Advanced Machine Learning

Lab 1

Rasmus Holm

2017-09-07

Assignment 1

```r
true_net <- empty.graph(names(learning.test))
modelstring(true_net) = "[A][C][F][B|A][D|A:C][E|B:F]"
```

```r
sapply(1:10, function(i) {
    set.seed(i)
    g1 <- hc(alarm, score="bde", iss=1, restart=10)

    set.seed(10 * i)
    g2 <- hc(alarm, score="bde", iss=1, restart=10)

    all.equal(cpdag(g1), cpdag(g2)) == TRUE
        ## all.equal(vstructs(g1, arcs=TRUE), vstructs(g2, arcs=TRUE)) == TRUE &&
        ## all.equal(vstructs(g1), vstructs(g2)) == TRUE
})
#>  [1]  TRUE  TRUE FALSE FALSE FALSE FALSE  TRUE FALSE FALSE FALSE
```
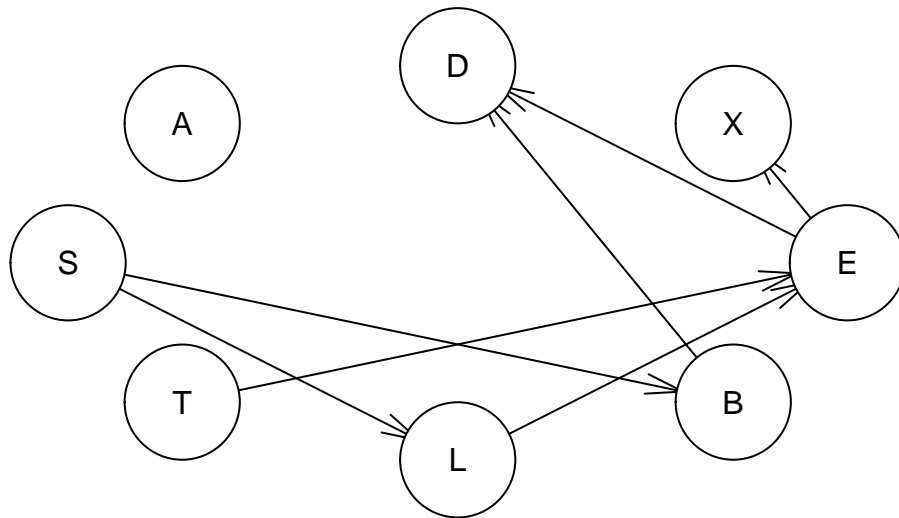
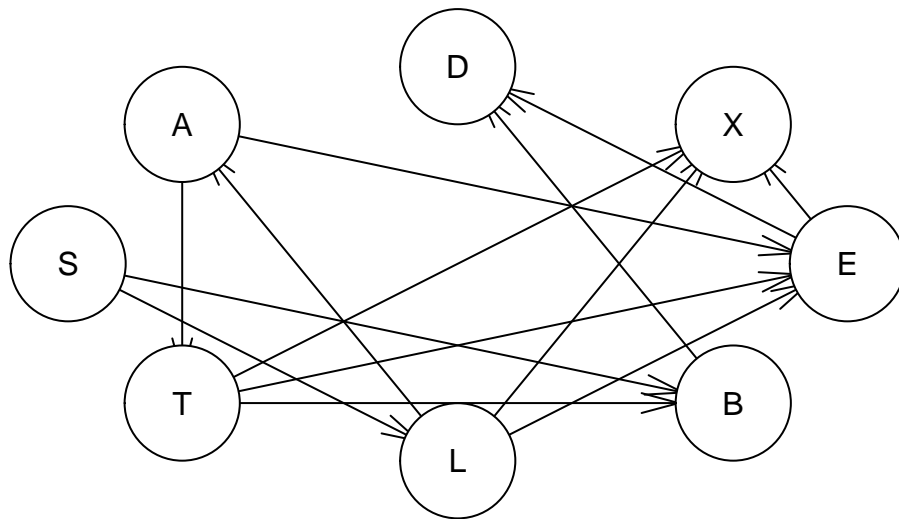## Assignment 2

```r
data <- asia

g1 <- hc(data, score="bde", iss=1, restart=10)
g10 <- hc(data, score="bde", iss=10, restart=10)
g100 <- hc(data, score="bde", iss=100, restart=10)
g1000 <- hc(data, score="bde", iss=1000, restart=10)

score(g1, data, score="bde")
#> [1] -11107.29
score(g10, data, score="bde")
#> [1] -11160.48
score(g100, data, score="bde")
#> [1] -11324.21
score(g1000, data, score="bde")
#> [1] -11672.71
```
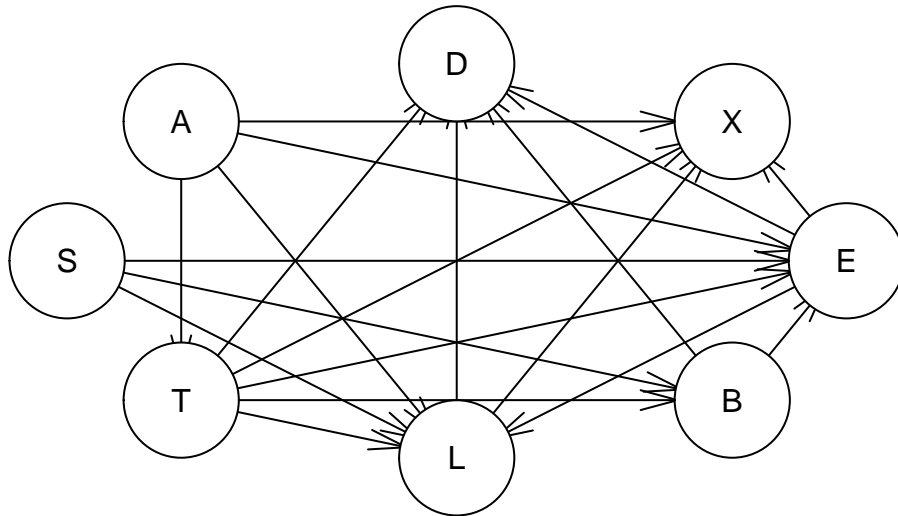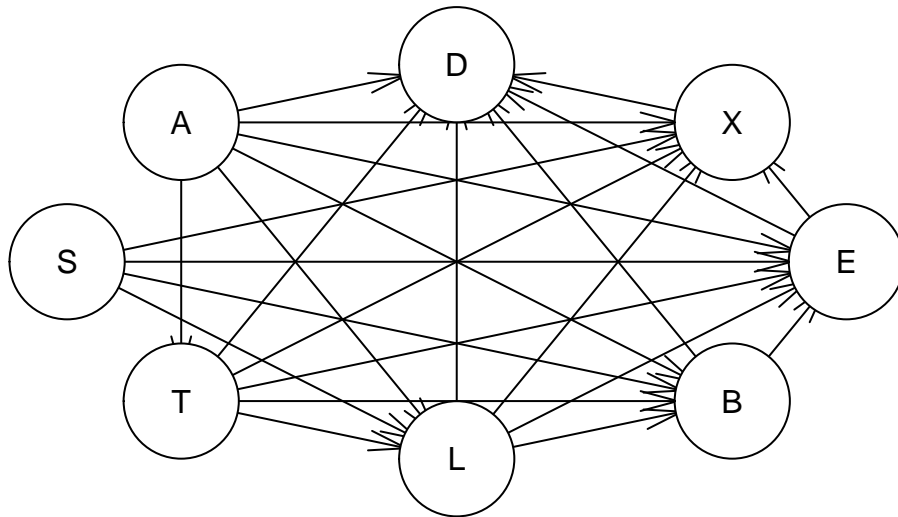
iss=1

iss=10

iss=100

iss=1000

In the BDe metric we assume that the prior distribution of nodes being connected is a Dirichlet. The imaginary sample size determine the parameters and the greater the number, the stronger we assume the prior and the data have less influence on the model we choose. That is to say that if the imaginary sample size is low the data determine the model which will most likely be more sparse, i.e. less arcs/edges, than having a stronger prior (uniform). Given that there are 8 nodes in the graph we would expect the average branching factor to be around $7/2 = 3.5$ and we get 3 with a imaginary sample size of 1000 which is close and Bayesian networks also have certain constraints that prevent edges from being added arbitrarly.

Assignment 3

Assignment 4a