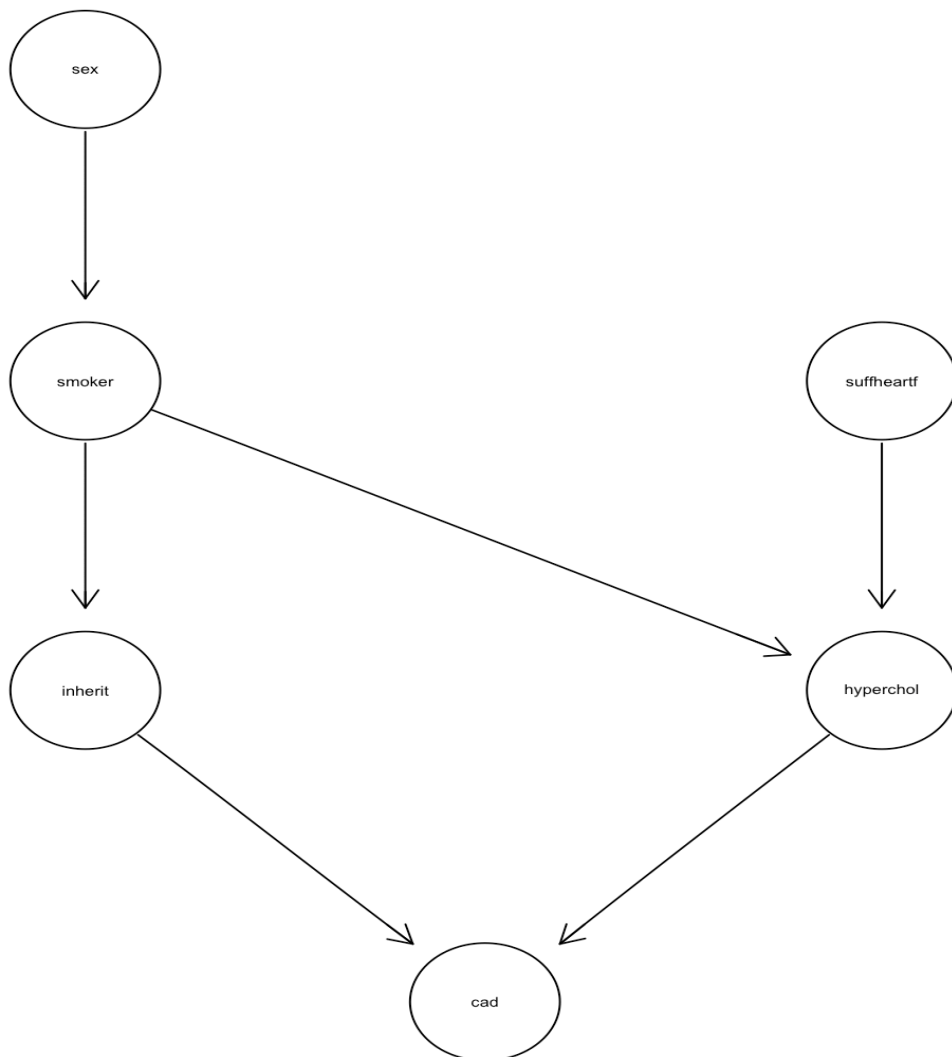# EAS 507

HW -3

By - Mohit Tripathi

Ans -1

a.)

Conditional Probability Tables:

```
> fullandsex
Sex
Female    Male
    47     189
> smokerandsex
       Sex
Smoker Female Male
   No        17   34
   Yes       30  155
> suffheartfandobj
SuffHeartF
 No Yes
167  69
> inheritandsmoker
       Smoker
Inherit  No Yes
    No   42 120
    Yes   9  65
> cadandhypercholandinherit
, , Hyperchol = No

      Inherit
CAD    No Yes
  No  69  12
  Yes 15  12


, , Hyperchol = Yes

      Inherit
CAD    No Yes
  No  35  13
  Yes 43  37
```

```
> hypercholandsuffheartfandsmoker
, , SuffHeartF = No

           Smoker
Hyperchol  No  Yes
      No   27   59
      Yes  13   68

, , SuffHeartF = Yes

           Smoker
Hyperchol  No  Yes
      No    3   19
      Yes   8   39
```

D- separation in the graph :

- Sex and suffheartf

- Smoker and suffheartf

- Inherit and suffheartf

- Sex and hyperchol

b.)

New observation is female with Hypercholesterolemia (high cholesterol)

Change in Conditional Prob :-

```
          CAD
SuffHeartF         No        Yes
       No  0.3957368 0.3118903
       Yes 0.1443930 0.1479799
attr(,"class")
[1] "parray" "array"
> querygrain(gcomp.ev, nodes = c("SuffHeartF", "CAD"), type = "joint")
          CAD
SuffHeartF         No        Yes
       No  0.4078210 0.2998061
       Yes 0.1453059 0.1470670
attr(,"class")
[1] "parray" "array"
>
>
> querygrain(gcomp, nodes = c("SuffHeartF", "CAD"), type = "conditional")
     SuffHeartF
CAD          No        Yes
  No  0.7326698 0.2673302
  Yes 0.6782138 0.3217862
> querygrain(gcomp.ev, nodes = c("SuffHeartF", "CAD"), type = "conditional
     SuffHeartF
CAD          No        Yes
  No  0.7373010 0.2626990
  Yes 0.6708976 0.3291024
>
>
> querygrain(gcomp, nodes = c("SuffHeartF", "CAD"), type = "marginal")
$SuffHeartF
SuffHeartF
       No        Yes
0.7076271 0.2923729

$CAD
CAD
       No        Yes
0.5401298 0.4598702
```

```
> querygrain(gcomp.ev, nodes = c("SuffHeartF", "CAD"), type = "marginal")
$SuffHeartF
SuffHeartF
       No       Yes
0.7076271 0.2923729

$CAD
CAD
       No       Yes
0.5531269 0.4468731
```

on observing marginal prob., Chances of heart failure remains the same whereas CAD prob. Has been decreased slightly.
We can observe change in joint and conditional prob. As well as stated in above figures.

c.)
   New data set with 5 observations conditional upon this new information :

```
     Sex Smoker SuffHeartF Inherit CAD Hyperchol
1 Female    No        Yes      No  No        No
2 Female    No         No     Yes  No        No
3 Female   Yes         No     Yes  No       Yes
4 Female   Yes         No      No  No        No
5 Female   Yes         No     Yes  No       Yes
>
```

```
$pred$Smoker
[1] "Yes" "Yes" "Yes" "No"  "Yes"

$pred$CAD
[1] "No"  "No"  "Yes" "No"  "Yes"


$pEvidence
[1] 0.01262758 0.02075442 0.01984580 0.05544145 0.01984580
```

The prediction for smoker and CAD is shown in above figure

d.)

Misclassification rate for smoker is 33.6

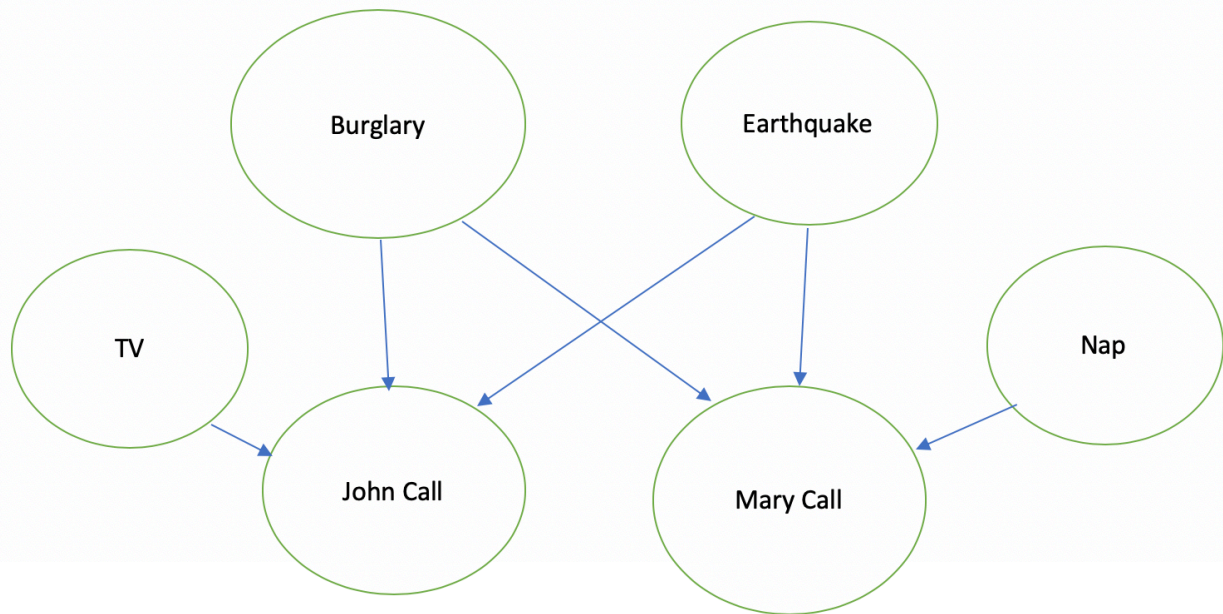Misclassification rate for CAD is 32.4

Network performance is not good as the misclassification is high for smoker and CAD.

To improve it, we can consider more permutations of the network which may increase accuracy.
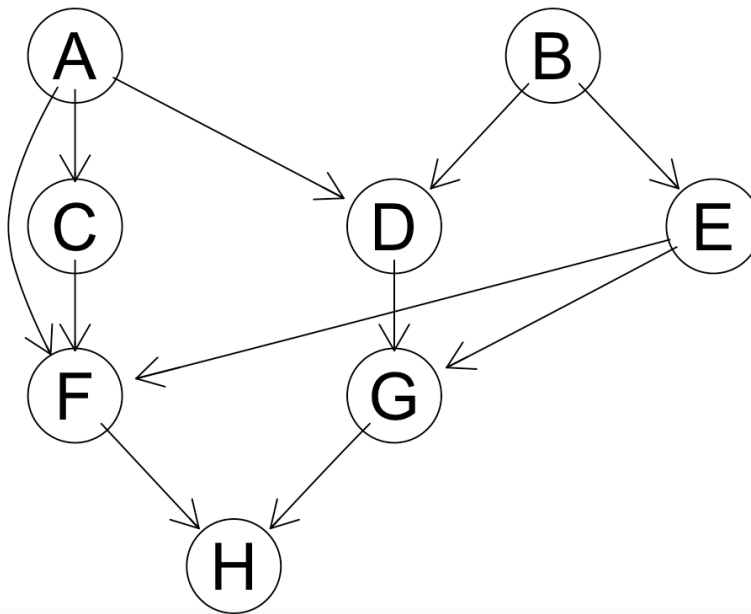Domain knowledge can also be added along with the network to contribute towards accuracy.

Ans 2

Without Alarm(I-Map): -

Ans 3 DAG plot for the diagram in the problem statement: -

A) C and G are d-separated.

Ans)False

B) C and E are d-separated.

Ans)True

**C) C and E are d-connected given
evidence about G.**

Ans)True

D) A and G are d-connected given evidence about D and E.

Ans)False

E) A and G are d-connected given evidence on D.

Ans)True