Elisabetta Fersini

DISCo

Università degli Studi di Milano-

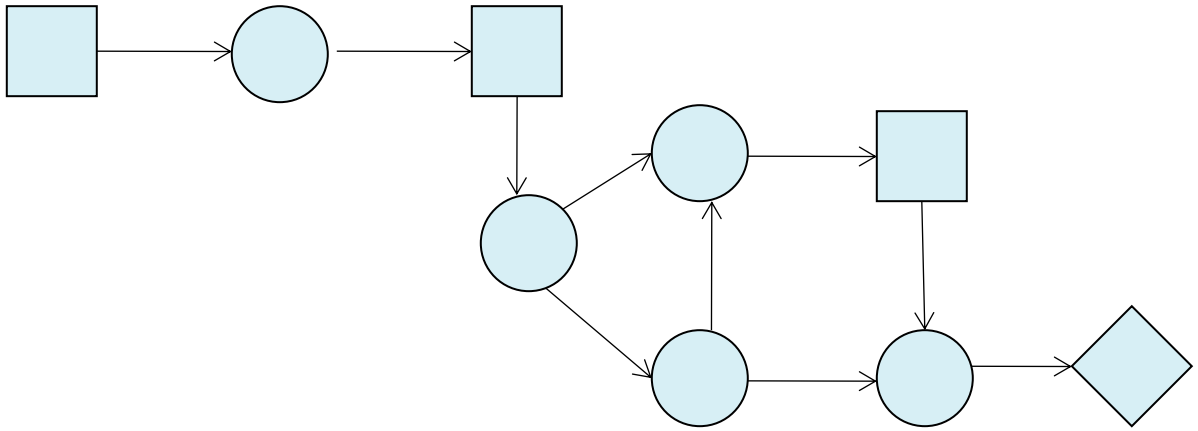
Bicocca

Viale Sarca, 336

20126 Milano

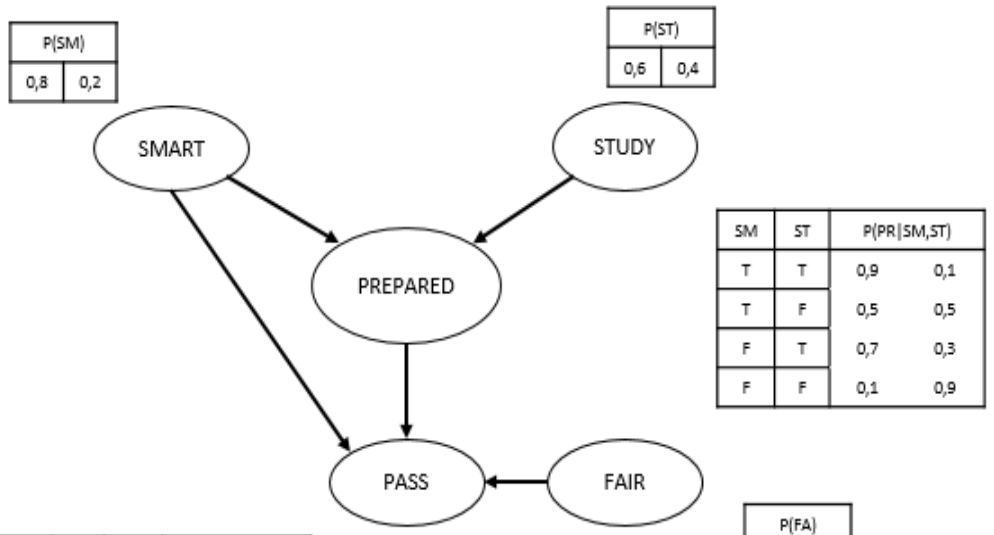
elisabetta.fersini@unimib.it

Esercitazione



Esercizio 1

- Data la seguente rete bayesiana, generare N=1000 campioni mediante Direct Sampling.



P(SM)	
0,8	0,2

P(ST)	
0,6	0,4

SM	ST	P(PR|SM,ST)	
T	T	0,9	0,1
T	F	0,5	0,5
F	T	0,7	0,3
F	F	0,1	0,9

F	PR	SM	P(PA|PR,SM,ST)	
T	T	T	0,9	0,1
T	T	F	0,7	0,3
T	F	T	0,7	0,3
T	F	F	0,2	0,8
F	T	T	0,1	0,9
F	T	F	0,1	0,9
F	F	T	0,1	0,9
F	F	F	0,1	0,9

P(FA)	
0,9	0,1

Esercizio 1

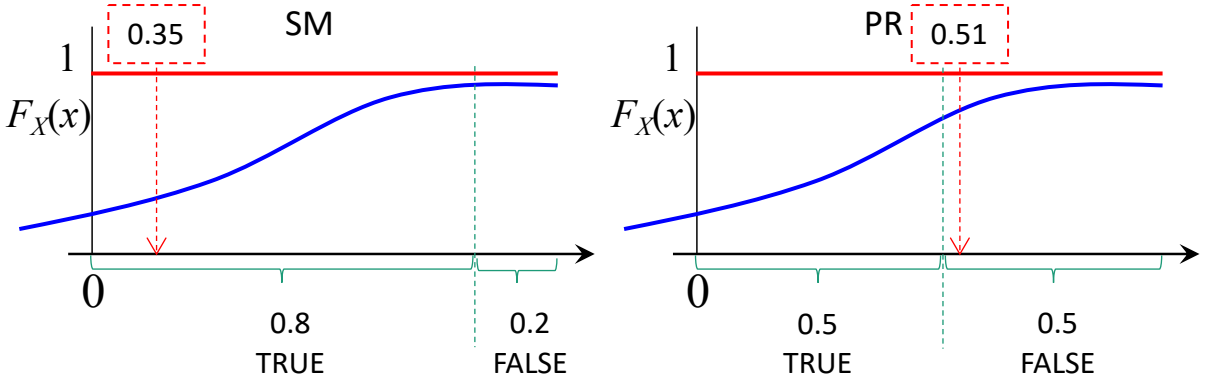
- Data la seguente rete bayesiana, generare N=1000 campioni mediante Direct Sampling.

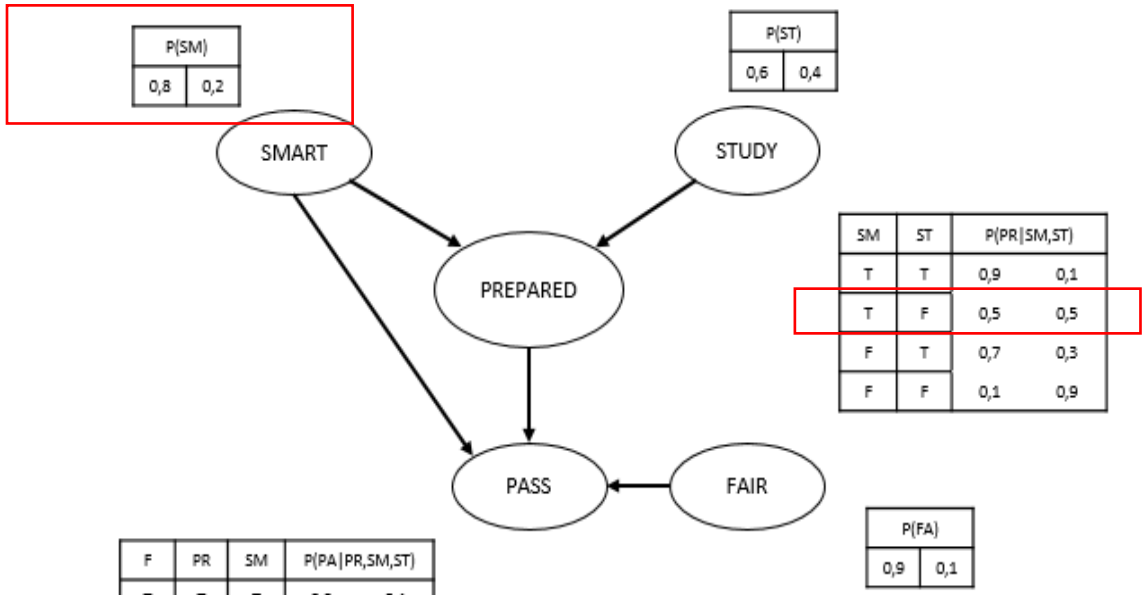
ORDINE TOPOLOGICO: SM, ST, PR, FA, PA

RANDOM NUMBER GENERATOR:

- [0.35, 0.76, 0.51, 0.44, 0.08] ➔ S1

S1 → [0.35 → SM = T ; 0.76 → ST = F ; 0.51 → PR = F ; 0.44 → FA = T ; 0.08 → PA = T]





SM	ST	P(PR|SM,ST)	
T	T	0,9	0,1
T	F	0,5	0,5
F	T	0,7	0,3
F	F	0,1	0,9

P(FA)	
0,9	0,1

F	PR	SM	P(PA|PR,SM,ST)	
T	T	T	0,9	0,1
T	T	F	0,7	0,3
T	F	T	0,7	0,3
T	F	F	0,2	0,8
F	T	T	0,1	0,9
F	T	F	0,1	0,9
F	F	T	0,1	0,9
F	F	F	0,1	0,9

Esercizio 1

- Data la seguente rete bayesiana, generare N=1000 campioni mediante Direct Sampling.

ORDINE TOPOLOGICO: SM, ST, PR, FA, PA

RANDOM NUMBER GENERATOR:

- [0.35, 0.76, 0.51, 0.44, 0.08] ➔ S1

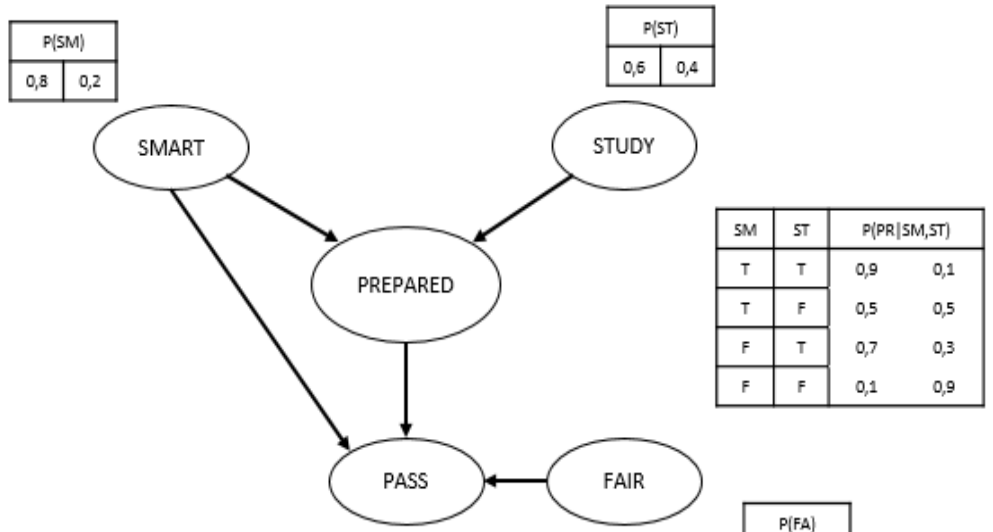
S1 → [0.35 → SM = T ; 0.76 → ST = F ; 0.51 → PR = F ; 0.44 → FA = T ; 0.08 → PA = T]

- [0.28, 0.03, 0.92, 0.92, 0.42] ➔ S2

S2 → [0.28 → SM = T ; 0.03 → ST = T ; 0.92 → PR = F ; 0.92 → FA = F ; 0.42 → PA = F]

…

S1000 ….



P(SM)	
0,8	0,2

P(ST)	
0,6	0,4

SM	ST	P(PR|SM,ST)	
T	T	0,9	0,1
T	F	0,5	0,5
F	T	0,7	0,3
F	F	0,1	0,9

P(FA)	
0,9	0,1

F	PR	SM	P(PA|PR,SM,ST)	
T	T	T	0,9	0,1
T	T	F	0,7	0,3
T	F	T	0,7	0,3
T	F	F	0,2	0,8
F	T	T	0,1	0,9
F	T	F	0,1	0,9
F	F	T	0,1	0,9
F	F	F	0,1	0,9

Esercizio 1

- Data la seguente rete bayesiana, generare N=1000 campioni mediante Direct Sampling.

ORDINE TOPOLOGICO: SM, ST, PR, FA, PA

RANDOM NUMBER GENERATOR:

- [0.35, 0.76, 0.51, 0.44, 0.08] ➔ S1

S1 → [0.35 → SM = T ; 0.76 → ST = F ; 0.51 → PR = F ; 0.44 → FA = T ; 0.08 → PA = T]

- [0.28, 0.03, 0.92, 0.92, 0.42] ➔ S2

S2 → [0.28 → SM = T ; 0.03 → ST = T ; 0.92 → PR = F ; 0.92 → FA = F ; 0.42 → PA = F]

…

S1000 ….

Con 1000 campioni:

- 690 PA = T
- 310 PA = F

< 0.69 ; 0.31 >

Esercizio 1

- Data la seguente rete bayesiana, generare N=1000 campioni mediante Direct Sampling.

ORDINE TOPOLOGICO: SM, ST, PR, FA, PA

RANDOM NUMBER GENERATOR:

- [0.35, 0.76, 0.51, 0.44, 0.08] ➔ S1

S1 → [0.35 → SM = T ; 0.76 → ST = F ; 0.51 → PR = F ; 0.44 → FA = T ; 0.08 → PA = T]

- [0.28, 0.03, 0.92, 0.92, 0.42] ➔ S2

S2 → [0.28 → SM = T ; 0.03 → ST = T ; 0.92 → PR = F ; 0.92 → FA = F ; 0.42 → PA = F]

…

S1000 ….

Con 1000 campioni:

- 690 PA = T
- 310 PA = F

< 0.69 ; 0.31 >

Esercizio 2

- Data la seguente rete bayesiana, calcolare $P(PA = T | ST = F)$ tramite Rejection Sampling.

ORDINE TOPOLOGICO: SM, ST, PR, FA, PA

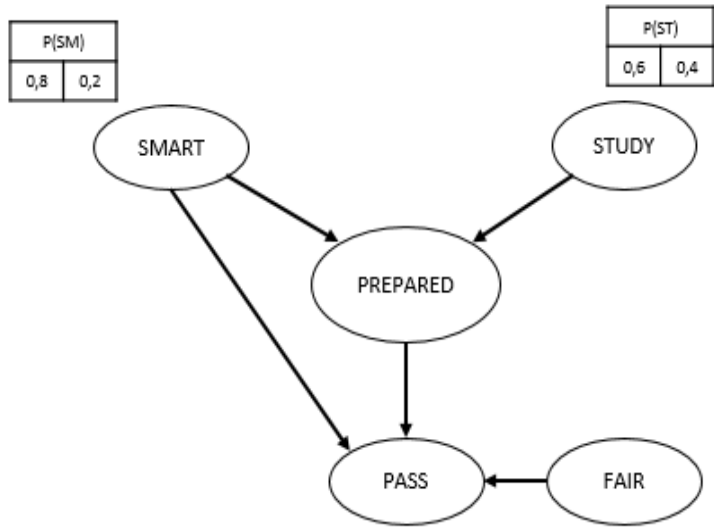
RANDOM NUMBER GENERATOR:

- [0.35, 0.76, 0.51, 0.44, 0.08] ➔ S1

S1 → [0.35 → SM = T ; 0.76 → ST = F ; 0.51 → PR = F ; 0.44 → FA = T ; 0.08 → PA = T] ➔ ACCEPT

- [0.28, 0.03, 0.92, 0.92, 0.42] ➔ S2

S2 → [0.28 → SM = T ; 0.03 → ST = T ; 0.92 → PR = F ; 0.92 → FA = F ; 0.42 → PA = F] ➔ REJECT



P(SM)	
0,8	0,2

P(ST)	
0,6	0,4

SM	ST	P(PR|SM,ST)	
T	T	0,9	0,1
T	F	0,5	0,5
F	T	0,7	0,3
F	F	0,1	0,9

P(FA)	
0,9	0,1

F	PR	SM	P(PA|PR,SM,ST)	
T	T	T	0,9	0,1
T	T	F	0,7	0,3
T	F	T	0,7	0,3
T	F	F	0,2	0,8
F	T	T	0,1	0,9
F	T	F	0,1	0,9
F	F	T	0,1	0,9
F	F	F	0,1	0,9

Esercizio 2

- Data la seguente rete bayesiana, calcolare $P(PA = T|ST = F)$ tramite Rejection Sampling (1000 campioni).

ORDINE TOPOLOGICO: SM, ST, PR, FA, PA

RANDOM NUMBER GENERATOR:

- [0.35, 0.76, 0.51, 0.44, 0.08] ➔ S1

S1 → [0.35 → SM = T ; 0.76 → ST = F ; 0.51 → PR = F ; 0.44 → FA = T ; 0.08 → PA = T] ➔ ACCEPT

- [0.28, 0.03, 0.92, 0.92, 0.42] ➔ S2

S2 → [0.28 → SM = T ; 0.03 → ST = T ; 0.92 → PR = F ; 0.92 → FA = F ; 0.42 → PA = F] ➔ REJECT

Abbiamo generato 1000 campioni:

- 730 campioni con ST=T ➔ REJECT!
- 270 campioni con ST=F
  - 130 campioni con PA=T
  - 140 con PA=F

$$P(PA = T|ST = F) = \alpha < 130; 140 >$$

$$= < 0.48; 0.52 >$$

Esercizio 2

- Data la seguente rete bayesiana, calcolare $P(PA = T|ST = F)$ tramite Rejection Sampling (1000 campioni).

ORDINE TOPOLOGICO: SM, ST, PR, FA, PA

RANDOM NUMBER GENERATOR:

- [0.35, 0.76, 0.51, 0.44, 0.08] ➔ S1

S1 → [0.35 → SM = T ; 0.76 → ST = F ; 0.51 → PR = F ; 0.44 → FA = T ; 0.08 → PA = T] ➔ ACCEPT

- [0.28, 0.03, 0.92, 0.92, 0.42] ➔ S2

S2 → [0.28 → SM = T ; 0.03 → ST = T ; 0.92 → PR = F ; 0.92 → FA = F ; 0.42 → PA = F] ➔ REJECT

Abbiamo generato 1000 campioni:

- 730 campioni con ST=T ➔ REJECT!
- 270 campioni con ST=F
  - 130 campioni con PA=T
  - 140 con PA=F

$$P(PA = T|ST = F) = \alpha < 130; 140 >$$

$$= < 0.48; 0.52 >$$

Esercizio 3

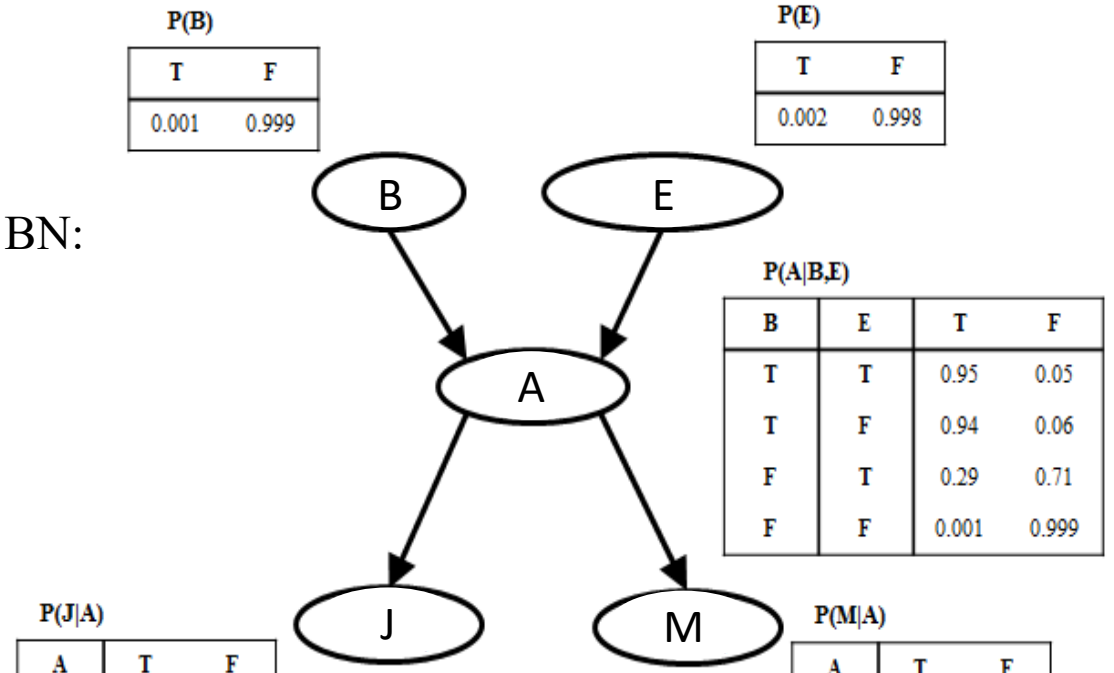
- Data la seguente rete bayesiana, calcolare $P(B = true | J = true, M = false)$ tramite Likelihood Weighting.

Ordine topologico: B, E, A, J, M

Numeri pseudocasuali per la generazioni dei campioni della BN:

- [0; 0.89; 0.3; /; /] ➔ S1
- [0.7; 0.9; 0.6; /;/] ➔ S2
- ….
- […..] ➔ S1000

In questo esercizio genereremo SOLO 2 campioni



P(B)

T	F
0.001	0.999

P(E)

T	F
0.002	0.998

P(A|B,E)

B	E	T	F
T	T	0.95	0.05
T	F	0.94	0.06
F	T	0.29	0.71
F	F	0.001	0.999

P(J|A)

A	T	F
T	0.9	0.1
F	0.05	0.95

P(M|A)

A	T	F
T	0.7	0.3
F	0.01	0.99

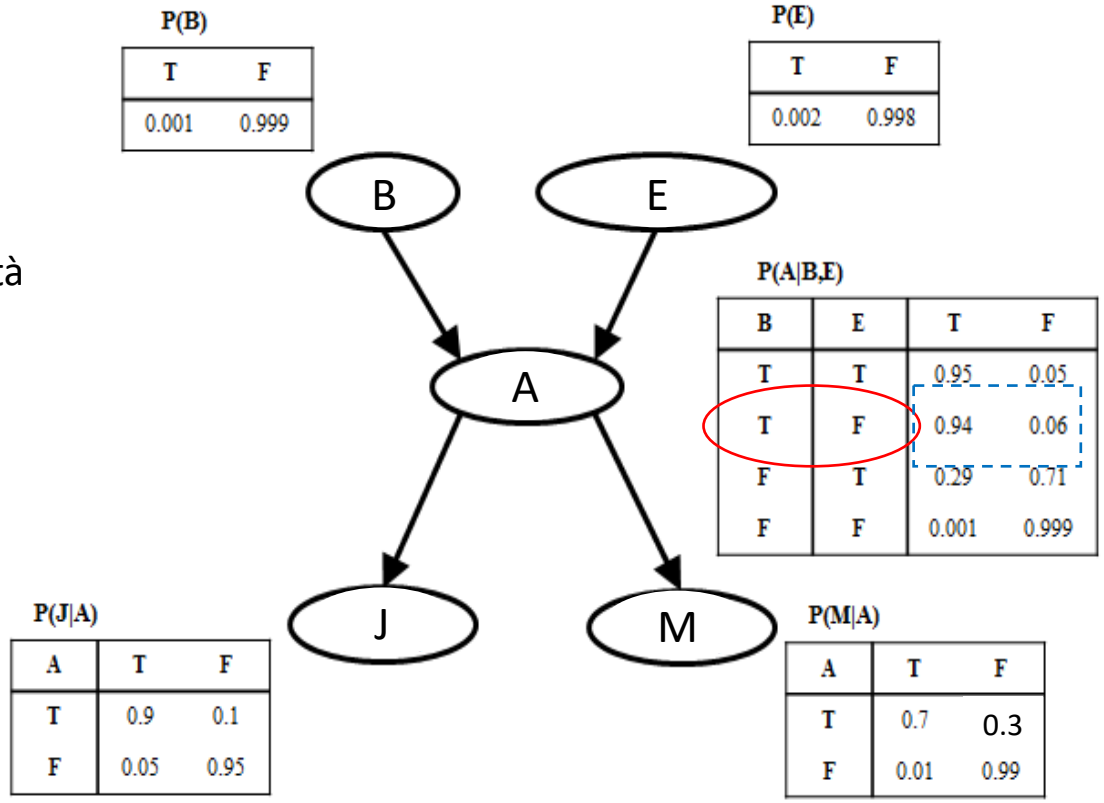
Esercizio 3

- Data la seguente rete bayesiana, calcolare $P(B = true | J = true, M = false)$ tramite Likelihood Weighting.

Ordine topologico: B, E, A, J, M

1. Fisso le evidenze
2. Genero 1000 campioni in accordo alle distribuzioni di probabilità (usando i numeri pseudocasuali) e seguendo l'ordine topologico:

- S1 = [0.9994; 0.89; 0.3; /; /] ➔ [true, false, true, true, false]
  - w_{s1} = P(J=true|A=true)*P(M=false|A=true) = 0.9*0.3 = 0.27
- S2 = [0.7; 0.9; 0.6; /;/] ➔ [false, false, false, true, false]
  - w_{s2} = P(J=true|A=false)*P(M=false|A=false) = 0.05*0.99 = 0.0495



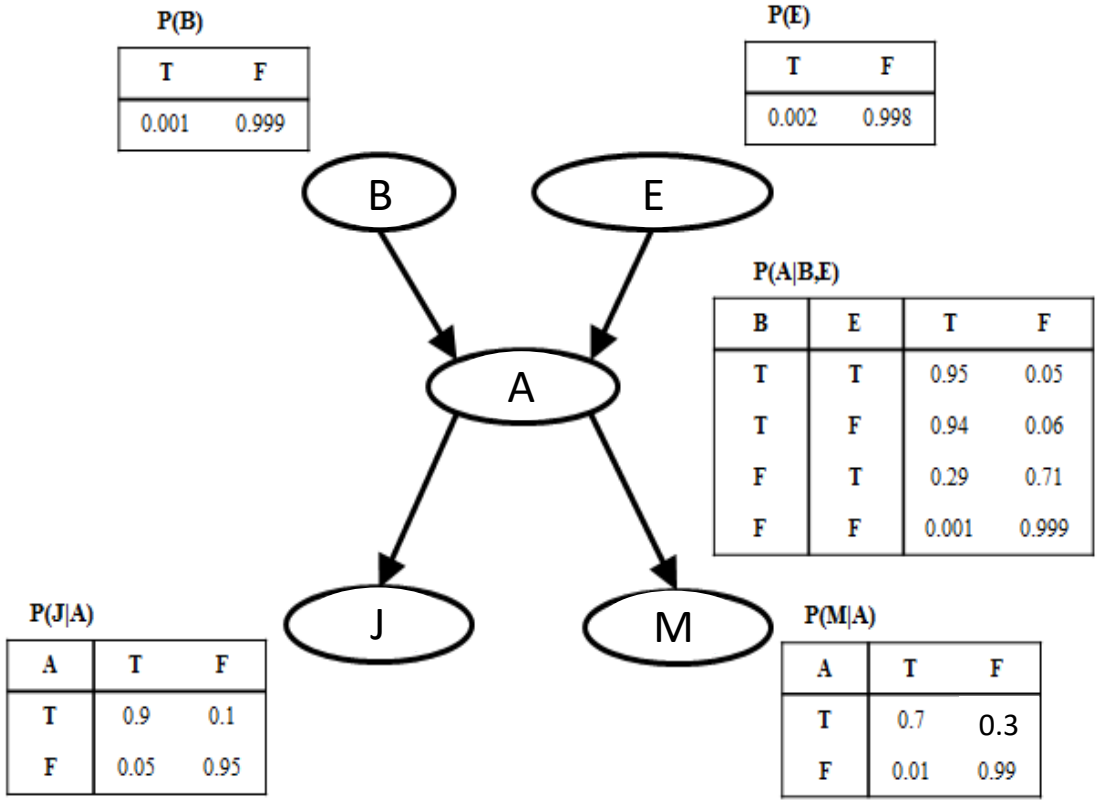
Esercizio 3

- Data la seguente rete bayesiana, calcolare $P(B = true|J = true, M = false)$ tramite Likelihood Weighting.

Stimiamo quindi P(B=true|J=true,M=false):

$$\mathrm{P}(\mathrm{B} = \mathrm{true}|\mathrm{J} = \mathrm{true}, \mathrm{M} = \mathrm{false}) = \frac{\sum_{s_i:B=true} w_{s_i}}{\sum_{s_i} w_{s_i}}$$

$$= \frac{0.27}{0.27 + 0.0495} = 0.8450$$

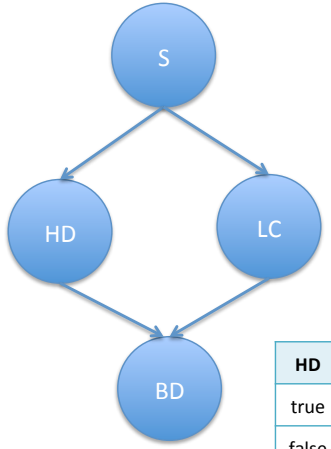


Esercizio 4

- Data la seguente rete bayesiana, calcolare P(HD|S=t,BD=t) mediante il metodo Markov Chain Monte Carlo.

1. Fisso le evidenze e campiono una variabile alla volta.
2. Inizializzo in modo random le variabili non osservate, LC=t e HD=f, e determino quindi lo **stato iniziale** S0 = [S=t, HD=f, LC=t, BD=t]
3. Campiono le variabili non osservate (variabili nascoste e la variabile query) dato la relativa Markov Blanket

$$P(x|MB(x)) = \alpha \cdot P(x|Pa(x)) \cdot \prod_{y \in Children(x)} P(y|Pa(y))$$



P(S=true)	P(S=false)
0.4	0.6

S	P(HD=true)	P(HD=false)
true	0.4	0.6
false	0.3	0.7

S	P(LC=true)	P(LC=false)
true	0.7	0.3
false	0.2	0.8

HD	LC	P(BD=true)	P(BD=false)
true	true	0.4	0.6
false	true	0.7	0.3
true	false	0.5	0.5
false	false	0.2	0.8

Esercizio 4

- Data la seguente rete bayesiana, calcolare P(HD|S=t,BD=t) mediante il metodo Markov Chain Monte Carlo.

3. Campiono LC dato le variabili della sua MB:
   - $P(LC = t|S = t, HD = f, BD = t) =$

     $= P(LC = t|S = t)P(BD = t|LC = t, HD = f)$

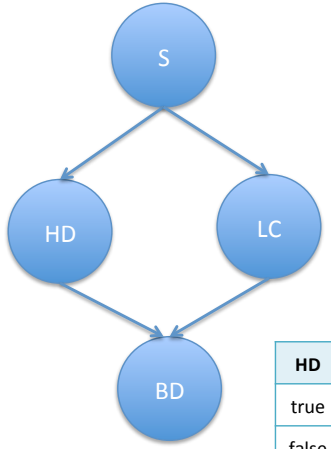
     $= 0.7 * 0.7 = 0.49$
   - $P(LC = f|S = t, HD = f, BD = t) =$

     $= P(LC = f|S = t)P(BD = t|LC = f, HD = f)$

     $= 0.3 * 0.2 = 0.06$

     $P(LC|S = t, HD = f, BD = t) =$

     $\alpha < 0.49; 0.06 > = < 0.89; 0.11 >$
   - campiono LC, ottenendo LC=f



P(S=true)	P(S=false)
0.4	0.6

S	P(HD=true)	P(HD=false)
true	0.4	0.6
false	0.3	0.7

S	P(LC=true)	P(LC=false)
true	0.7	0.3
false	0.2	0.8

HD	LC	P(BD=true)	P(BD=false)
true	true	0.4	0.6
false	true	0.7	0.3
true	false	0.5	0.5
false	false	0.2	0.8

S0 = [S=t, HD=f, LC=t, BD=t]

S1 = [S=t, HD=f, LC=f, BD=t]

Esercizio 4

- Data la seguente rete bayesiana, calcolare P(HD|S=t,BD=t) mediante il metodo Markov Chain Monte Carlo.

4. Campiono HD dato le variabili della sua MB:

- $P(HD = t|S = t, LC = f, BD = t)$

$$= P(HD = t|S = t)P(BD = t|LC = f, HD = t)$$

$$= 0.4 * 05 = 0.2$$

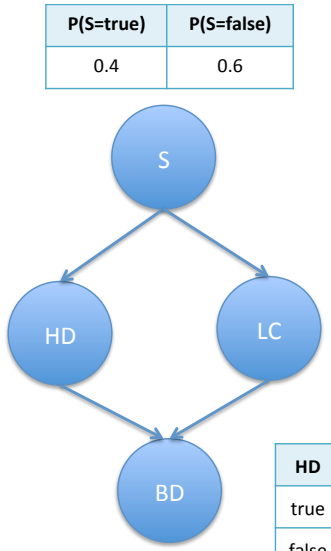
- $P(HD = f|S = t, LC = f, BD = t)$

$$= P(HD = f|S = t)P(BD = t|LC = f, HD = f)$$

$$= 0.6 * 0.7 = 0.42$$

$$P(HD|S = t, LC = f, BD = t) =$$
$$\alpha < 0.2; 0.42 > = < 0.32; 0.68 >$$

campiono HD, ottenendo HD=t



P(S=true)	P(S=false)
0.4	0.6

S	P(HD=true)	P(HD=false)
true	0.4	0.6
false	0.3	0.7

S	P(LC=true)	P(LC=false)
true	0.7	0.3
false	0.2	0.8

HD	LC	P(BD=true)	P(BD=false)
true	true	0.4	0.6
false	true	0.7	0.3
true	false	0.5	0.5
false	false	0.2	0.8

S0 = [S=t, HD=f, LC=t, BD=t]

S1 = [S=t, HD=f, LC=f, BD=t]

S2 = [S=t, HD=t, LC=f, BD=t]

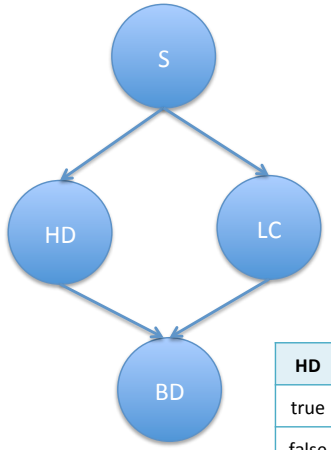
Esercizio 4

- Data la seguente rete bayesiana, calcolare P(HD|S=t,BD=t) mediante il metodo Markov Chain Monte Carlo.

5. Itero il processo di campionamento generando N=1000 stati della rete. Supponendo che in 800 campioni HD=t e 200 campioni HD=f:

$$P(HD|S=t,BD=t) = \langle 0.8;0.2 \rangle$$

P(S=true)	P(S=false)
0.4	0.6



S	P(HD=true)	P(HD=false)
true	0.4	0.6
false	0.3	0.7

S	P(LC=true)	P(LC=false)
true	0.7	0.3
false	0.2	0.8

HD	LC	P(BD=true)	P(BD=false)
true	true	0.4	0.6
false	true	0.7	0.3
true	false	0.5	0.5
false	false	0.2	0.8