



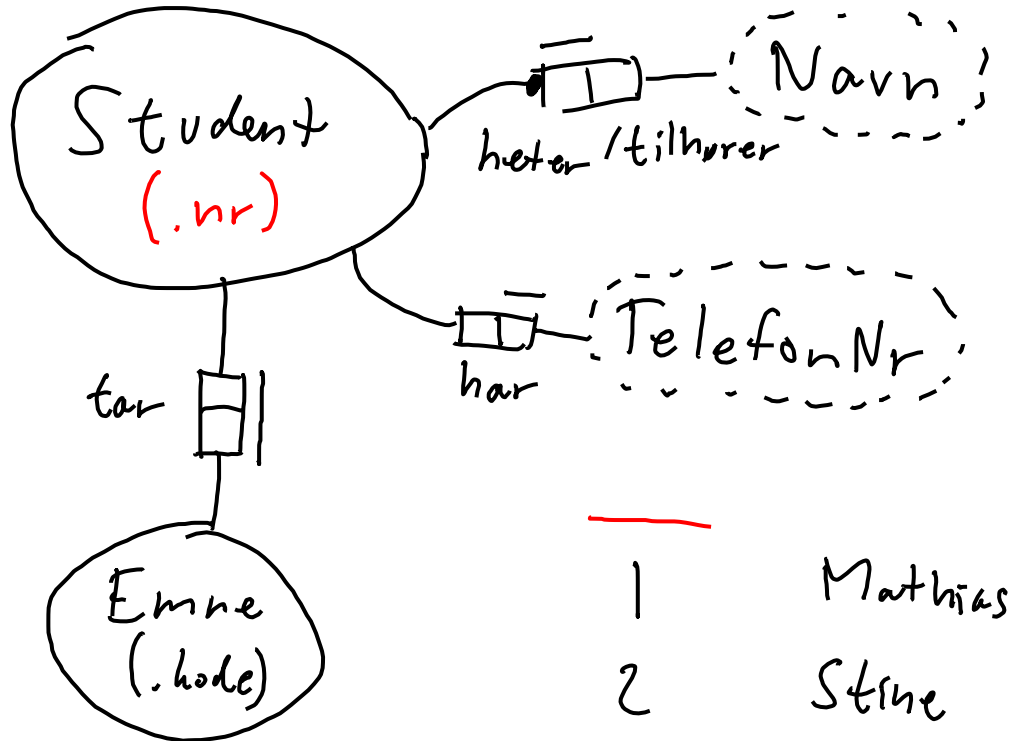
INF1300

Introduksjon til databaser

Dagens tema:

- Begrepsdannelse
- Eksterne entydighetsskranker
- Representasjon
- n-1-regelen
- Verdiskranker
- Mengdeskranker

Fra forrige forelesning



1	Matthias
2	Stine
1	~

interesseområde (UoD, domene)

begrep

verditype
(representasjonstype)

elementær setning

forekomst

fakta

faktatype

objekttype

en setnings aritet

rolle

bro

setningstype

entydighetsskranke

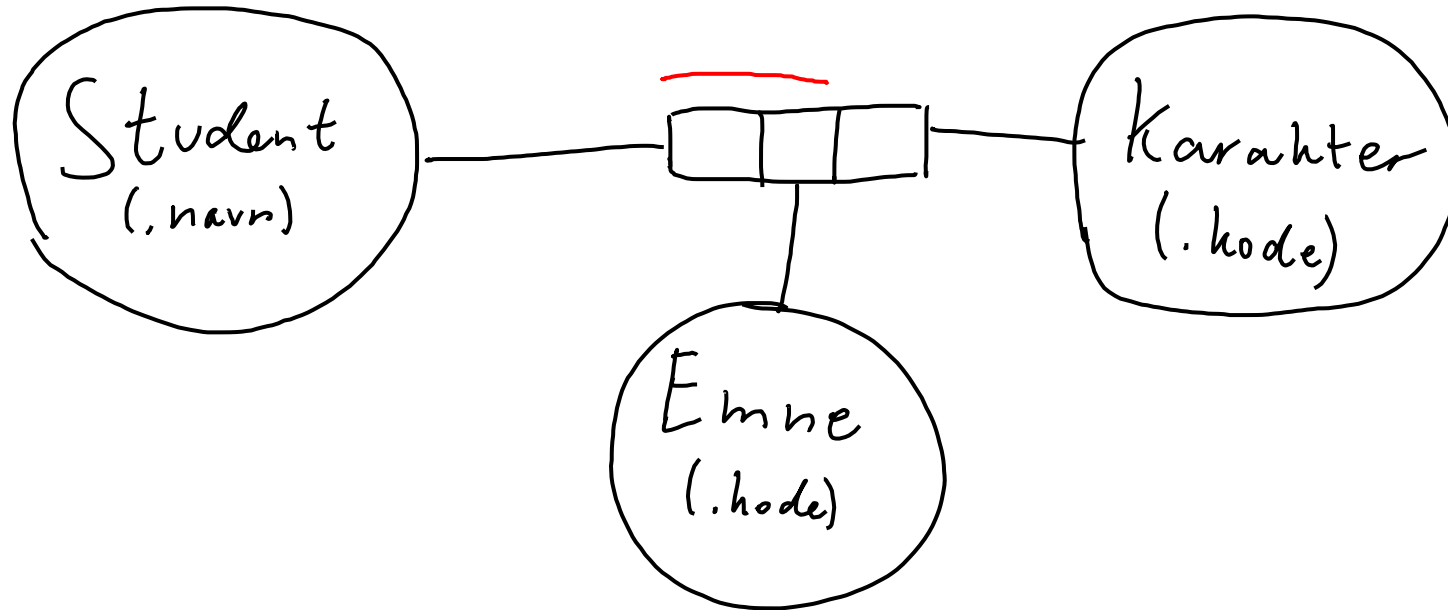
forekomsttabell

påkrevd rolle

perfekt bro

Entydighetsskranker på faktatyper med høyere aritet

«Student har fullført emne med karakter»



C	INF1300	A
A	INF1300	F
B	INF1300	B

~~A INF1300 A~~

Begrepsdannelse

- Alle mange-til-mange-entydighetsskranger ($m:n$) kan være opphav til et nytt begrep
- For å gjøre dette, kan vi ofte følge en algoritme
- La oss se på faktatypen «Student tar Emne»



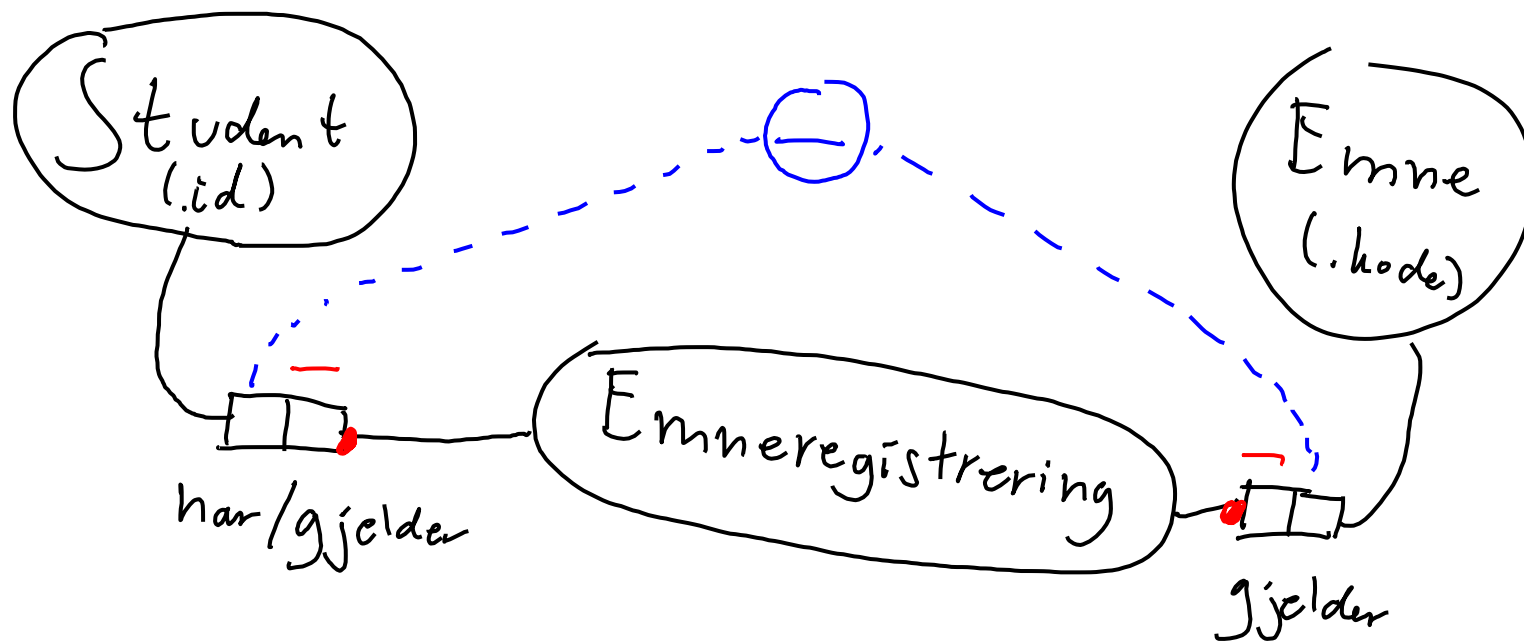
Begrepsdannelse

1 INF1300
 2 INF1300
 3 INF1000
 1 INF1000



1	r1
2	r2
3	r3
1	r4

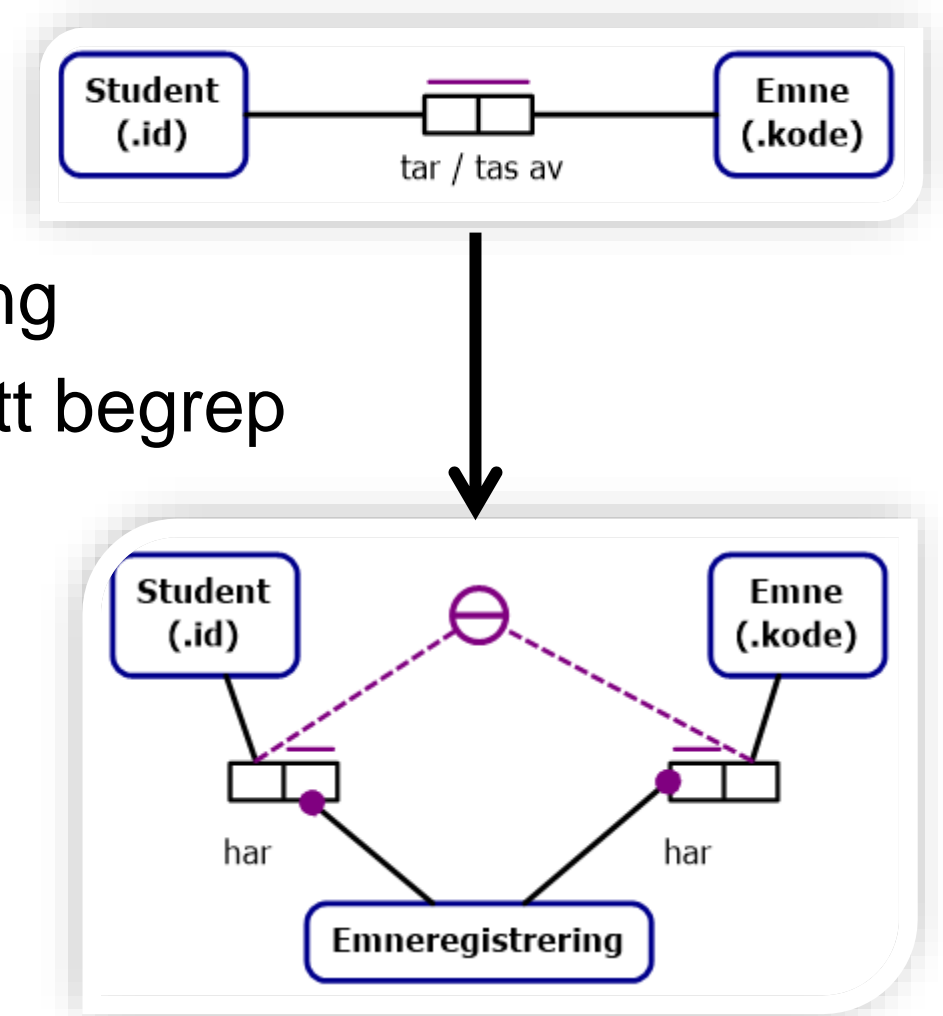
r1	INF1300
r2	INF1300
r3	INF1000
r4	INF1000



Begrepsdannelse

- Litt forenklet oppsummering:

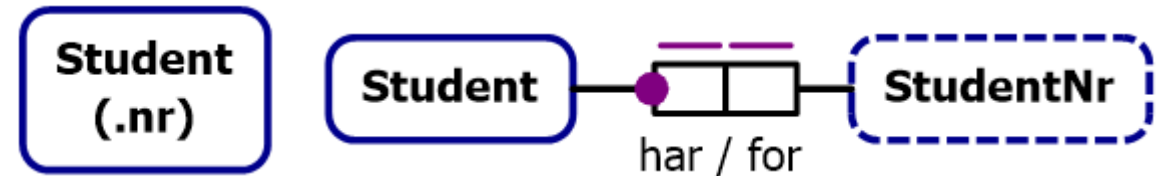
1. Lag begrep for å erstatte $m:n$ -setning
2. Sett $n:1$ -entydighetsskranke fra nytt begrep
3. Sett påkrevde roller
4. Sett eksternt entydighetsskranke



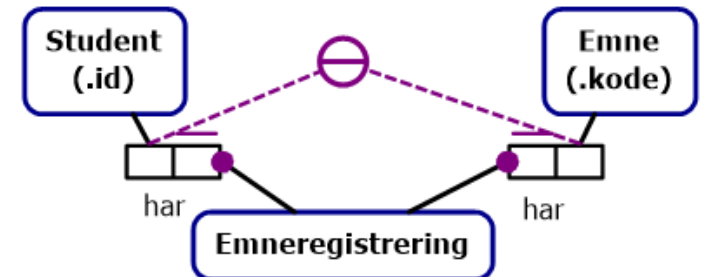
Representasjon

- Alle begreper må ha en representasjon på en av følgende måter:

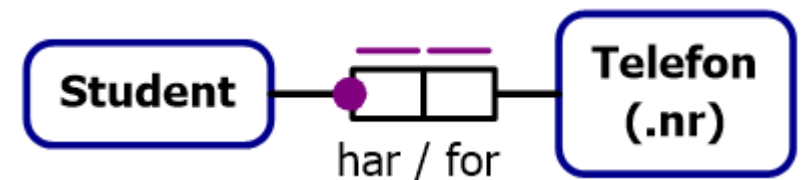
1. Verditype som er knyttet til begrepet med perfekt bro



2. Gjennom ekstern entydighetskranke og påkrevde roller (som i en begrepsdannelse)



3. Representasjonen til et annet begrep som er knyttet til begrepet via en 1:1-faktatype med påkrevd rolle



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?

Sted Tidspunkt Person Temp

– Hvilken **representasjon** bruker vi for disse begrepene?

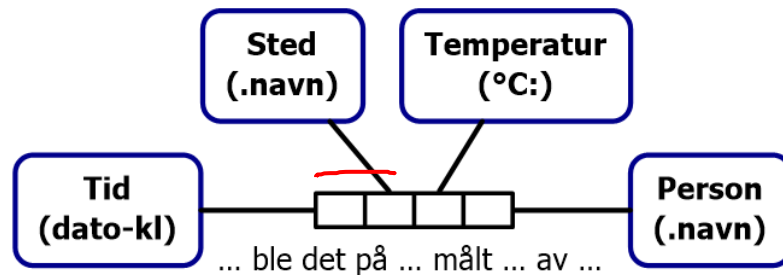
navn klokkeslett navn °C

– 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**, kan vi lage en **forekomsttabell**

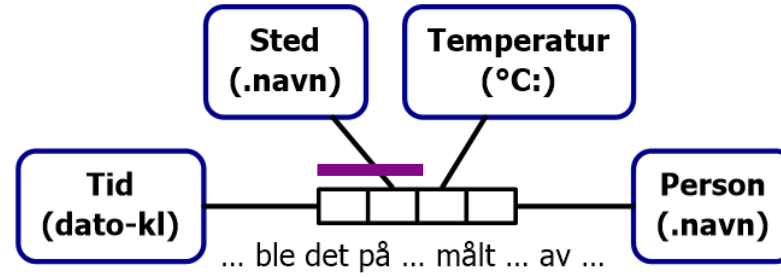
Et eksempel fra virkeligheten



3.8.17 kl 8:00	Blindern	9	Jens
3.8.17 kl 8:00	Værnes	9	Siri
3.8.17 kl 8:00	Blindern	11	Lars
4.8.17 kl 8:00	Blindern	9	Jens

- Er det noen av forekomstene som må strykes?
- Hvor skal entydighetsskranken(e) stå?

Et eksempel fra virkeligheten



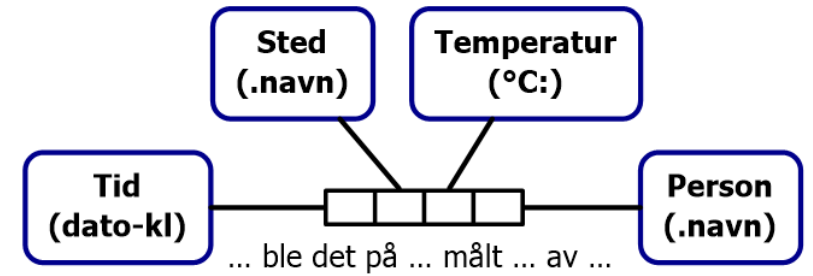
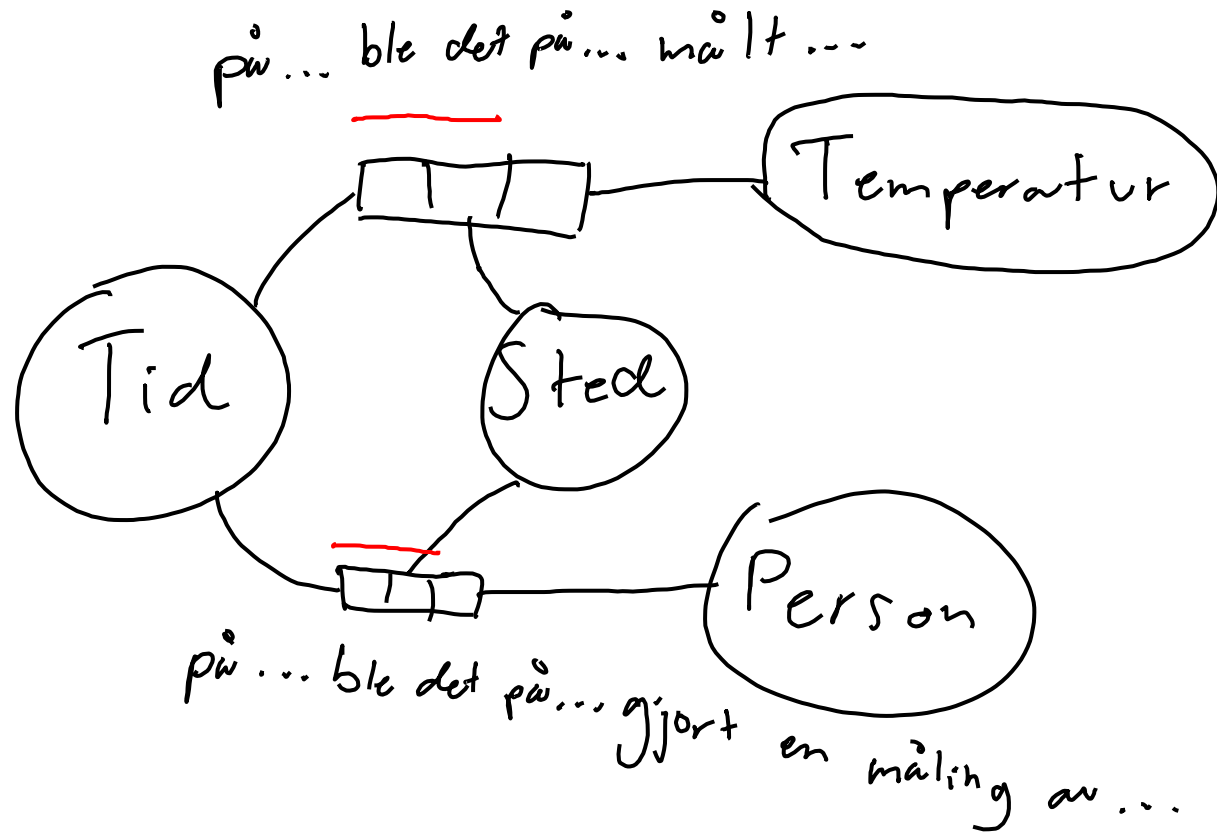
3.8.17 kl 8:00	Blindern	9	Jens
3.8.17 kl 8:00	Værnes	9	Siri
3.8.17 kl 8:00	Blindern	11	Lars
4.8.17 kl 8:00	Blindern	9	Jens

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

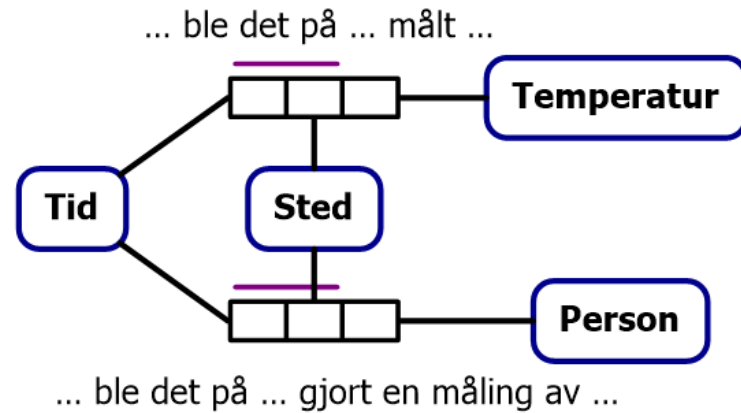
n-1-regelen

- *En setning med aritet n er aldri elementær hvis den tilhørende forekomsttabellen 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*

Hvordan gjør vi denne gyldig?



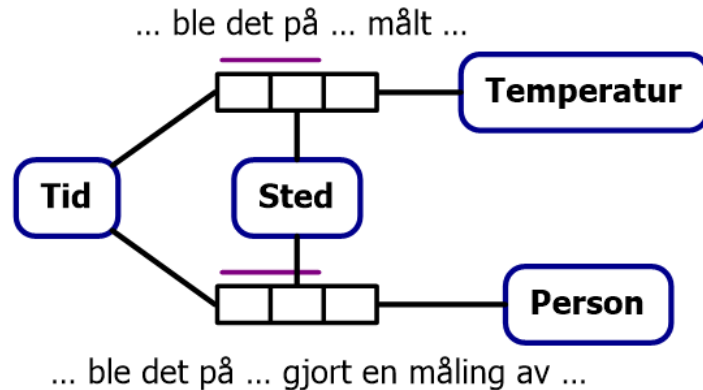
Et eksempel fra virkeligheten



8 ~~Oslo~~ 11°C
8 Bergen 11°C
1 Oslo 11°C

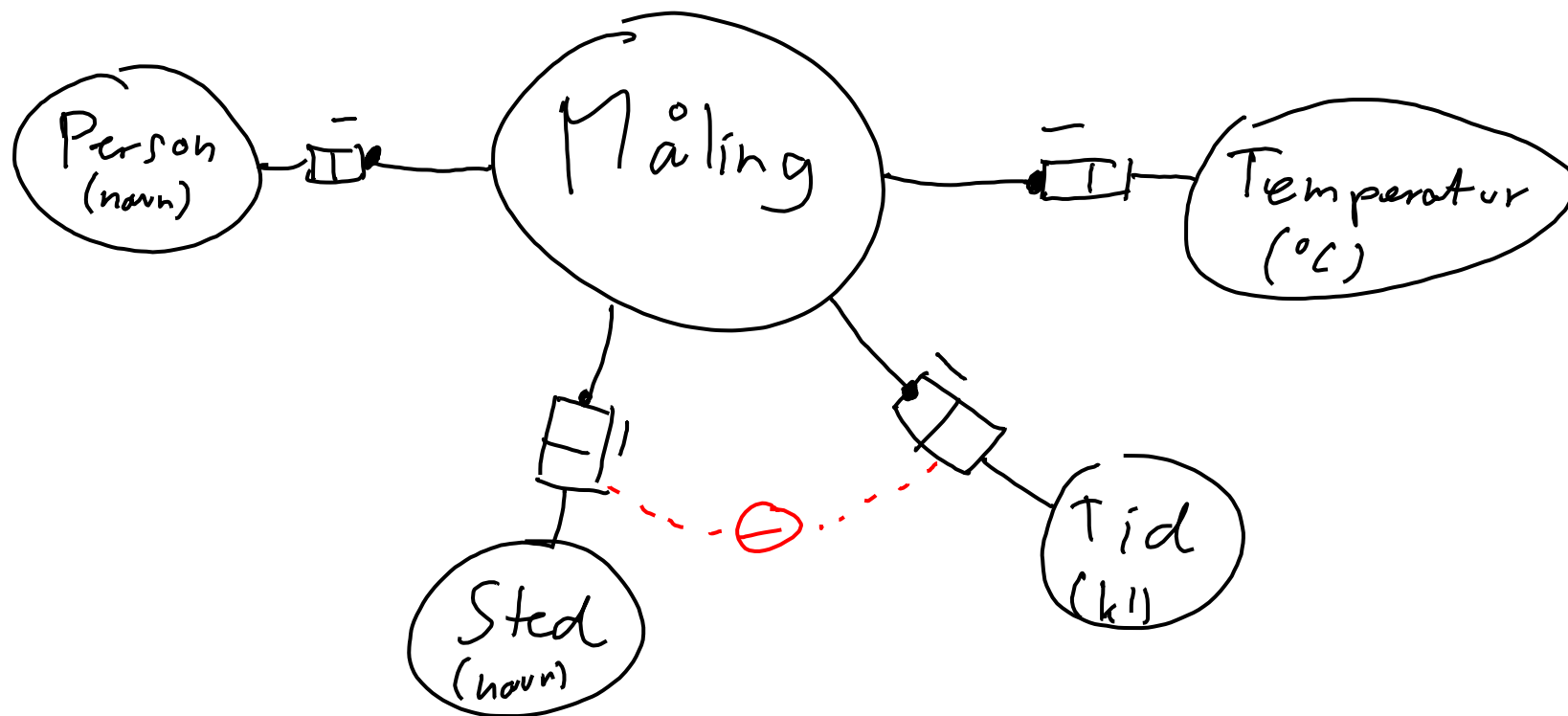
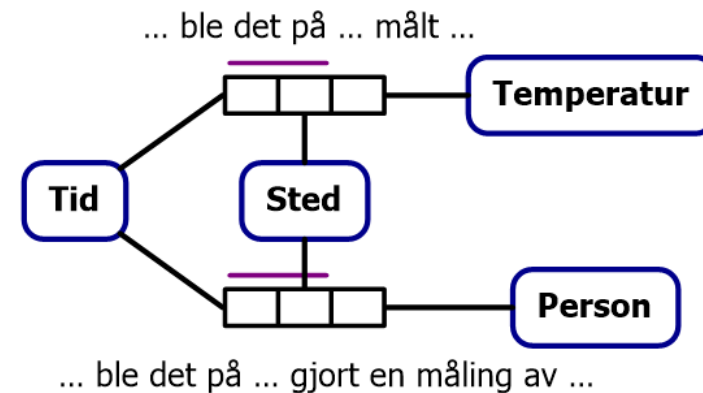
- Vi har nå et skjema for elementære ternære setninger med entydighetsskranker av lengde to
- Entydighetsskrankene sier at Tid og Sted sammen bestemmer Temperatur og sammen bestemmer Person.

Et eksempel fra virkeligheten

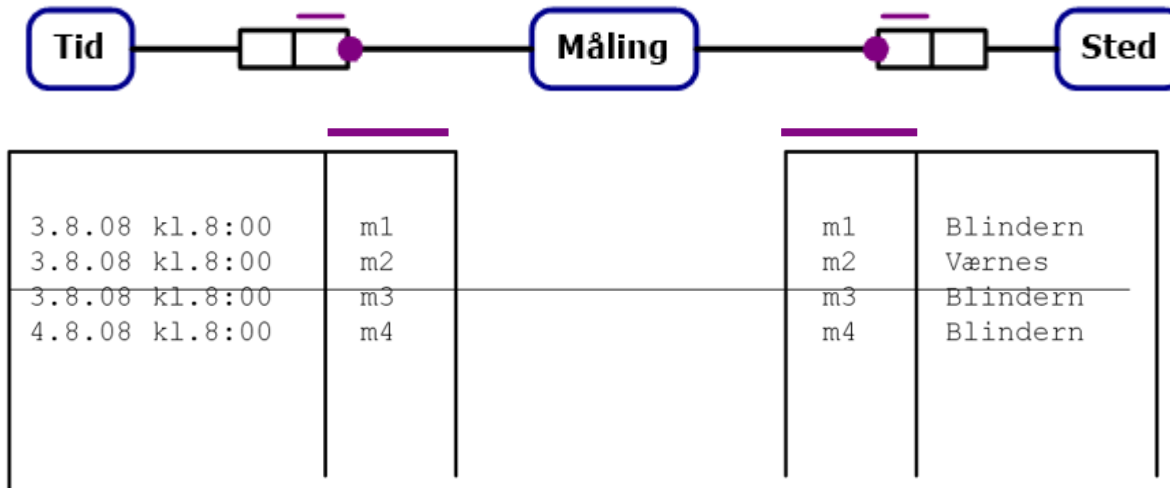


- De to setningene har til felles at Tid og Sted *sammen* bestemmer en tredje verdi
- ORM-diagrammet til venstre er korrekt, men tydeliggjør ikke at for både Temperatur og Person er forekomster bestemt av **en og samme** kombinasjon av Tid og Sted

Begrepsdannelse

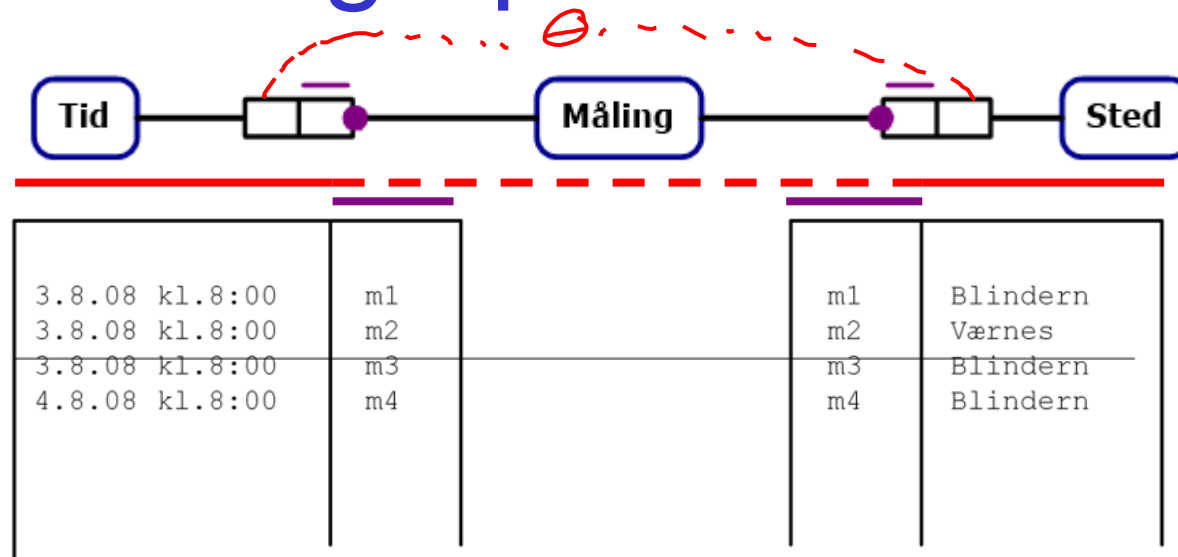


Begrepsdannelse



- 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

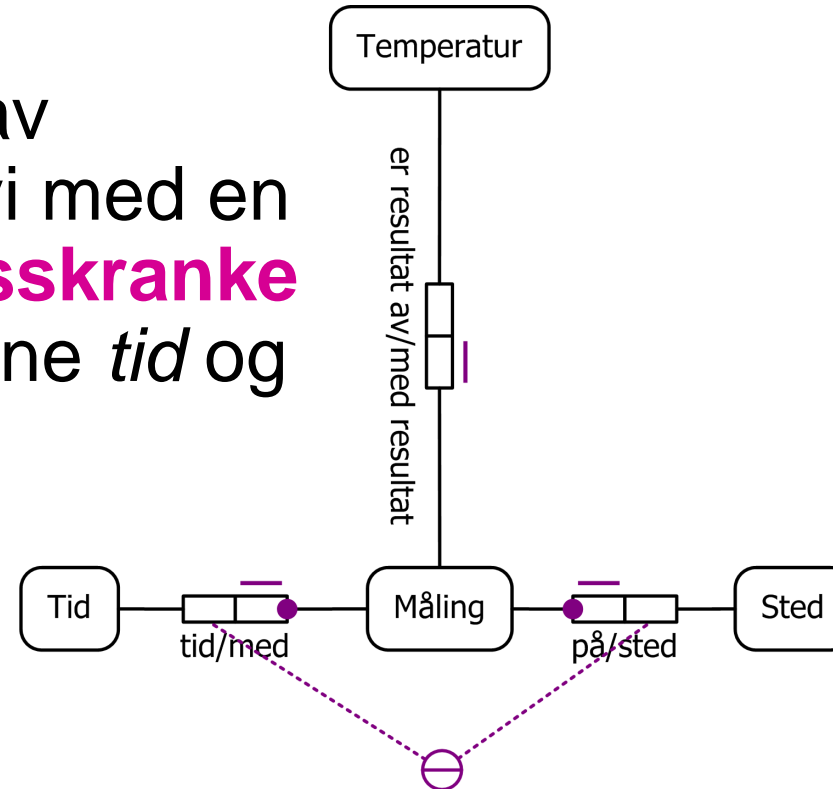
Begrepsdannelse



- 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 entydighetsskranke *på tvers* av faktatypene

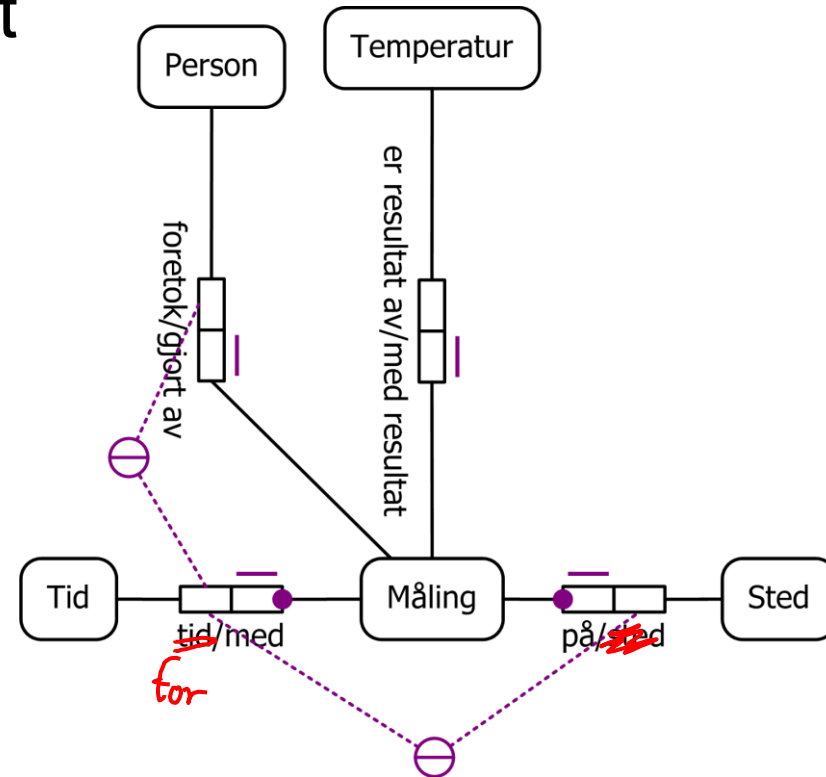
Eksterne entydighetsskranker

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



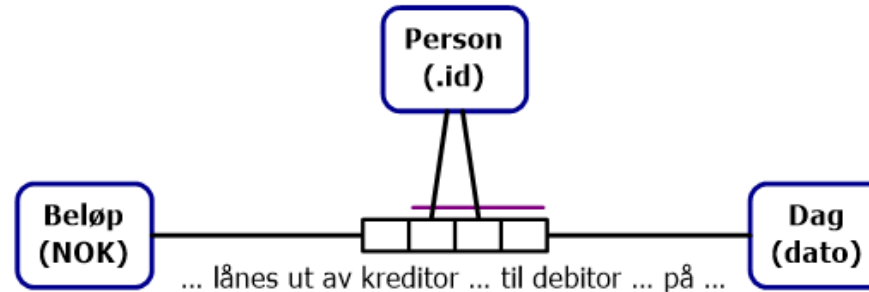
Eksterne entydighetsskranke

- 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*

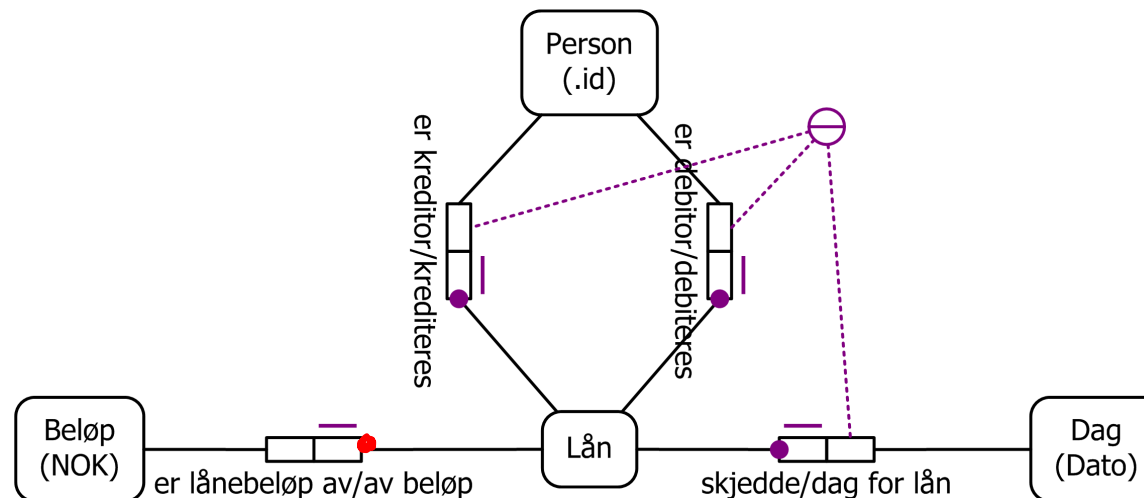


Eksempel med aritet 4

- 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:

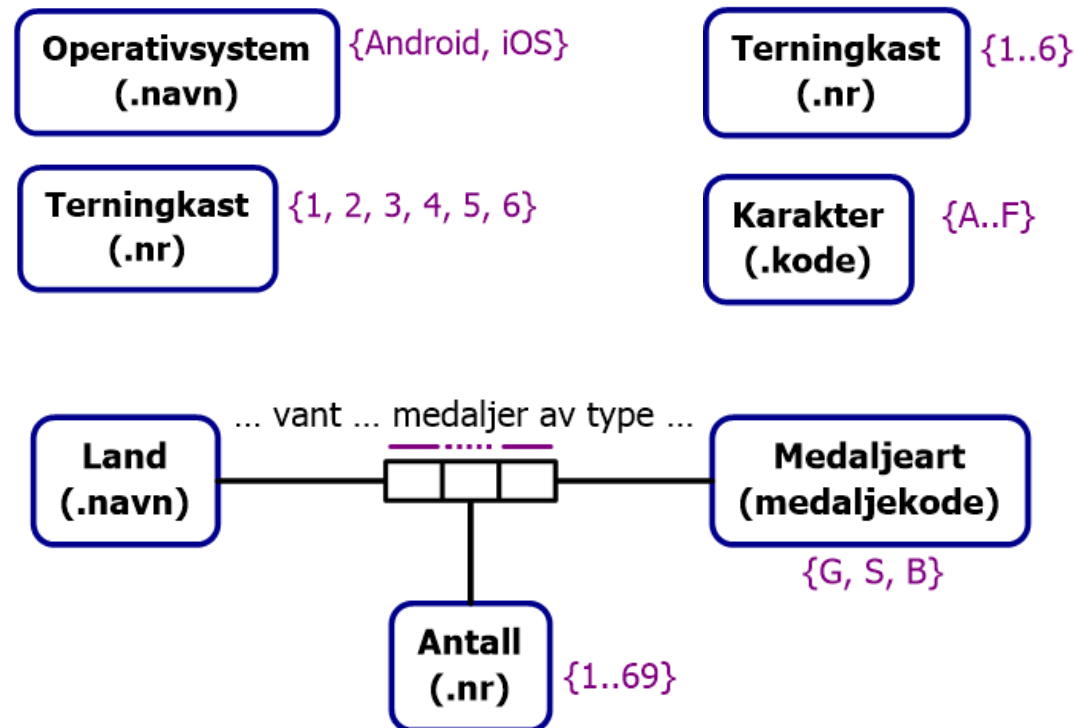


Avsluttende om begrepsdannelser

- Alle entydighetsskranker som går over mer enn én rolle i en faktatype, skjuler et (nytt) begrep. Man skal alltid vurdere om det skal lages nye begreper når man får faktatyper med lange entydighetsskranker.
 - En faktatype med aritet 3 eller 4 (eller mer) kan gjøres om til binære setninger ved å lage ett eller flere nye begreper.
- **IKKE bruk begrepsdannelse om du ikke trenger det – det er enklere (og bedre) å bruke mange-til-mange-skranker!**

Verdiskranker

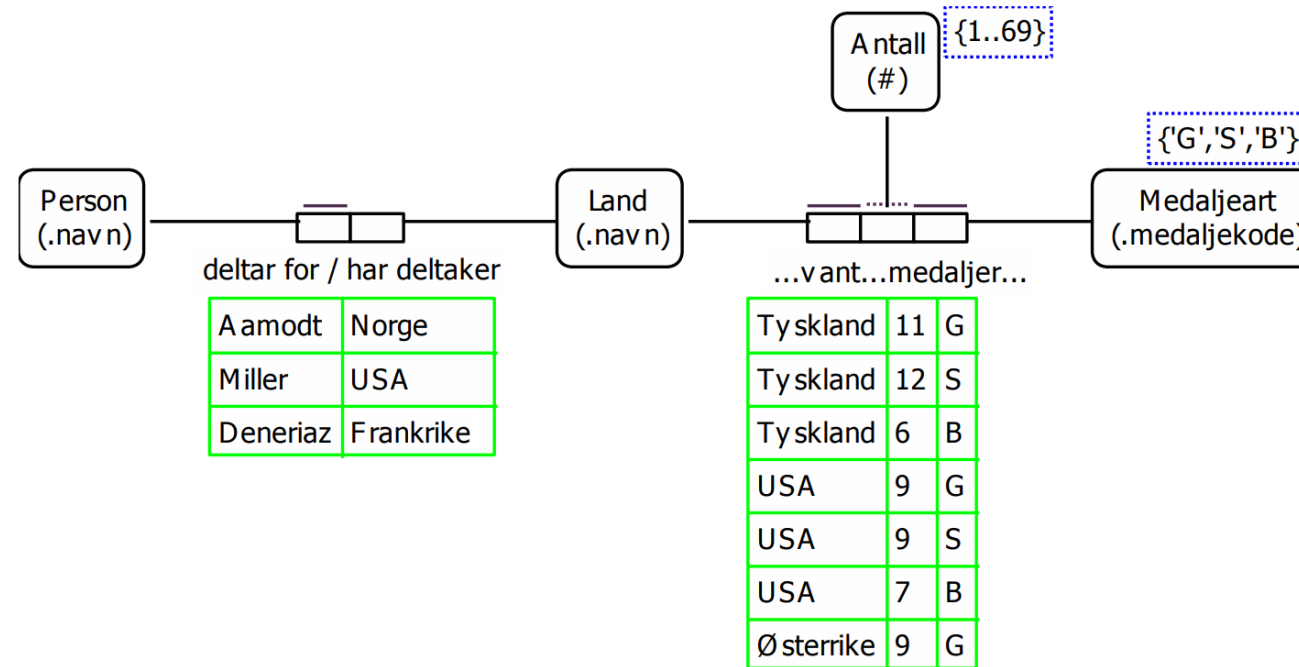
- Begrenser mulige forekomster av et begrep
- I praksis: Angir en mengde verdier som er lovlige representasjoner – f.eks. ved direkte opprams av verdiene eller angivelse av en nedre og/eller øvre grense



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



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

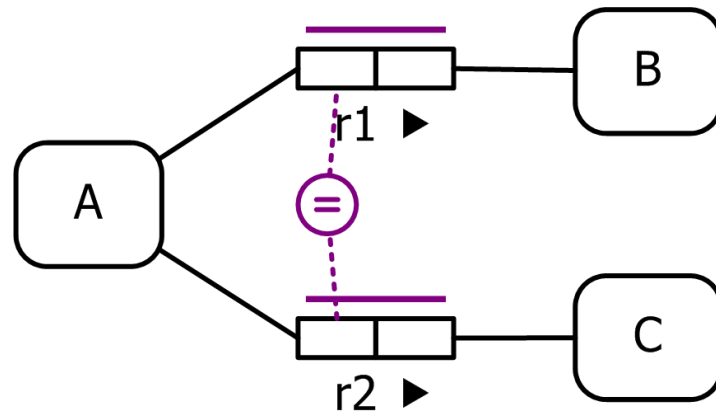
Mengdeskranker

- Mengdeskrankene begrenser mengden av forekomster i en eller flere roller i forhold til forekomstene i andre roller
- Finnes i følgende varianter:
 - Likhetskranker
 - Ulikhetskranker
 - Delmengdeskranker

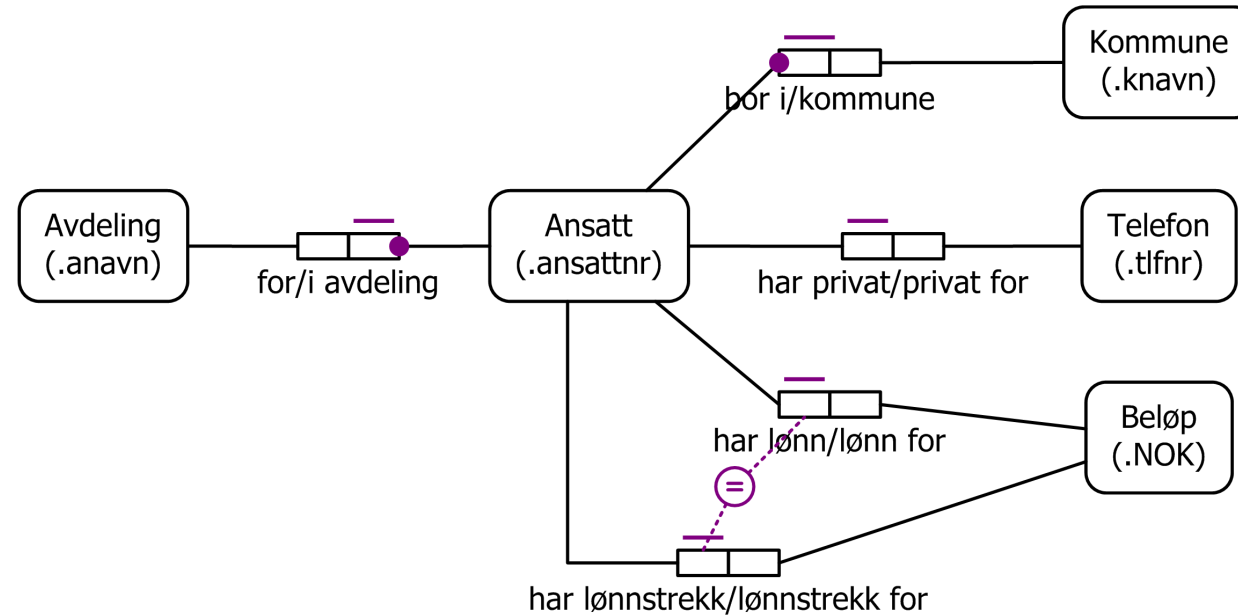
Mengdelikhetsranken

A skal ha rollen r1
hvis og bare hvis
A har rollen r2.

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



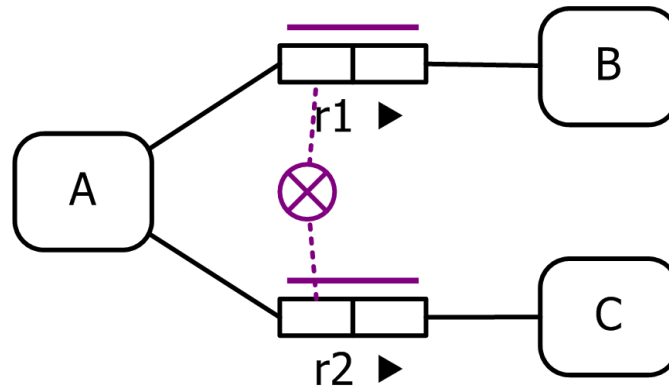
Mengdelikhetsranken, eksempel



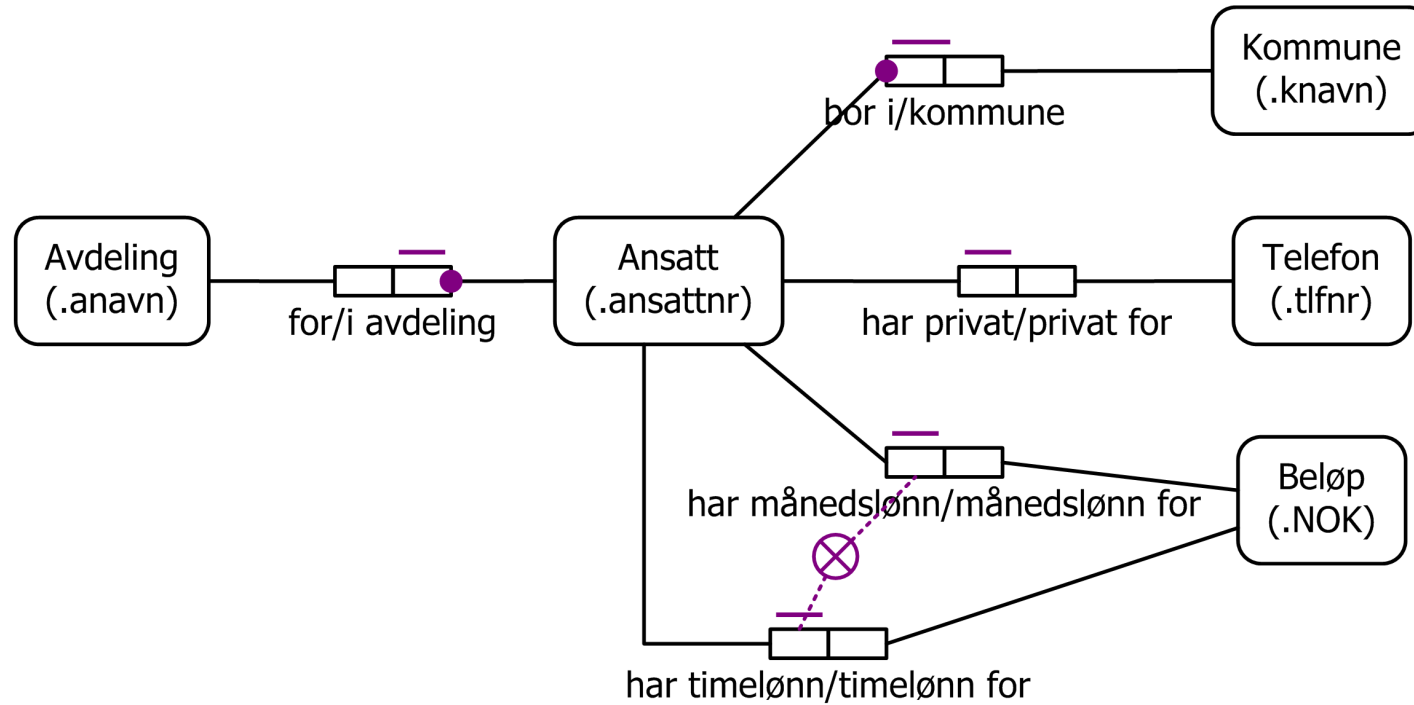
Konsekvens: En person som **har lønn** må også **ha lønnstrekk** og omvendt

Mengdeulikhetsranken

A skal ikke ha
både rollen r1 og r2.
 $\text{pop}(r1) \cap \text{pop}(r2) = \emptyset$
for alle tilstander

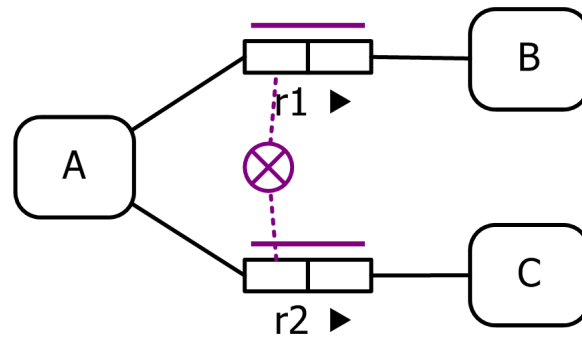


Mengdeulikhetsranken, eksempel

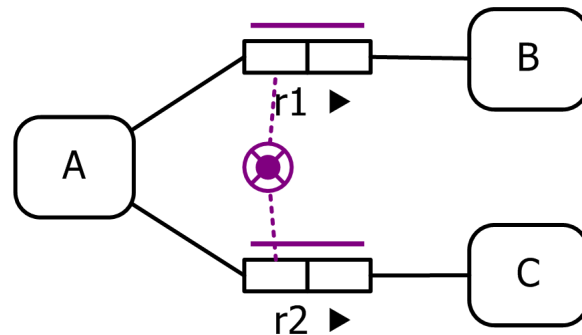


Mengdeulikhetsskranken

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

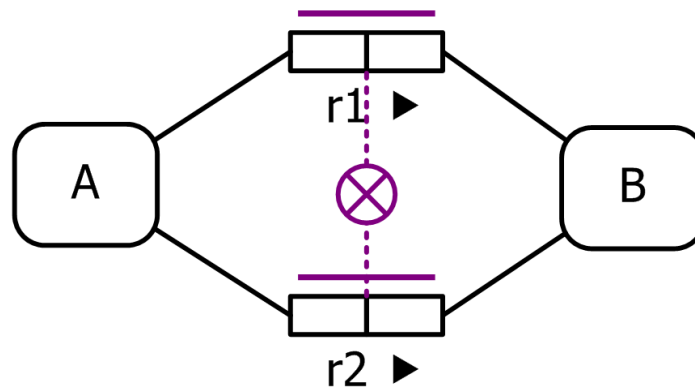


A skal ha en og bare en av rollene r1 og r2

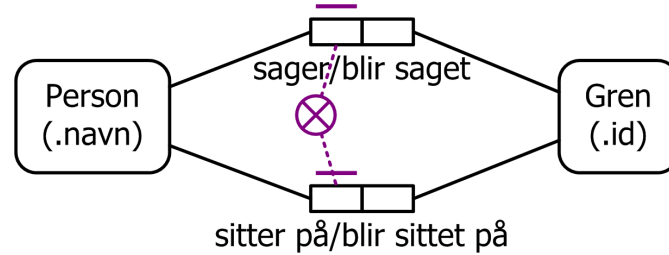


Mengdeulikhetsskranke over to roller

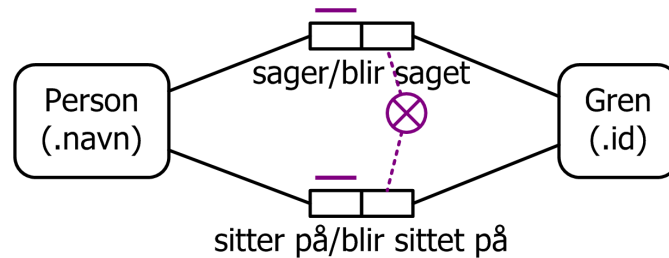
Det skal ikke være forekomster av A og B som er relatert gjennom begge faktatypene



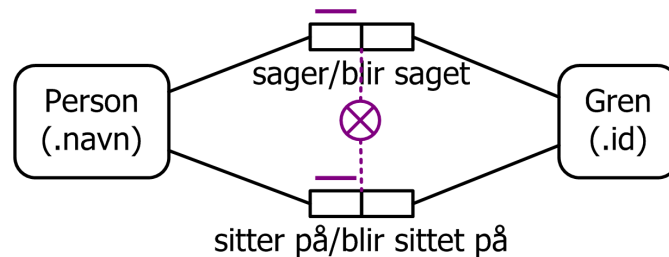
Mengdeulikhetsranken, eksempel



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

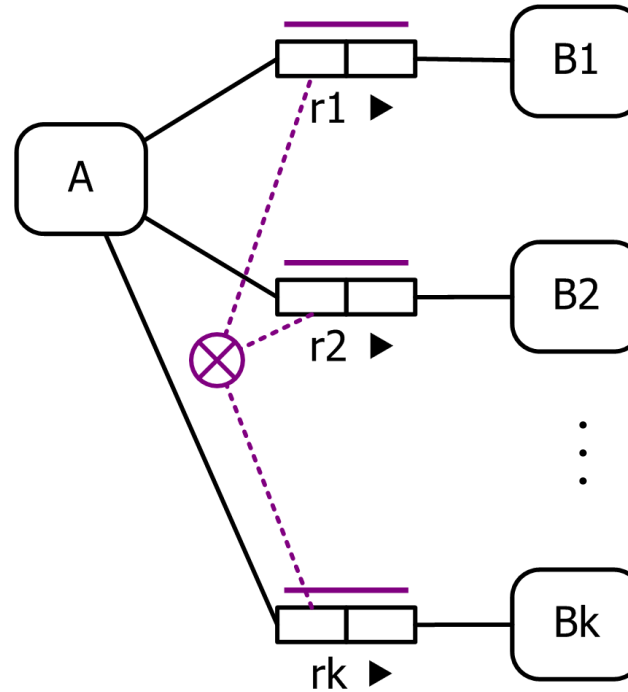


”Sag ikke av en gren som noen andre sitter på”



”Sag ikke av den grenen du selv sitter på”

Den generelle mengdeulikhetskranken

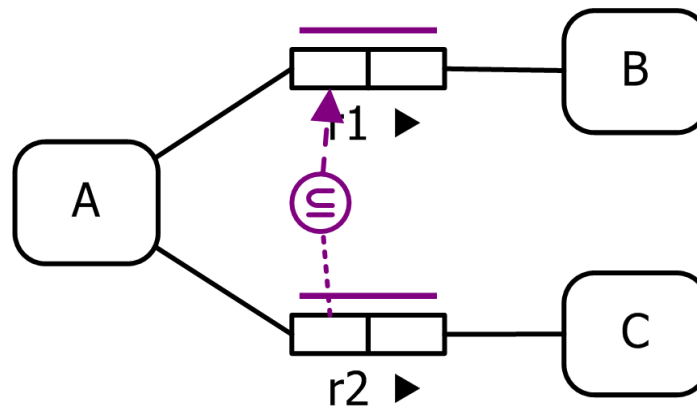


$$\text{pop}(r_i) \cap \text{pop}(r_j) = \emptyset \text{ for alle } i \neq j (1 \leq i \leq k, 1 \leq j \leq k)$$

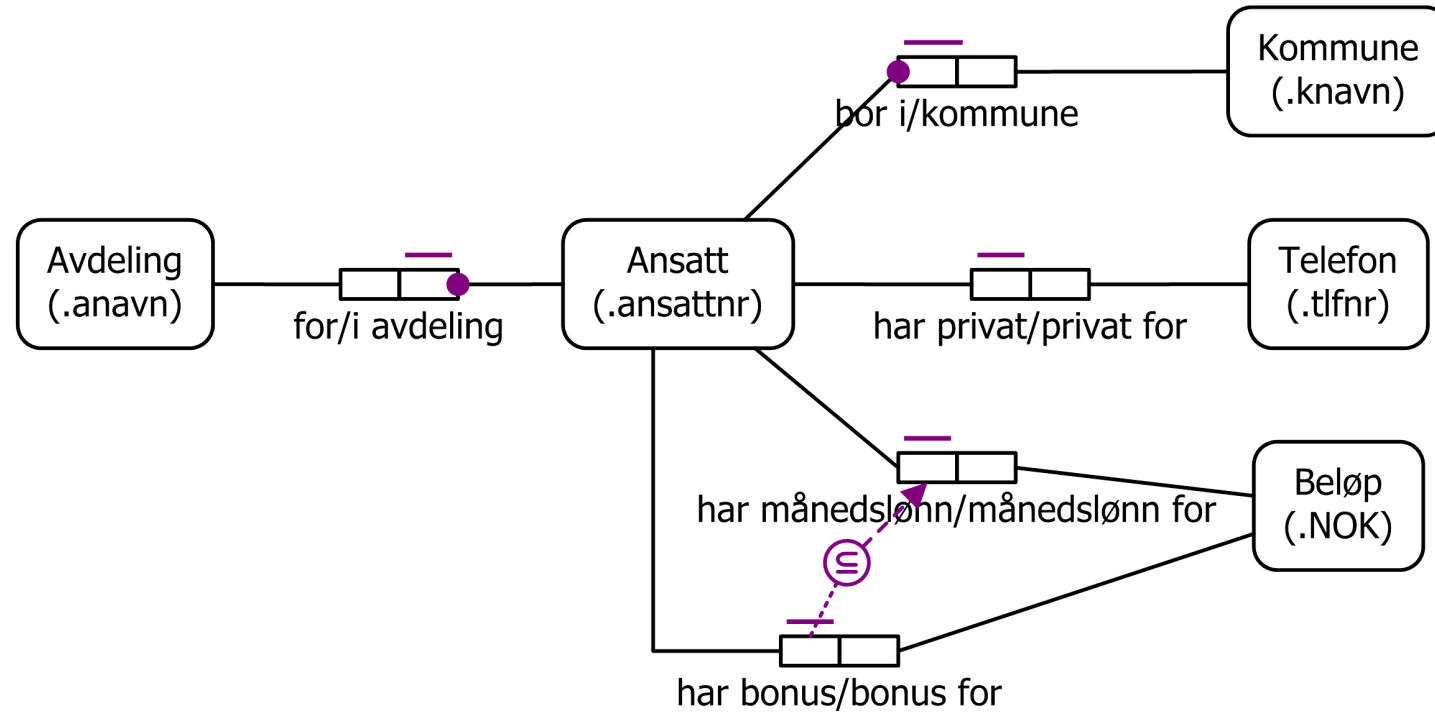
Delmengdeskranken

Hvis A har rollen r_2 ,
så skal A også ha rollen r_1 .

$\text{pop}(r_2) \subseteq \text{pop}(r_1)$ for alle tilstander

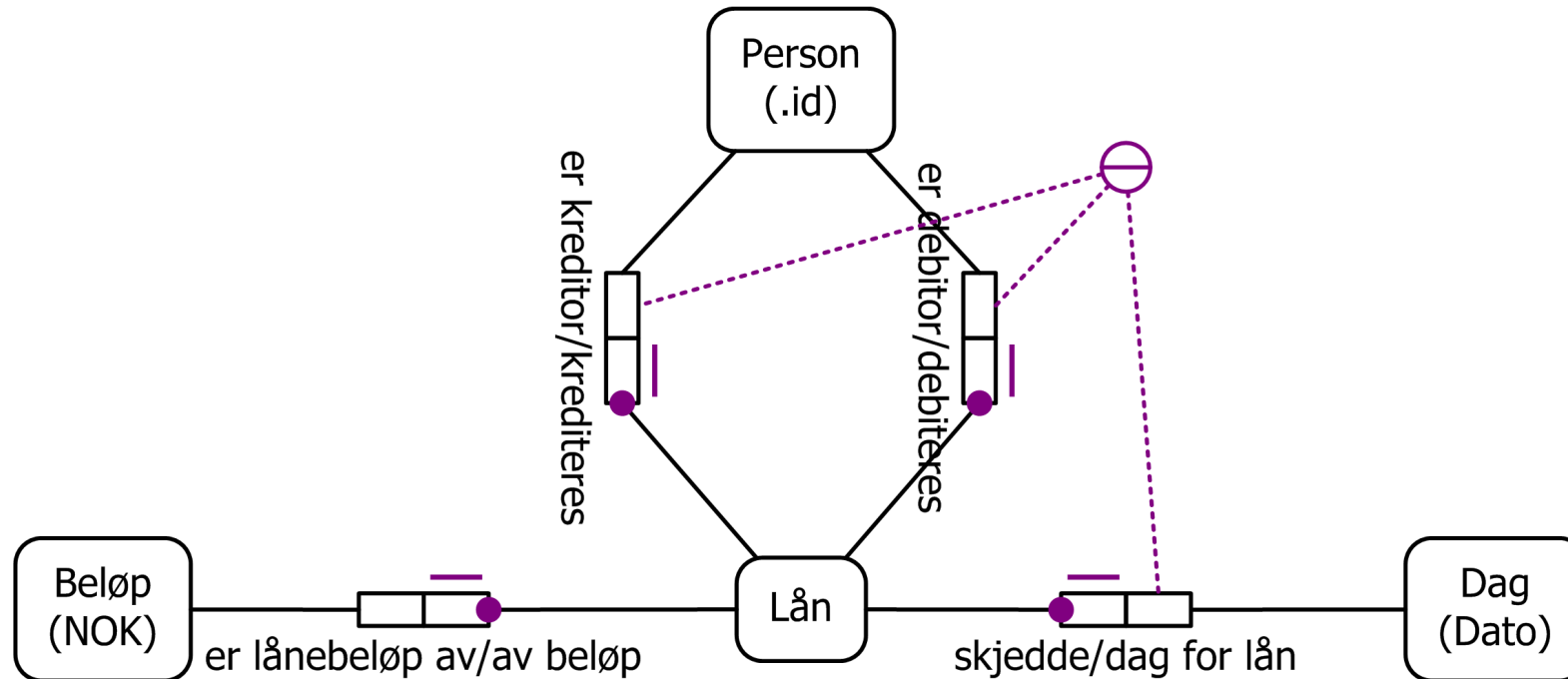


Delmengdeskranken, eksempel



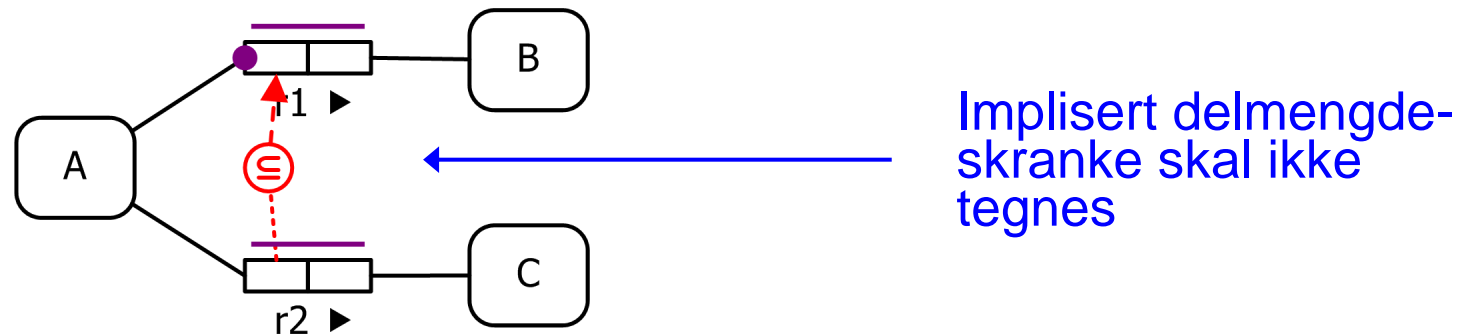
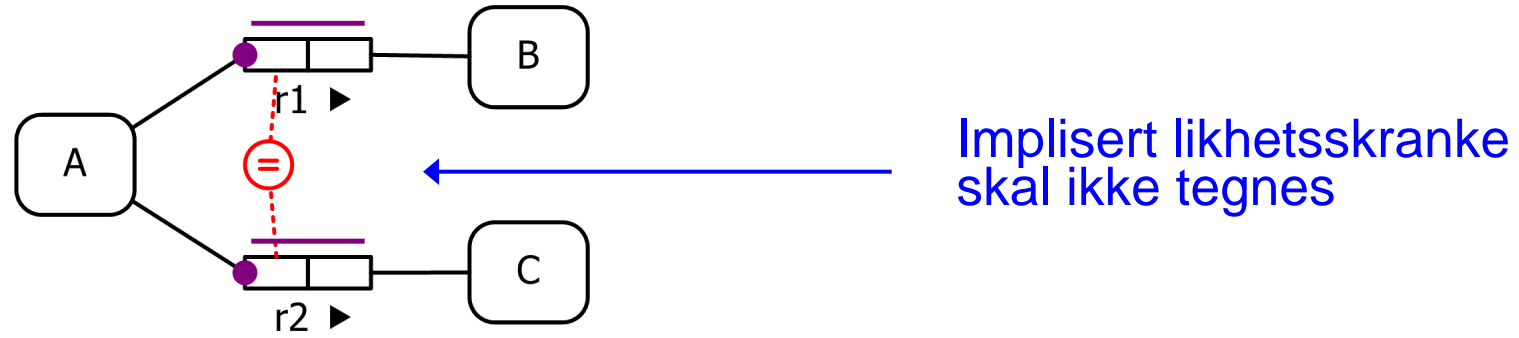
Eksempel med aritetet 4 – forts.

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

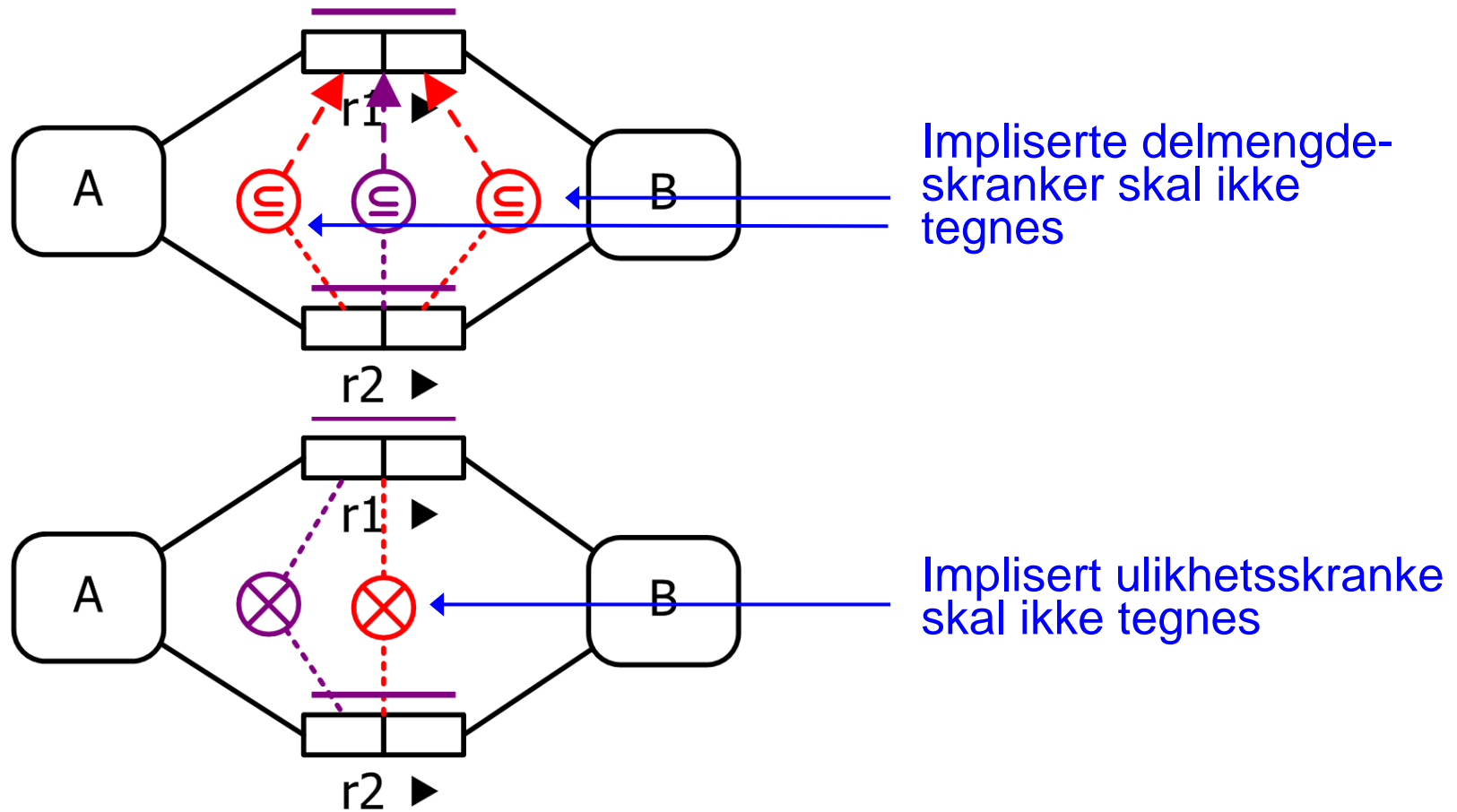


Oppgave: Hvilken skranke mangler?
Hvor skal den plasseres?

Impliserte skranke skal ikke tegnes



Impliserte skranke skal ikke tegnes



Mengdelikhetskranke over flere roller er vanligvis ikke lov

