Teste Melhor Impressão de Barcharts
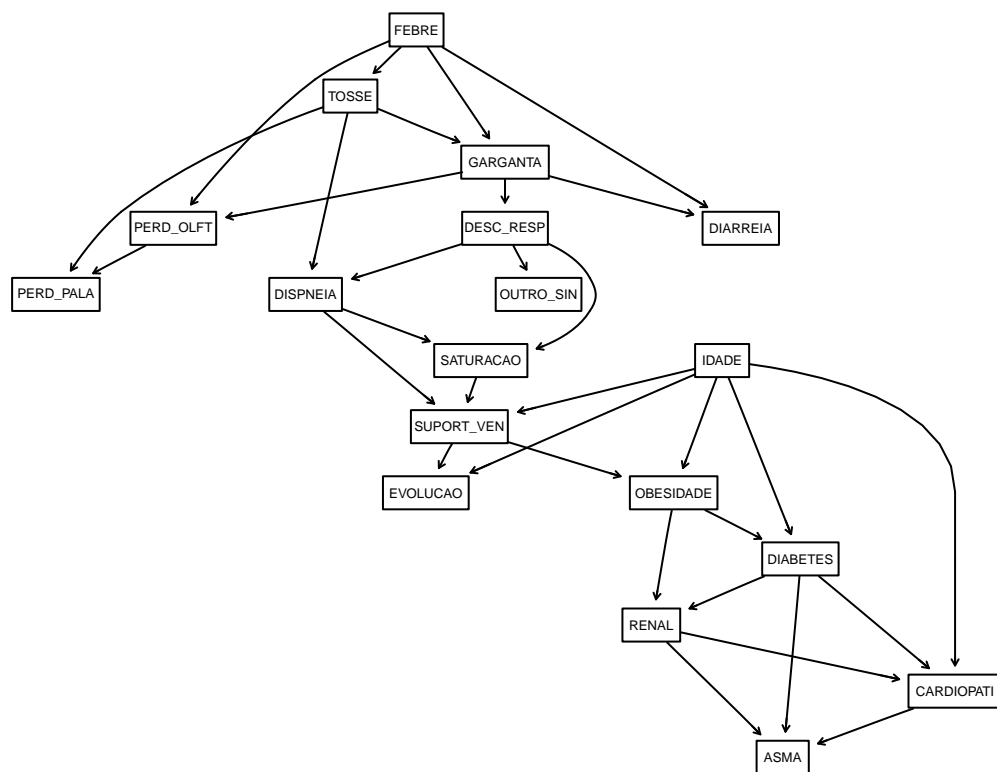
José Elvano Moraes
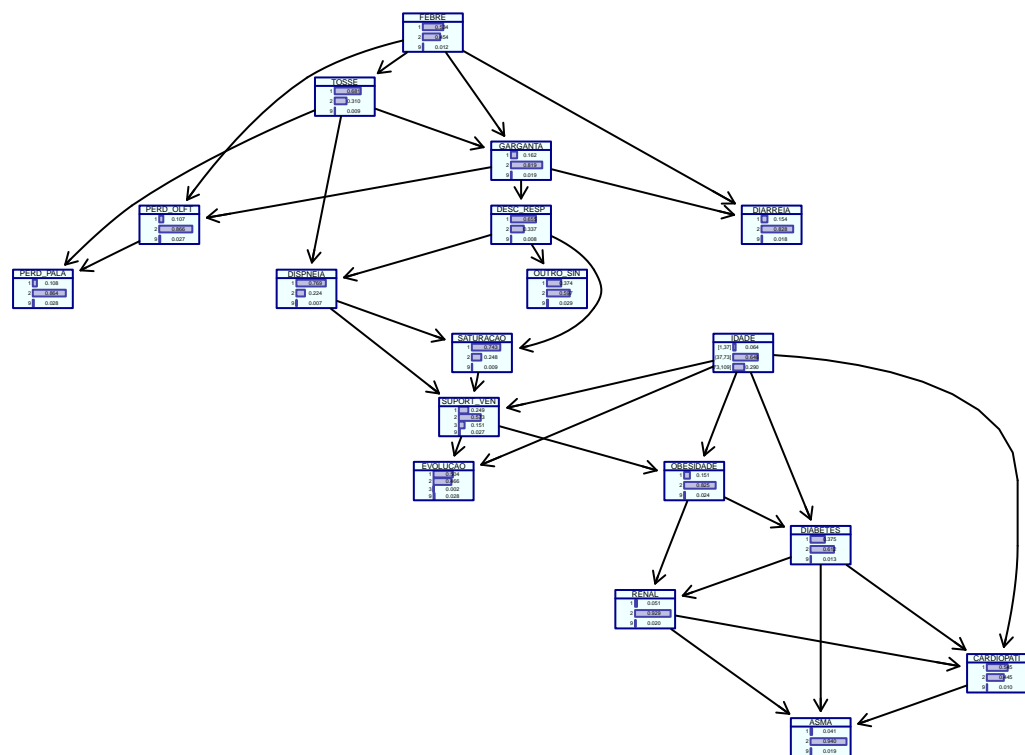
4/15/2021

```
## Rows: 76,344
## Columns: 19
## $ IDADE      <fct> "(73,109]", "(37,73]", "(37,73]", "(73,109]", "(37,73]", "(~
## $ FEBRE      <fct> 2, 2, 2, 2, 1, 2, 1, 1, 1, 1, 2, 1, 2, 1, 2, 2, 1, 1, 1, 2,~
## $ TOSSE      <fct> 1, 2, 2, 2, 2, 2, 2, 1, 2, 1, 1, 2, 1, 1, 2, 2, 1, 2, 2, 1,~
## $ GARGANTA   <fct> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 1, 2, 2, 2, 1, 2, 2, 2,~
## $ DISPNEIA   <fct> 1, 1, 2, 2, 1, 2, 2, 2, 2, 1, 1, 1, 1, 2, 1, 2, 1, 2, 1, 1,~
## $ DESC_RESP  <fct> 1, 1, 2, 2, 2, 2, 2, 2, 1, 1, 1, 2, 1, 1, 2, 2, 1, 2, 1, 1,~
## $ SATURACAO  <fct> 1, 1, 1, 2, 2, 2, 2, 2, 1, 2, 2, 1, 1, 1, 1, 2, 2, 2, 2, 1,~
## $ DIARREIA   <fct> 2, 2, 1, 2, 2, 2, 2, 1, 2, 2, 2, 2, 1, 1, 2, 2, 1, 1, 2, 2,~
## $ OUTRO_SIN  <fct> 2, 2, 1, 1, 2, 1, 2, 1, 2, 2, 2, 1, 2, 1, 1, 1, 2, 1, 2, 2,~
## $ EVOLUCAO   <fct> 1, 2, 1, 1, 1, 2, 1, 2, 2, 1, 2, 1, 2, 2, 2, 1, 1, 1, 2, 1,~
## $ RENAL      <fct> 2, 2, 2, 2, 1, 2, 1, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2,~
## $ DIABETES   <fct> 1, 2, 1, 2, 1, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2,~
## $ OBESIDADE  <fct> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 1, 2, 2, 2, 2, 2,~
## $ PERD_OLFT  <fct> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2,~
## $ PERD_PALA  <fct> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,~
## $ CARDIOPATI <fct> 2, 2, 2, 2, 1, 1, 2, 2, 1, 1, 1, 1, 1, 1, 1, 2, 2, 1, 2, 1,~
## $ SUPORT_VEN <fct> 2, 2, 2, 2, 9, 3, 3, 2, 2, 3, 2, 2, 3, 1, 1, 3, 3, 2, 1, 3,~
## $ HEMATOLOGI <fct> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,~
## $ ASMA       <fct> 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,~

## White list:

##        from        to
## [1,] "OBESIDADE" "DIABETES"
## [2,] "IDADE"     "DIABETES"
## [3,] "IDADE"     "CARDIOPATI"
## [4,] "IDADE"     "SUPORT_VEN"
## [5,] "DIABETES"  "RENAL"
## [6,] "DIABETES"  "CARDIOPATI"
## [7,] "RENAL"     "CARDIOPATI"
```

...

FEBRE

TOSSE

GARGANTA

PERD_OLFT

DIARREIA

PERD_PALA

DESC_RESP

DISPNEIA

OUTRO_SIN

SATURACAO

IDADE

SUPORT_VEN

EVOLUCAO

OBESIDADE

DIABETES

RENAL

CARDIOPATI

ASMA

Rede de probabilidasdes

FEBRE

TOSSE

GARGANTA

PERD_OLFT

DIARREIA

PERD_PALA

DESC_RESP

DISPNEIA

OUTRO_SIN

SATURACAO

IDADE

SUPORT_VEN

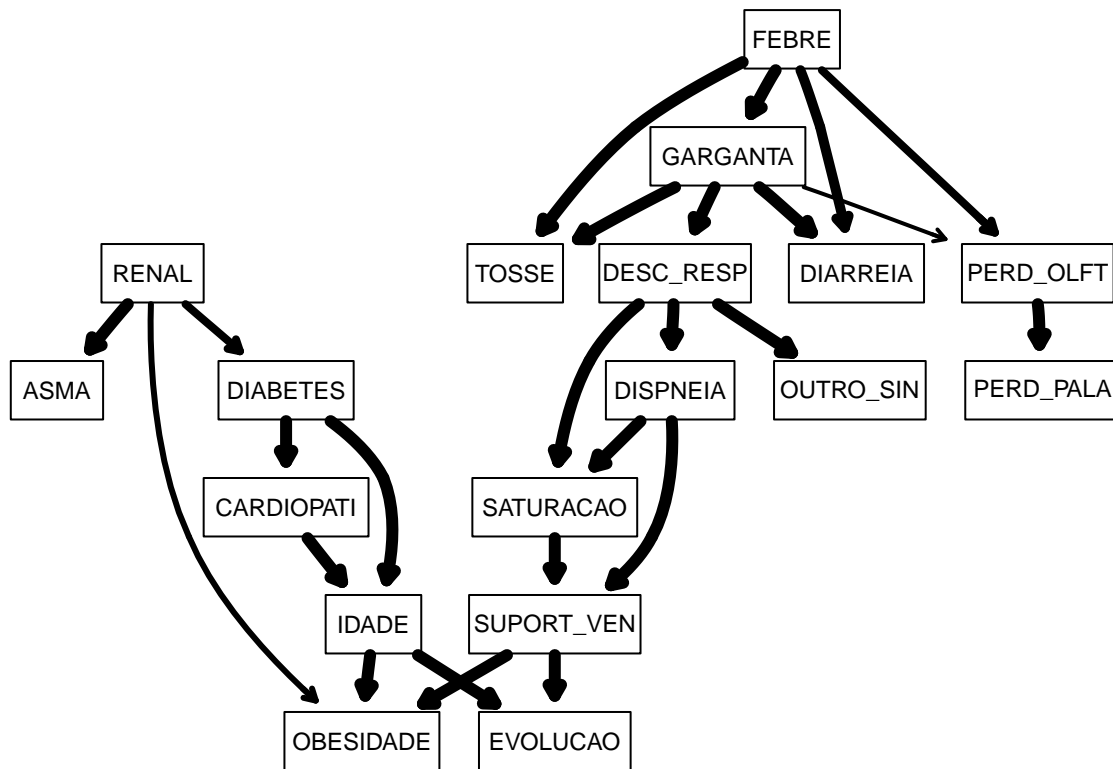EVOLUCAO

OBESIDADE

DIABETES

RENAL

CARDIOPATI

ASMA

Bootstraping

```
boots.trap  <- 100
str.diff = suppressMessages(boot.strength(s1, R = boots.trap, algorithm = "mmhc"))
cat(paste('Threshold: ', attr(str.diff, "threshold")))
```

```
## Threshold:  0.5
```

```
avg.diff = averaged.network(str.diff)
strength.plot(avg.diff, str.diff, shape = "rectangle", main = paste("Iterações = ", boots.trap))
```


How can we compare the averaged network (avg.diff) with the network we originally learned in from all the data? The most qualitative way is to plot the two networks side by side, with the nodes in the same positions, and highlight the arcs that appear in one network and not in the other, or that appear with different directions.

```
par(mfrow = c(1, 2))
graphviz.compare(avg.diff, bn1, shape = "rectangle", main = c("DAG médio", "DAG único"))
```
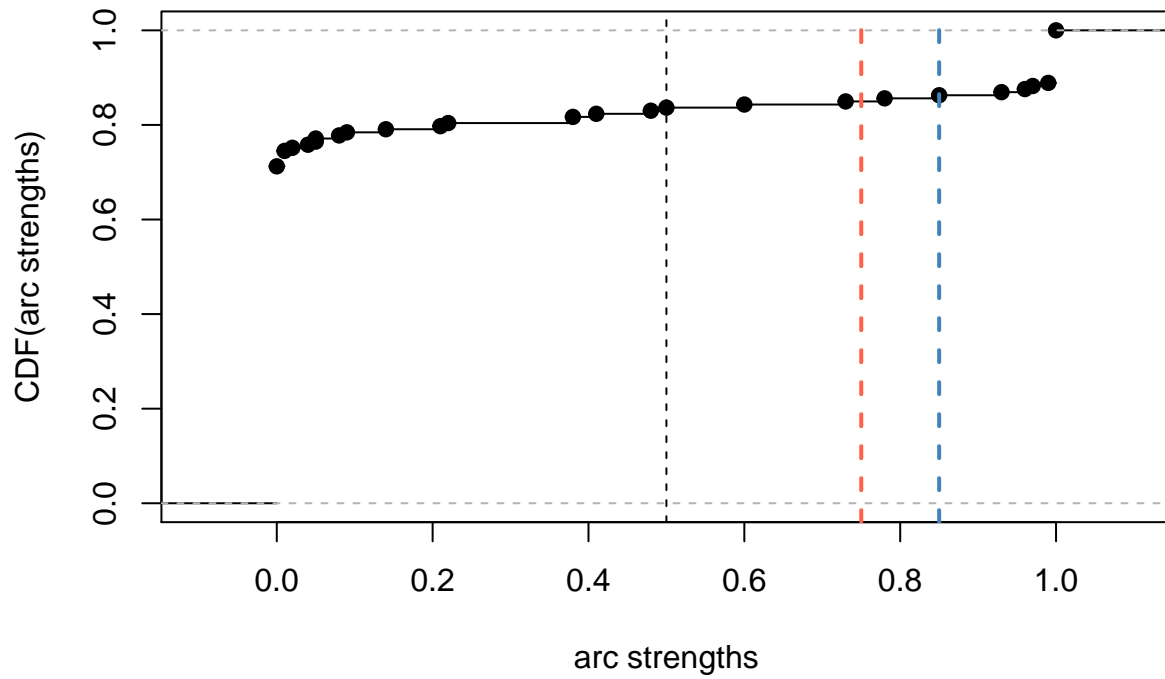
DAG médio
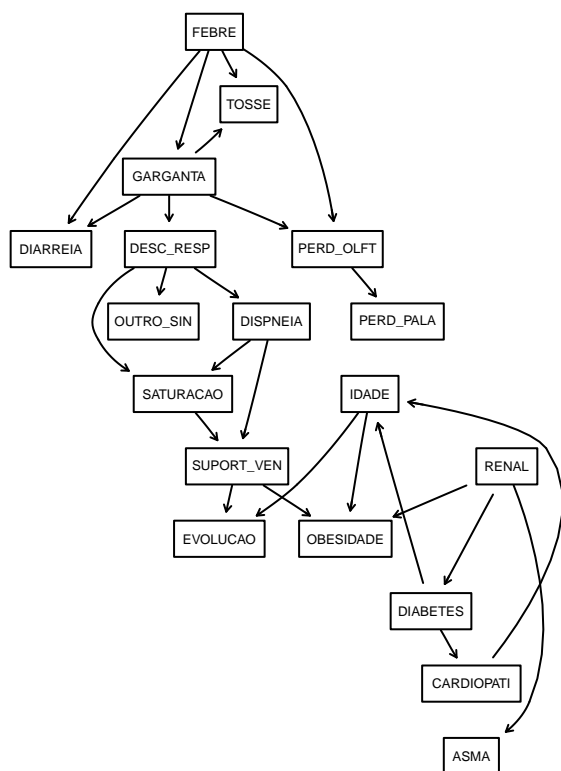
DAG único
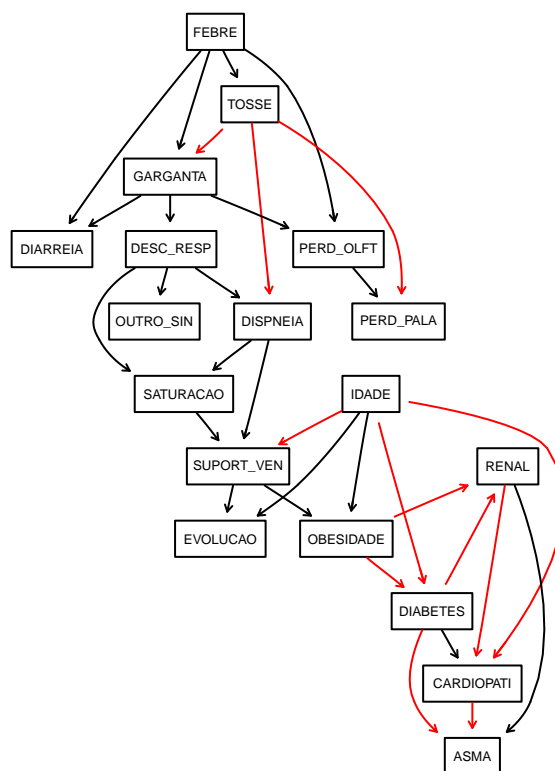
It is also a good idea to look at the threshold with respect to the distribution of the arc strengths

```r
plot(str.diff)
abline(v = 0.75, col = "tomato", lty = 2, lwd = 2)
abline(v = 0.85, col = "steelblue", lty = 2, lwd = 2)
```

threshold = 0.5

The simpler network we obtain by setting threshold = 0.8 in averaged.network() is shown below; it is certainly easier to reason with from a qualitative point of view.

```
par(mfrow = c(1, 2))
avg.simpler = averaged.network(str.diff, threshold = 0.75)
strength.plot(avg.diff, str.diff, shape = "rectangle", main = paste("Iterações = ", boots.trap, " Thr =
strength.plot(avg.simpler, str.diff, shape = "rectangle", main = 'Iterações = 200. Simplificado. Thr = (
```

Iterações = 100 Thr = 0.5   Iterações = 200. Simplificado. Thr = 0.7