Inferencia. Fase I

José Elvano Moraes

17/04/2021

Redes Bayesianas

Probabilistic reasoning on BNs works in the framework of Bayesian statistics and focuses on the computation of posterior probabilities or densities. For example, suppose we have learned a BN B with DAG G and parameters Θ. We want to use B to investigate the effects of a new piece of evidence E using the knowledge encoded in B, that is, to investigate the posterior distribution

$$P(\mathbf{X}|\mathbf{E}, \mathcal{B}) = P(\mathbf{X}|\mathbf{E}, \mathbf{G}, \Theta)$$

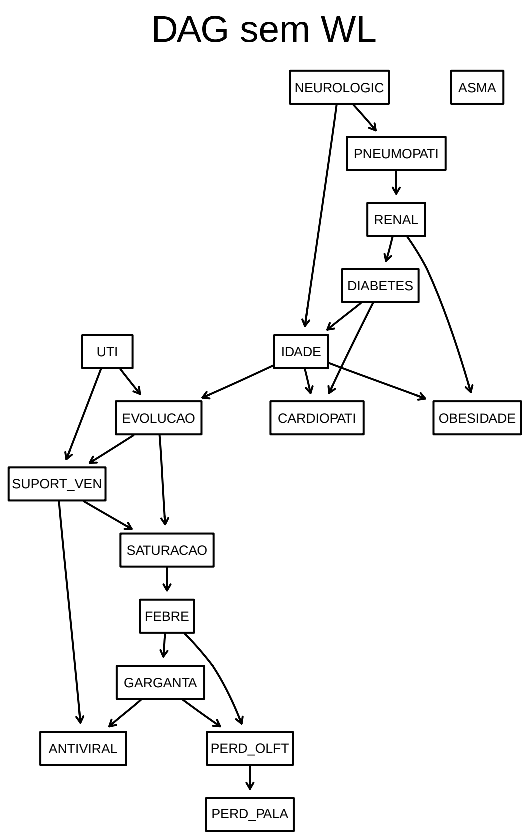
Questions that can be asked are called queries and are typically an event of interest. The two most common queries are conditional probability (CPQ) and maximum a posteriori (MAP) queries, also known as most probable explanation (MPE) queries

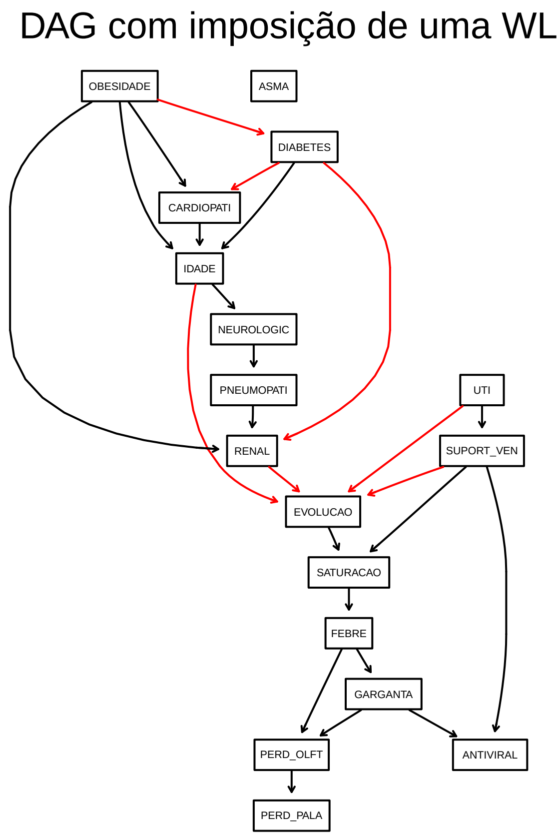
some content here

```r
glimpse(ddf)
```

```
## Rows: 76,666
## Columns: 17
## $ IDADE      <fct> "(37,73]", "(37,73]", "(73,109]", "(37,73]", "(73,109]", "(~
## $ FEBRE      <fct> 2, 2, 2, 1, 2, 1, 1, 1, 1, 2, 1, 2, 1, 2, 2, 1, 1, 2, 2, 2,~
## $ GARGANTA   <fct> 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 1, 2, 2, 2, 1, 2, 2, 2, 2,~
## $ SATURACAO  <fct> 1, 1, 2, 2, 2, 2, 2, 1, 2, 2, 1, 1, 1, 1, 2, 2, 2, 1, 2, 2,~
## $ EVOLUCAO   <fct> 2, 1, 1, 1, 2, 1, 2, 2, 1, 2, 1, 2, 2, 2, 1, 1, 1, 1, 1, 1,~
## $ RENAL      <fct> 2, 2, 2, 1, 2, 1, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,~
## $ DIABETES   <fct> 2, 1, 2, 1, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2,~
## $ OBESIDADE  <fct> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 1, 2, 2, 9, 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,~
## $ NEUROLOGIC <fct> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 9, 2, 2, 2, 2,~
## $ PNEUMOPATI <fct> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,~
## $ UTI        <fct> 1, 1, 2, 1, 2, 2, 1, 2, 2, 2, 2, 2, 1, 1, 2, 2, 1, 2, 2, 2,~
## $ CARDIOPATI <fct> 2, 2, 2, 1, 1, 2, 2, 1, 1, 1, 1, 1, 1, 1, 2, 2, 1, 1, 1, 1,~
## $ SUPORT_VEN <fct> 2, 2, 2, 9, 3, 3, 2, 2, 3, 2, 2, 3, 1, 1, 3, 3, 2, 3, 9, 3,~
## $ ASMA       <fct> 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,~
## $ ANTIVIRAL  <fct> 2, 2, 2, 9, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,~
```

some content here

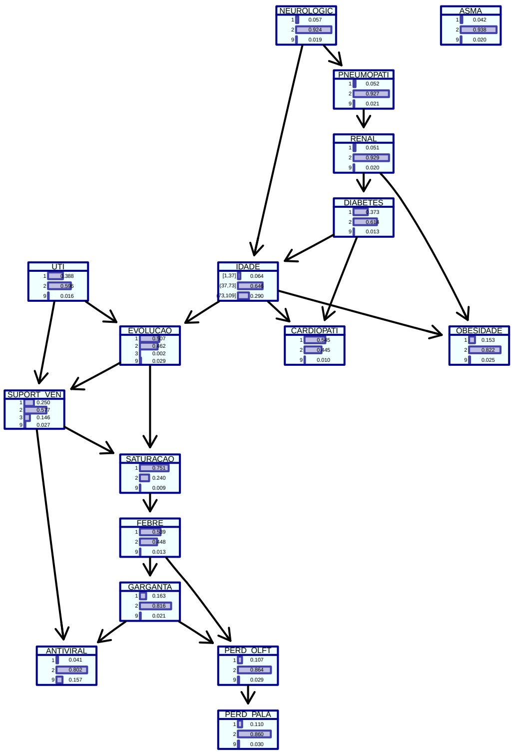




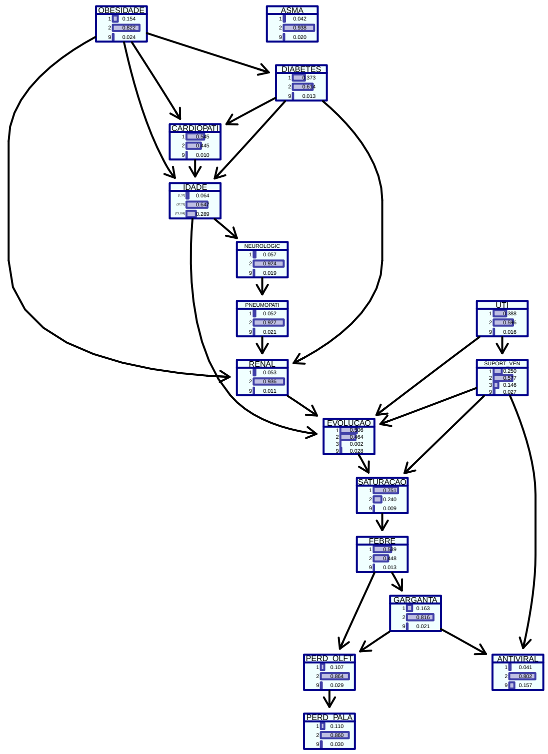
some content here

```
## Warning in from.bn.fit.to.grain(x): NaN conditional probabilities in EVOLUCAO,
## replaced with a uniform distribution.
```

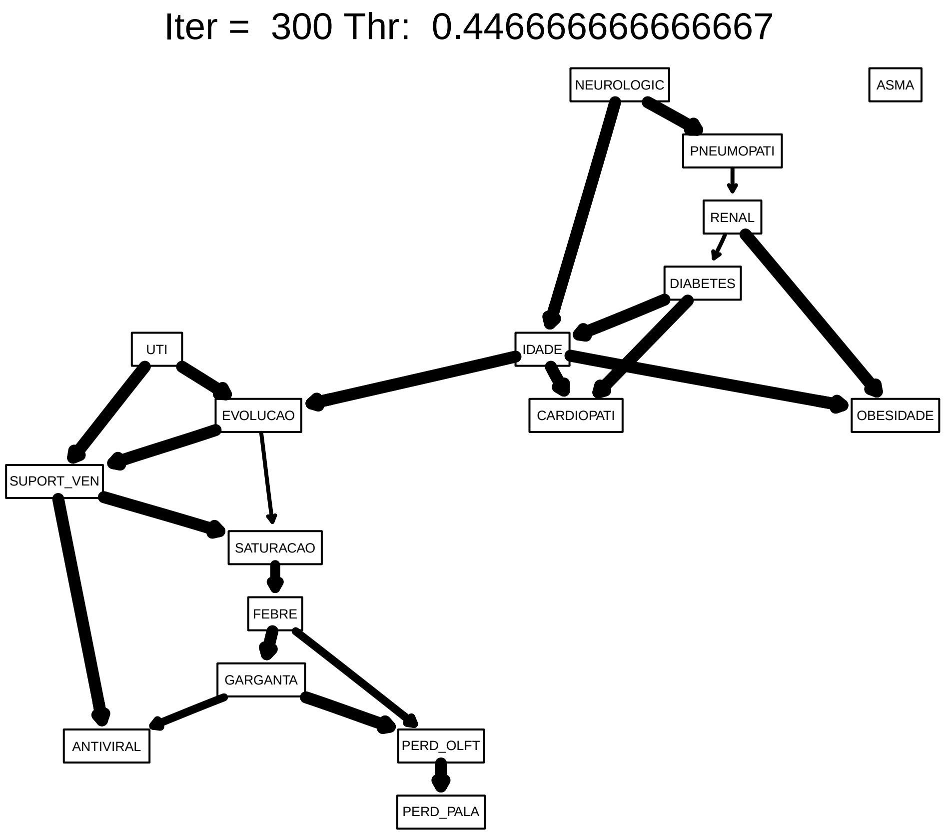
DAG sem WL



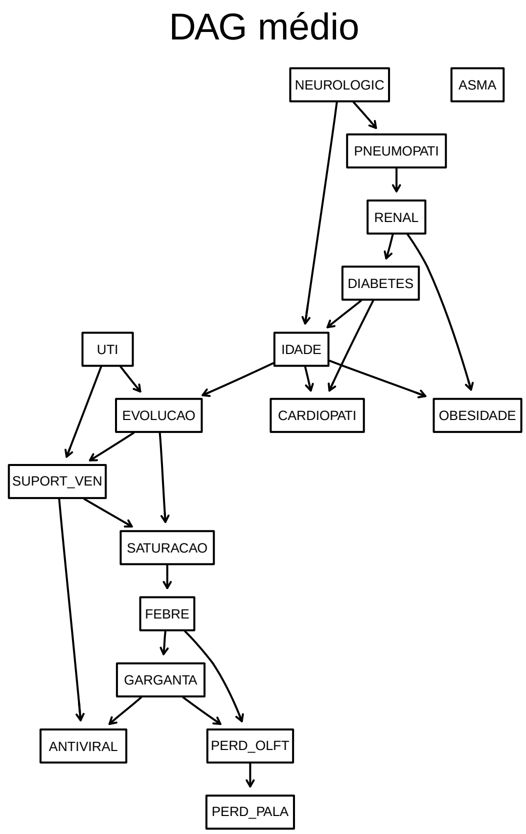
DAG com WL

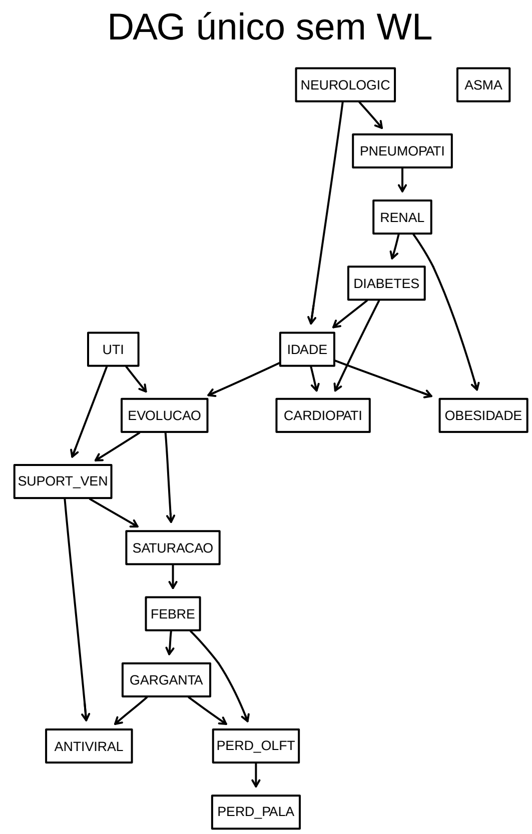


some content here

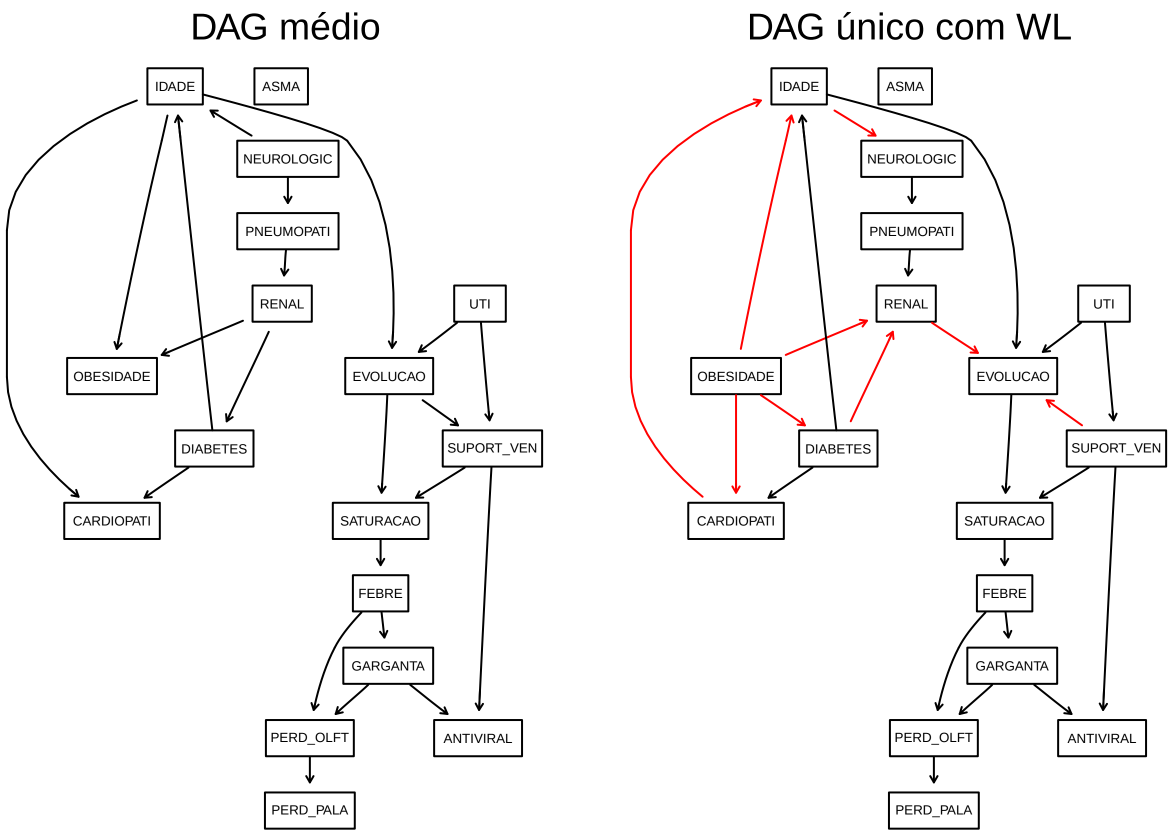


some content here





some content here



DAG mínimo obtido

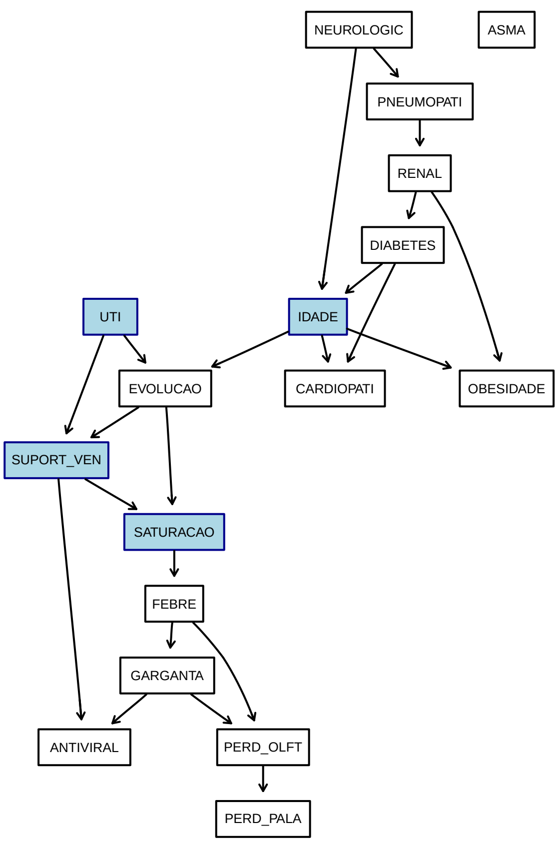
```
avg.diff
```

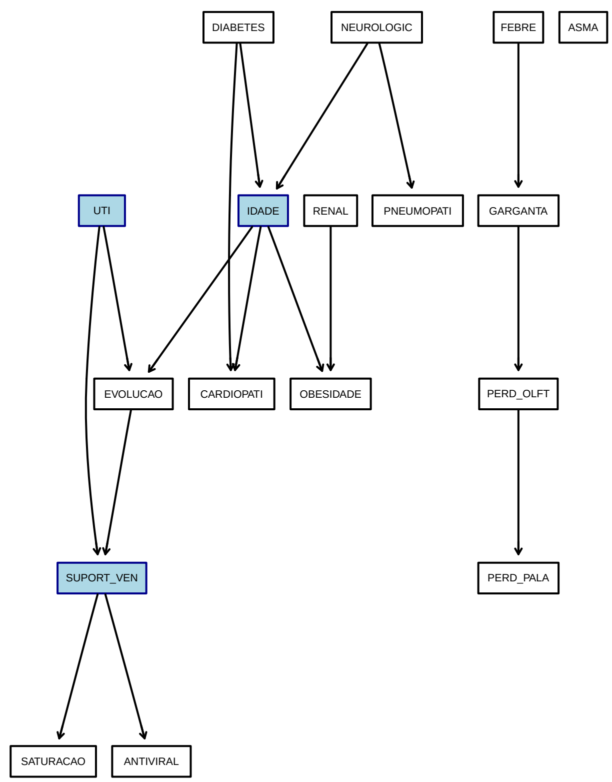
```
##
##   Random/Generated Bayesian network
##
##   model:
##    [NEUROLOGIC][UTI][ASMA][PNEUMOPATI|NEUROLOGIC][RENAL|PNEUMOPATI]
##    [DIABETES|RENAL][IDADE|DIABETES:NEUROLOGIC][EVOLUCAO|IDADE:UTI]
##    [OBESIDADE|IDADE:RENAL][CARDIOPATI|IDADE:DIABETES][SUPORT_VEN|EVOLUCAO:UTI]
##    [SATURACAO|EVOLUCAO:SUPORT_VEN][FEBRE|SATURACAO][GARGANTA|FEBRE]
##    [PERD_OLFT|FEBRE:GARGANTA][ANTIVIRAL|GARGANTA:SUPORT_VEN]
##    [PERD_PALA|PERD_OLFT]
##   nodes:                                 17
##   arcs:                                  22
##     undirected arcs:                     0
##     directed arcs:                       22
##   average markov blanket size:           3.06
##   average neighbourhood size:            2.59
##   average branching factor:              1.29
##
##   generation algorithm:                  Model Averaging
```

```
##   significance threshold:              0.45
```

```
avg.simpler
```

```
##
##   Random/Generated Bayesian network
##
##   model:
##    [FEBRE][RENAL][DIABETES][NEUROLOGIC][UTI][ASMA][IDADE|DIABETES:NEUROLOGIC]
##    [GARGANTA|FEBRE][PNEUMOPATI|NEUROLOGIC][EVOLUCAO|IDADE:UTI]
##    [OBESIDADE|IDADE:RENAL][PERD_OLFT|GARGANTA][CARDIOPATI|IDADE:DIABETES]
##    [PERD_PALA|PERD_OLFT][SUPORT_VEN|EVOLUCAO:UTI][SATURACAO|SUPORT_VEN]
##    [ANTIVIRAL|SUPORT_VEN]
##   nodes:                               17
##   arcs:                                16
##     undirected arcs:                   0
##     directed arcs:                     16
##   average markov blanket size:         2.24
##   average neighbourhood size:          1.88
##   average branching factor:            0.94
##
##   generation algorithm:                Model Averaging
##   significance threshold:              0.95
```





some content here