

Præsentation

Anders

28/10/2020

We have the dataset ChestSim1000

```
data(chestSim1000, package="gRbase")  
head(chestSim1000)
```

```
##   asia tub smoke lung bronc either xray dysp  
## 1   no  no   no   no   yes     no   no  yes  
## 2   no  no   yes  no   yes     no   no  yes  
## 3   no  no   yes  no   no      no   no  no  
## 4   no  no   no   no   no      no   no  no  
## 5   no  no   yes  no   yes     no   no  yes  
## 6   no  no   yes  yes  yes     yes  yes  yes
```

This is a hypothetical Chest Clinic problem, by Lauritzen and Spiegelhalter. (ref til <https://arxiv.org/pdf/1301.7394.pdf>)

Here is a short explanation of the variables in the dataset.

- asia → subject has visited asia
- tub → subject has tuberculosis
- smoke → subject is a smoker
- lung → subject has lung cancer
- bronc → subject has bronchitis
- either → subject has either tuberculosis or lungcancer
- xray → subject has positive X-ray
- dysp → Subject has dyspnoea

Shortness-of-breath (dyspnoea) may be due to tuberculosis, lung cancer, bronchitis, none of them, or more than one of them. A recent visit to Asia increases the chances of tuberculosis, while smoking is known to be a risk factor for both lung cancer and bronchitis. The results of a single chest X-ray do not discriminate between lung cancer and tuberculosis, as does neither the presence nor absence of dyspnoea. (citat direkte sat ind fra <https://arxiv.org/pdf/1301.7394.pdf>)

```
dg1 <- dag(~ S + L|S + X|L:S + B|S + D|L:B)  
plot(dg1)
```

