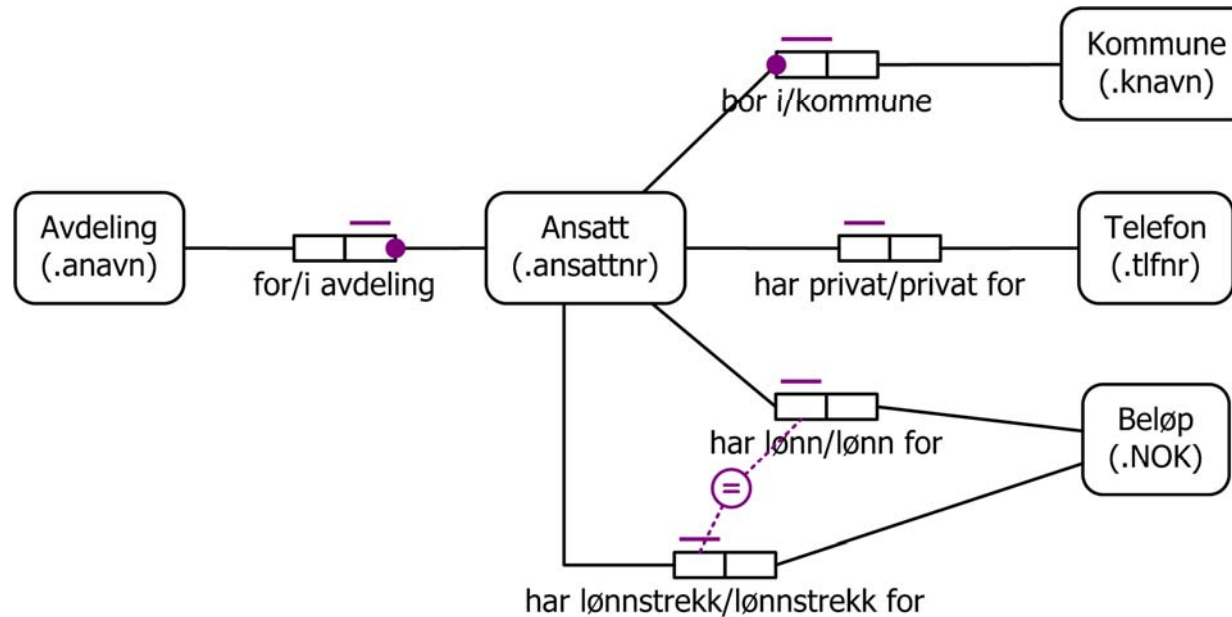
hvis og bare hvis

A har rollen r2.

$\text{pop}(r1) = \text{pop}(r2)$ for alle tilstander



Mengdelikhetsranken, eksempel

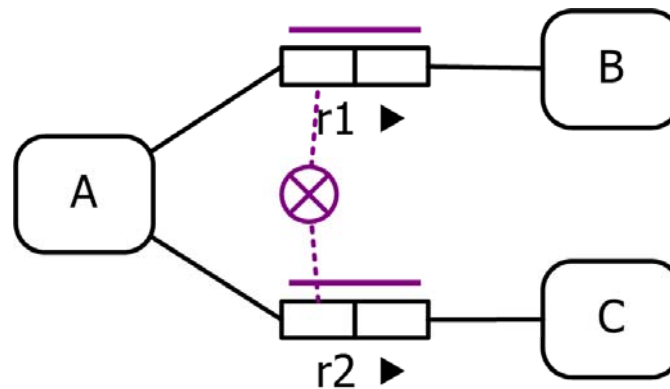


Konsekvens: Oppdatering av **har lønn** krever oppdatering av **har lønnstrekk** og omvendt

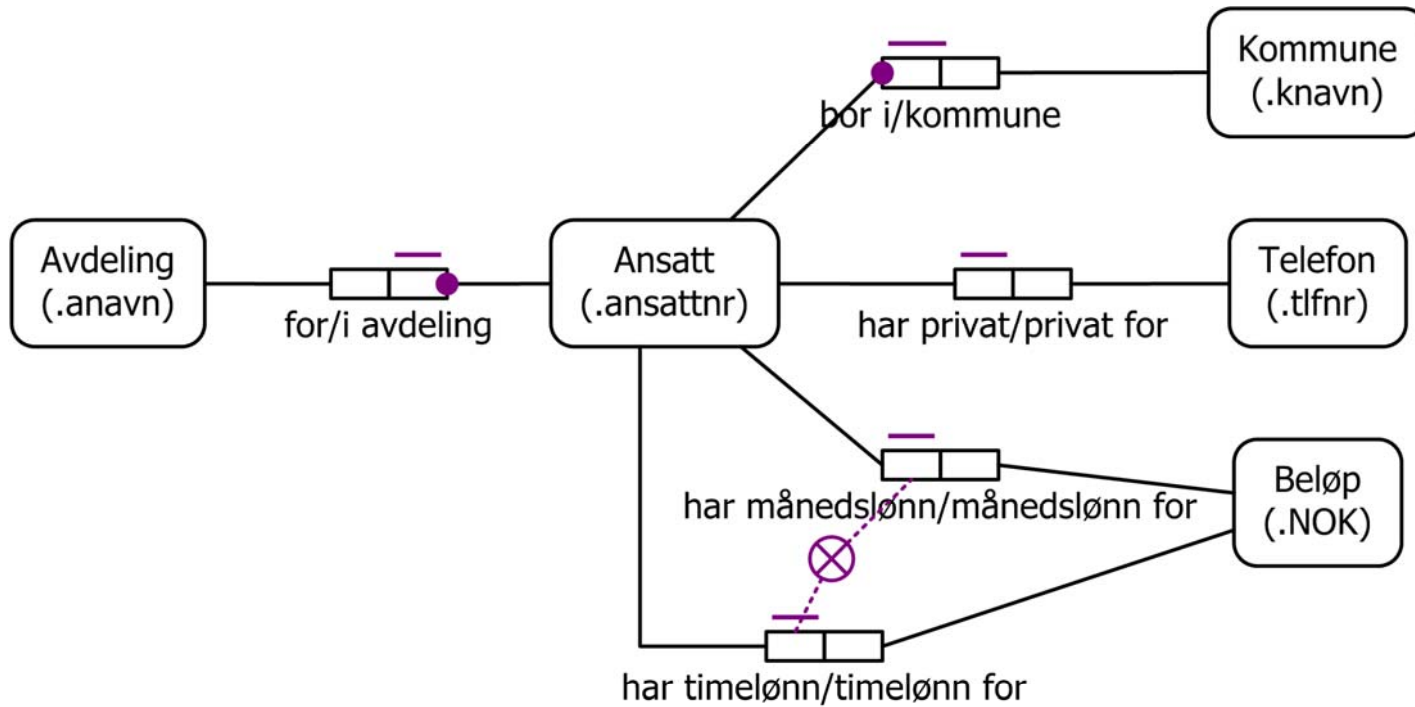
Mengdeulikhetsranken

A skal ikke ha
både rollen r1 og r2.

$\text{pop}(r1) \cap \text{pop}(r2) = \emptyset$
for alle tilstander

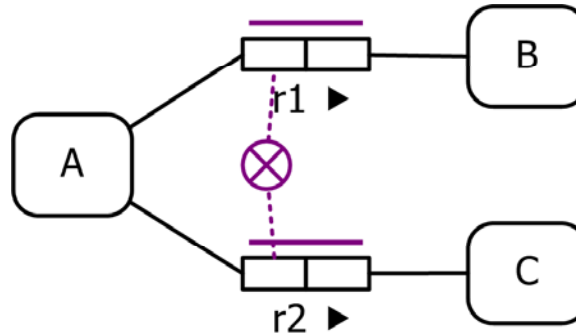


Mengdeulikhetsranken, eksempel

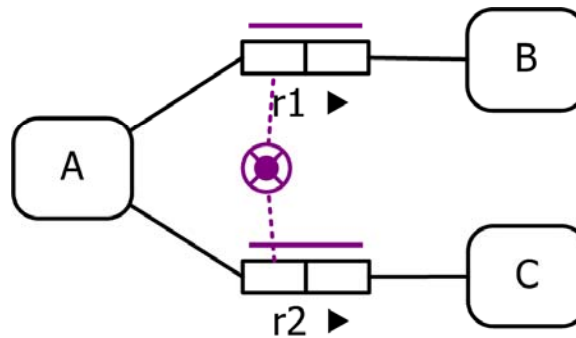


Mengdeulikhetsranken

A skal ikke ha både rollen r1 og r2. Det kan være forekomster av A som hverken er i r1 eller r2

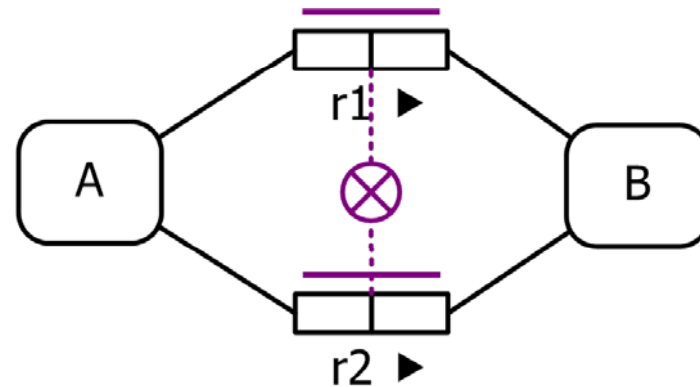


A skal ha en og bare en av rollene r1 og r2

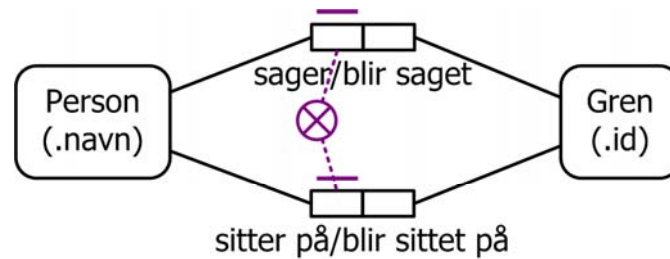


Mengdeulikhetskranke over to roller

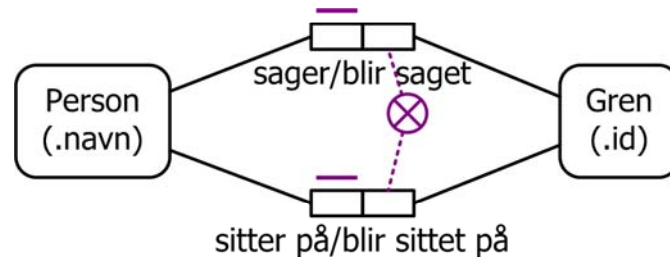
Det skal ikke være forekomster av A og B som er relatert gjennom begge faktatypene



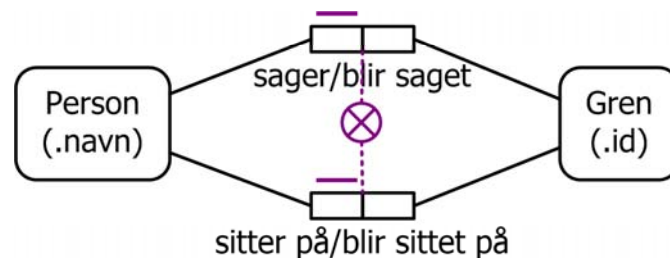
Mengdeulikhetsranken, eksempel



"Sag ikke av en gren når du sitter på en annen gren"

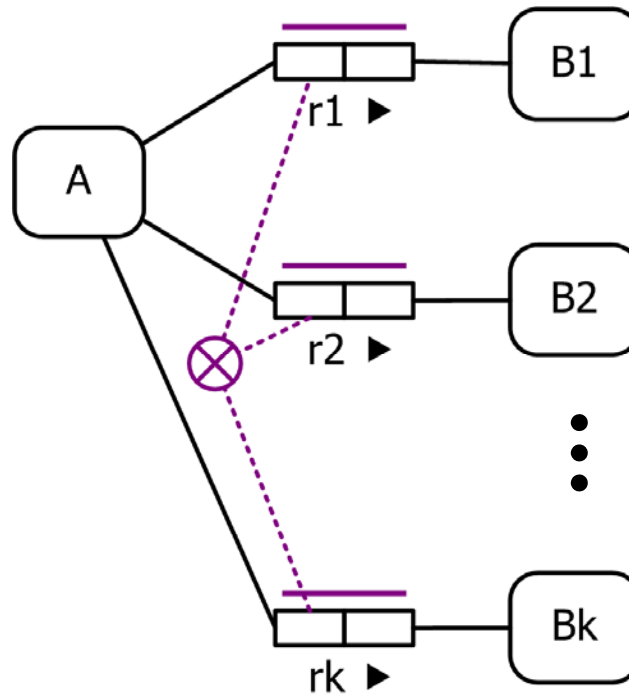


"Sag ikke av en gren som noen andre sitter på"



"Sag ikke av den grenen du selv sitter på"

Den generelle mengdeulikhetskranken

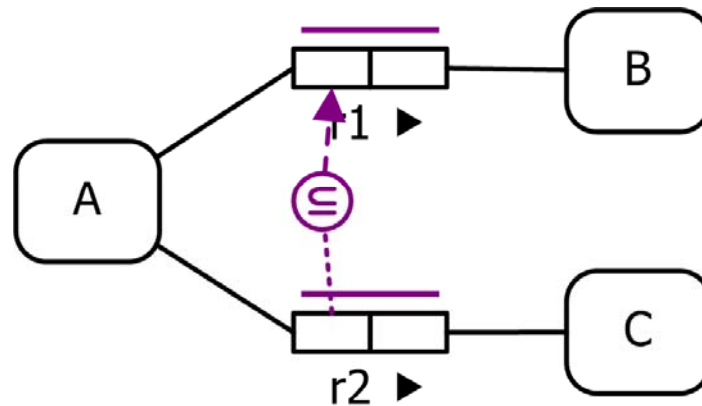


$$\text{pop}(r_i) \cap \text{pop}(r_j) = \emptyset \text{ for alle } i \neq j (1 \leq i \leq k, 1 \leq j \leq k)$$

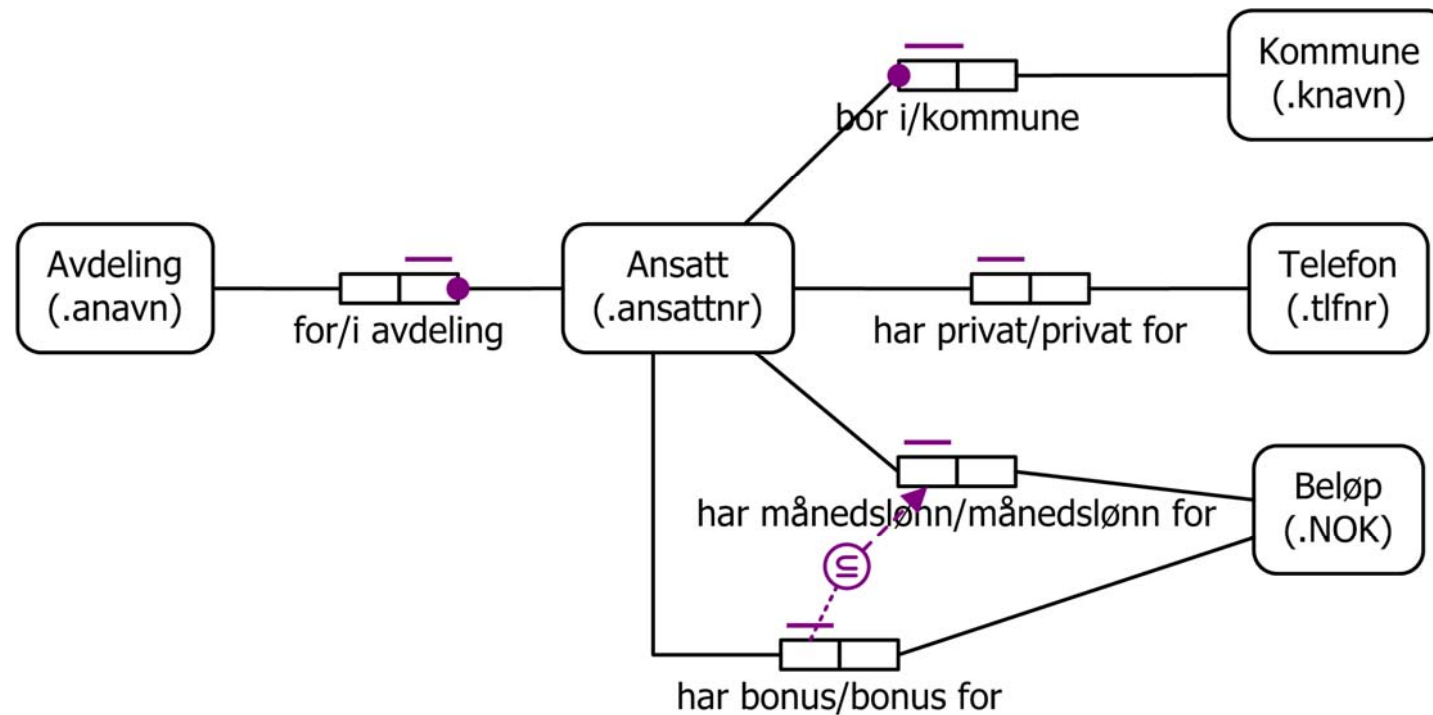
Delmengdeskranken

Hvis A har rollen r2,
så skal A også ha rollen r1.

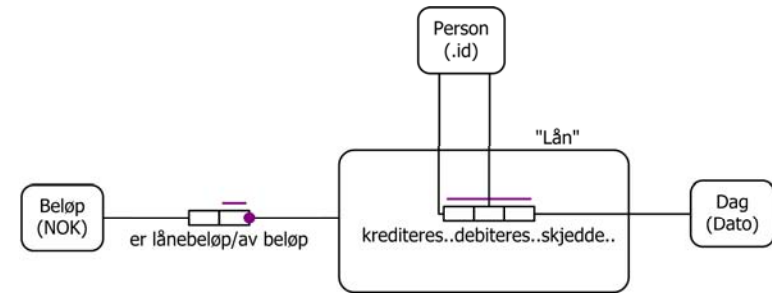
$\text{pop}(r2) \subseteq \text{pop}(r1)$ for alle tilstander



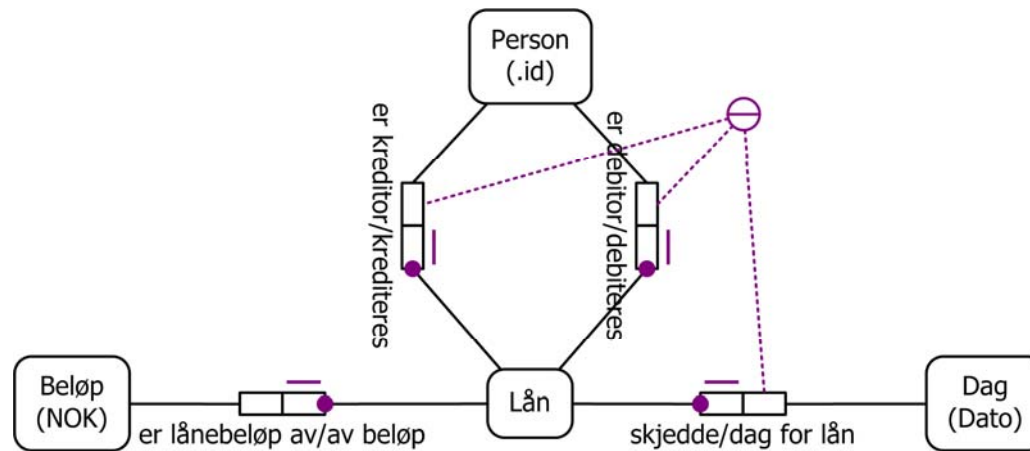
Delmengdeskranken, eksempel



Eksempel med aritet 4 – forts.



På en gitt dag låner en person – debitoren - et beløp fra en **annen** person - kreditoren

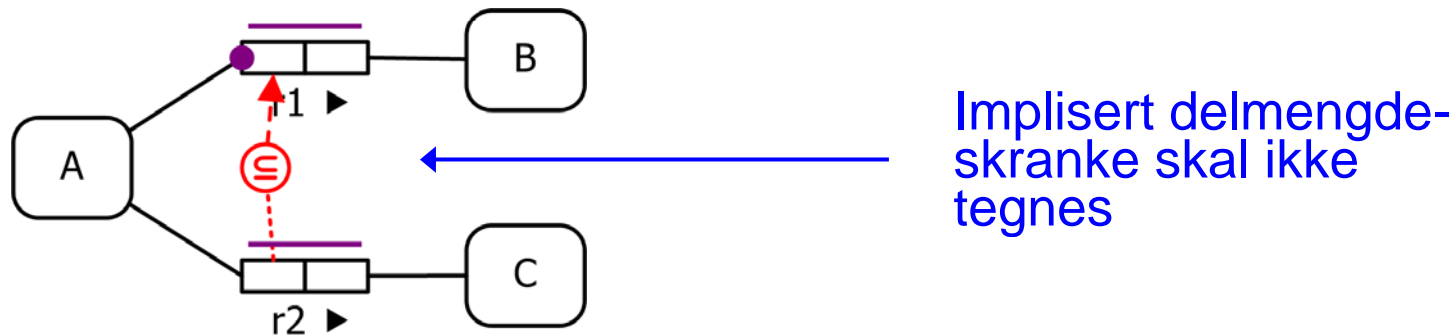
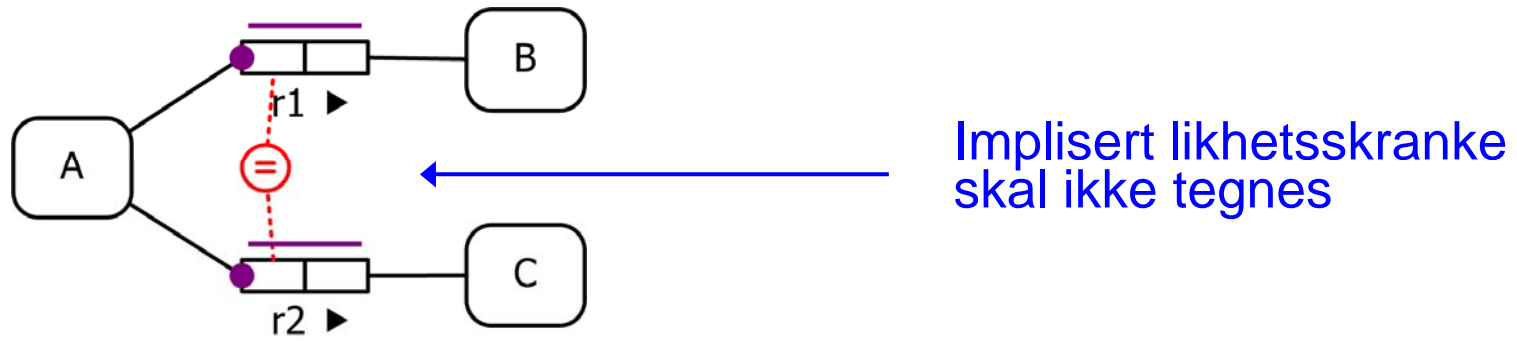


Oppgave: Hvilken skranke mangler?
Hvor skal den plasseres?

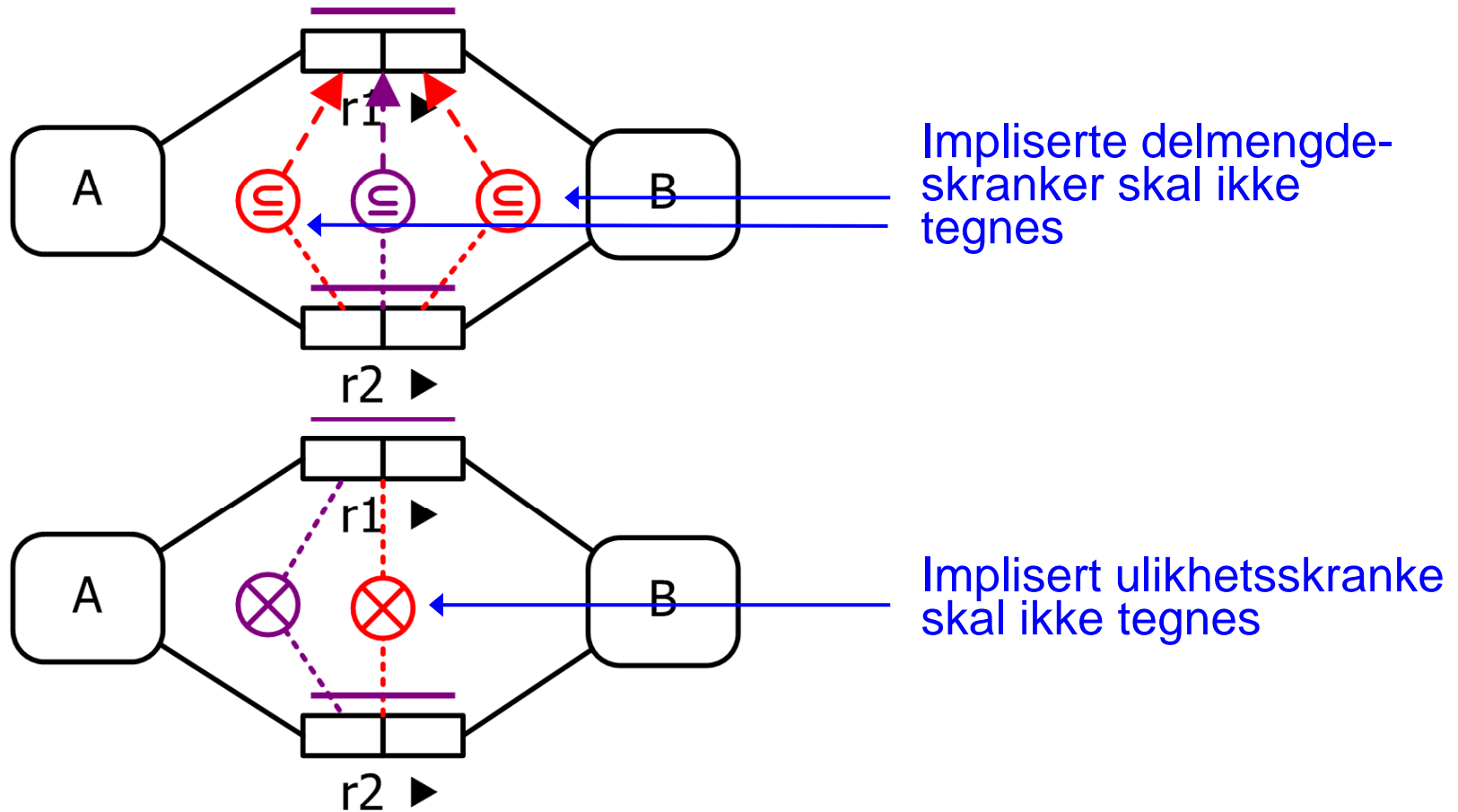
Oppgaver

- Ikke selg skinnet før bjørnen er skutt
- Kast ikke stein når du sitter i glasshus

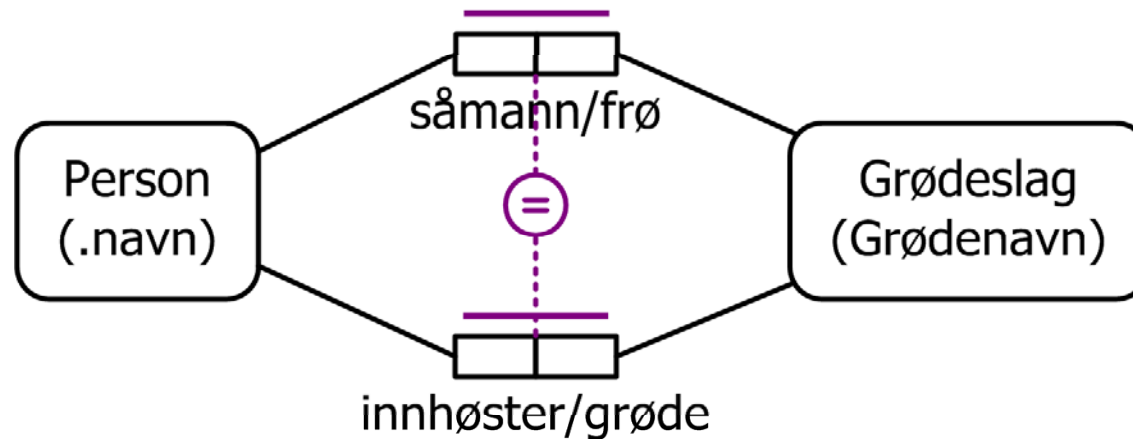
Impliserte skranke skal ikke tegnes



Impliserte skranke skal ikke tegnes

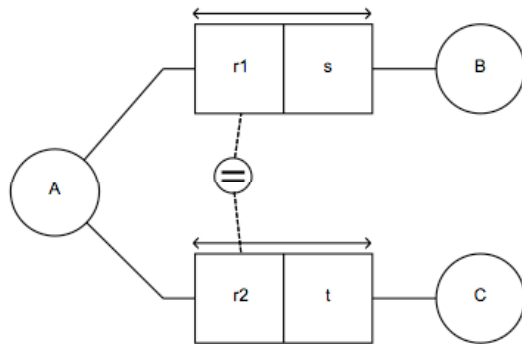


Mengdelikhetsskranke over flere roller er vanligvis ikke lov

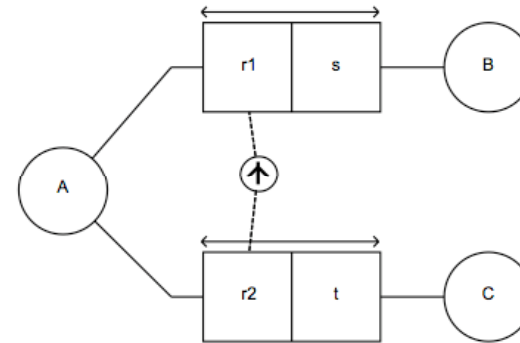


Mengdeskranker i stORM

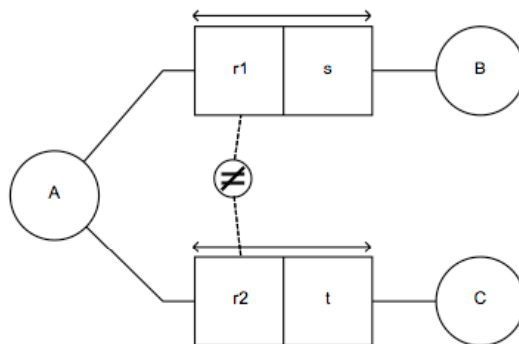
Mengdelikhetskranker:



Delmengdeskranker:



Mengdeulikhetskranker:



En og bare en rolle:

