1 Esercitazione 1

1.1 Esercizio 1

Dimostrare che $P(a|b \wedge a) = 1$.

Possiamo riscrivere $P(a|b \wedge a)$ come

$$P(a|b \wedge a) = \frac{P(a \wedge b \wedge a)}{P(b \wedge a)} \tag{1}$$

Questo per la definizione di probabilità condizionata che dice che:

$$P(A|B) = \frac{P(A \wedge B)}{P(B)}$$

Sappiamo anche che $P(a \wedge b \wedge a) = P(a \wedge b)$, per cui possiamo riscrivere l'equazione 1 come

$$P(a|b \wedge a) = \frac{P(a \wedge b)}{P(b \wedge a)}$$

e dato che $P(a \wedge b) = P(b \wedge a)$ possiamo concludere che $P(a|b \wedge a) = 1$.

1.2 Esercizio 2

Date le seguenti belief di un agente razionale:

$P(A) = 0.4$ $\quad$ $P(B) = 0.3$ $\quad$ $P(A \vee B) = 0.5$

Quale range di probabilità è ragionevole per $A \wedge V$?

Sapendo che

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$

y

$$0.4 + 0.3 - P(A \wedge B) = 0.5$$

Da cui:

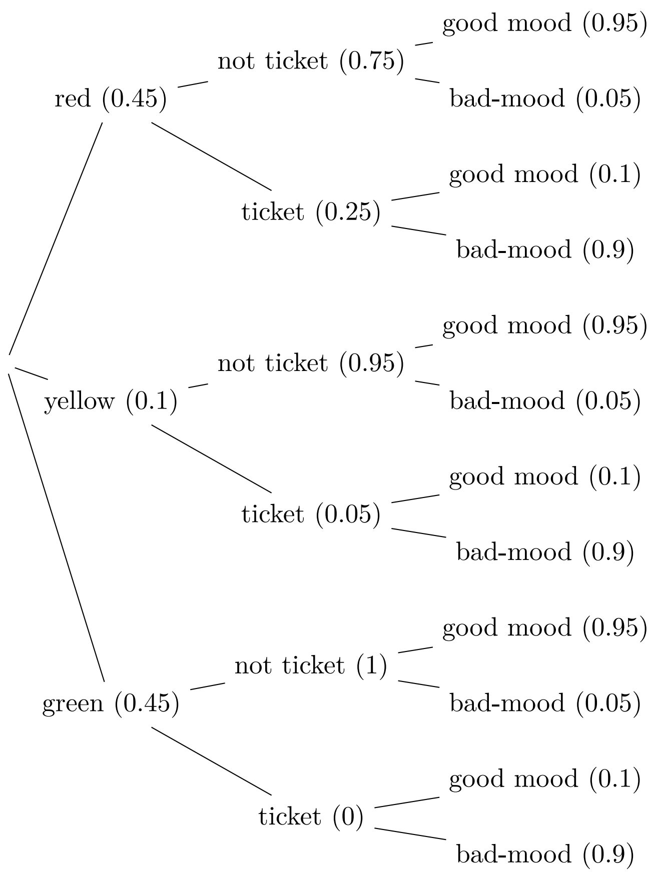
$$P(A \wedge B) = 0.2$$

1.3 Esercizio 3

Supponiamo di conoscere che la probabilità con cui un semaforo diventi verde sia pari a 0.45, arancione pari a 0.1 e che diventi rosso sia pari a 0.45. Inoltre, supponiamo di avere la probabilità del 25% di passare con il semaforo rosso senza prendere una multa , e il 5% di probabilità di prendere una multa passando con il semaforo arancione

In aggiunta, supponiamo che prendendo una multa, c'è il 90% di probabilità che si sia successivamente di cattivo umore; senza multa la probabilità è del 5%.

Qual è la probabilità totale di essere di cattivo umore?



Possiamo costruire la matrice delle probabilità congiunte:

	Green		Yellow		Red	
	Ticket	¬Ticket	Ticket	¬Ticket	Ticket	¬Ticket
Bad	0.45 * 0 * 0.9	0.45 * 1 * 0.05	0.1 * 0.05 *0.9	0.1 * 0.95 *0.05	0.45 * 0.25 *0.9	0.45 * 0.75 * 0.05
Good	0.45 * 0 * 0.1	0.45 * 1 * 0.95	0.1 * 0.05 *0.1	0.1 * 0.95 *0.95	0.45 * 0.25 *0.1	0.45 * 0.75 * 0.95

Da cui:

	Green		Yellow		Red	
	Ticket	¬Ticket	Ticket	¬Ticket	Ticket	¬Ticket
Bad	0	0.0225	0.0045	0.00475	0.10125	0.016875
Good	0	0.4275	0.00005	0.09025	0.01125	0.320625

1.4 Esercizio 4

Una società di consulenza ha creato un modello per prevedere le recessioni. Il modello prevede una recessione con l'80% di probabilità quando la recessione avviene realmente e con il 10% di probabilità quando non avviene. La probabilità incondizionata che si entri in una fase di recessione è del 20%. Se il modello prevede la recessione, qual è la probabilità che la recessione avvenga?

Sappiamo che

$$P(rec.\ pred. \mid rec.\ coming) = \frac{8}{10}$$

,che

$$P(rec.pred. \mid rec.\ not\ coming) = \frac{1}{10}$$

e che $P(rec.\ coming) = \frac{2}{10}$

Vogliamo trovare $P(rec.\ coming \mid rec.\ pred)$

Per la regola di Bayes sappiamo che :

$$P(rec.\ coming \mid rec.\ pred) = \frac{P(rec.\ pred \mid rec.\ coming) * P(rec.\ coming)}{P(rec.\ pred)}$$

Qual'è la $P(rec.\ pred)$?

Sappiamo che

$$P(rec.\ pred) = P(rec.\ pred \mid rec.coming) * P(rec.\ coming) + P(rec.\ pred \mid rec.\ not\ coming) * P(rec.\ not\ coming)$$

Inoltre se $P(rec.\ coming) = \frac{2}{10}$ segue che $P(rec.\ not\ coming) = \frac{8}{10}$

Da cui:

$$P(rec.\ pred) = \frac{8}{10} \cdot \frac{2}{10} + \frac{1}{10} \cdot \frac{8}{10} = \frac{24}{100}$$

Concludendo quindi che :

$$P(rec.\ coming \mid rec.\ pred) = \frac{\frac{8}{10} \cdot \frac{2}{10}}{\frac{24}{100}} = \frac{\frac{16}{100}}{\frac{24}{100}} = \frac{16}{100} \cdot \frac{100}{24} = \frac{2}{3}$$

1.5 Esercizio 5

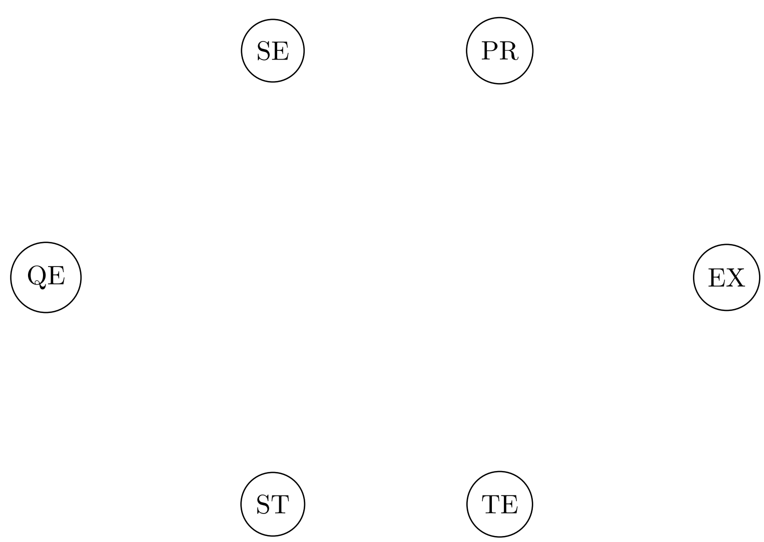
Sviluppare una rete di Bayes, per calcolare la probabilità che uno studente superi l'esame di MPD.
Le proprietà di interesse del problema sono:

- Il superamento dell'esame $EX \in \{true, false\}$
- L'acquisizione di buone capacità pratiche in MPD da parte dello studente $PR \in \{true, false\}$
- L'acquisizione di buone capacità teoriche in MPD da parte dello studente $TE \in \{true, false\}$
- Lo studente è efficientemente studioso $ST \in \{true, false\}$
- La quantità di esercitazioni seguite dallo studente $QE \in \{molte, poche, nessuna\}$
- L'aver fatto un numero sufficiente di esercizi $SE \in \{true, false\}$

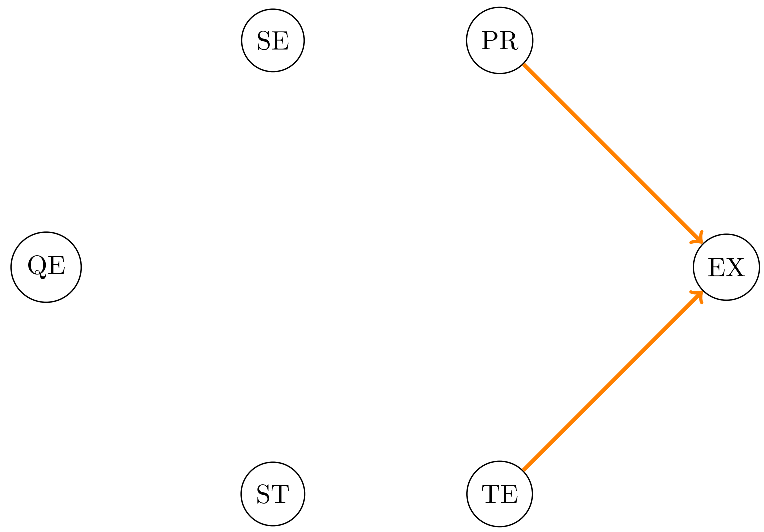
Costruire una rete bayesiana che rappresenti la conoscenza probabilistica relativa al dominio descritto dalle seguenti relazioni di dipendenza tra le variabili casuali:

- Il superamento dell'esame dipende dalle capacità teoriche e pratiche dello studente
- Se uno studente è studioso ha buone probabilità di acquisire capacità teoriche.
- Il numero di esercitazioni seguite dipende da quanto uno studente è studioso
- L'aver fatto sufficienti esercizi dipende dal numero di esercitazioni seguite, ed influenza le capacità pratiche dello studente.

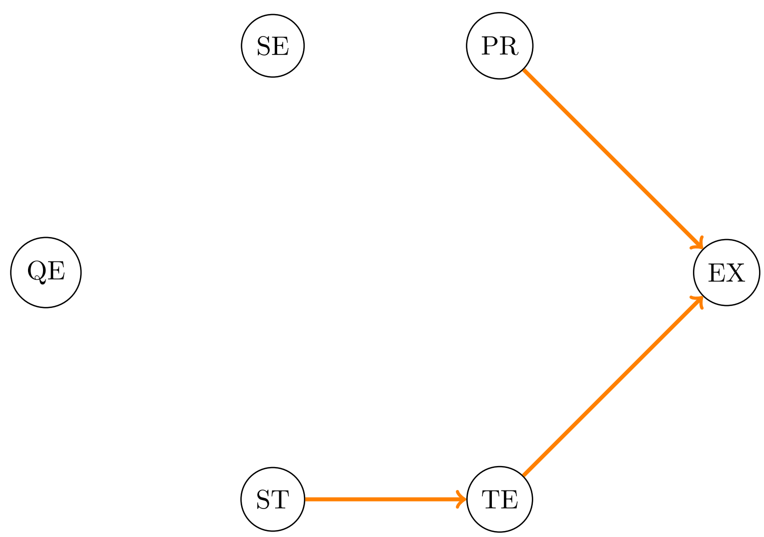
Dobbiamo un creare un DAG a partire dai sequenti nodi:



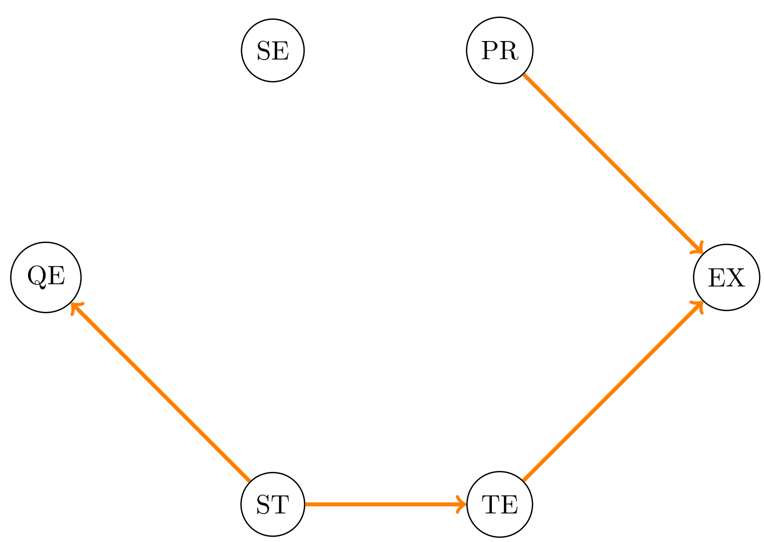
Sappiamo che il superamento dell'esame EX dipende dalle capacità teoriche (TE) e pratiche (PR) dello studente



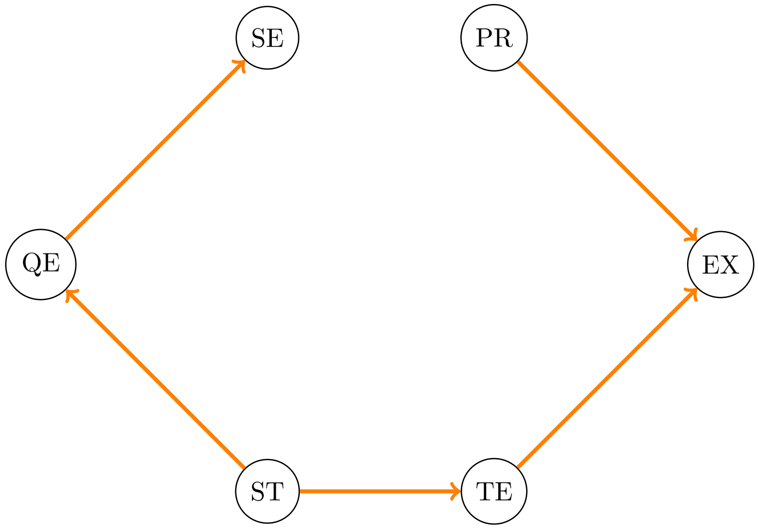
Sappiamo anche che se uno studente è studioso (ST) ha buone probabilità di acquisire capacità teoriche (TE).



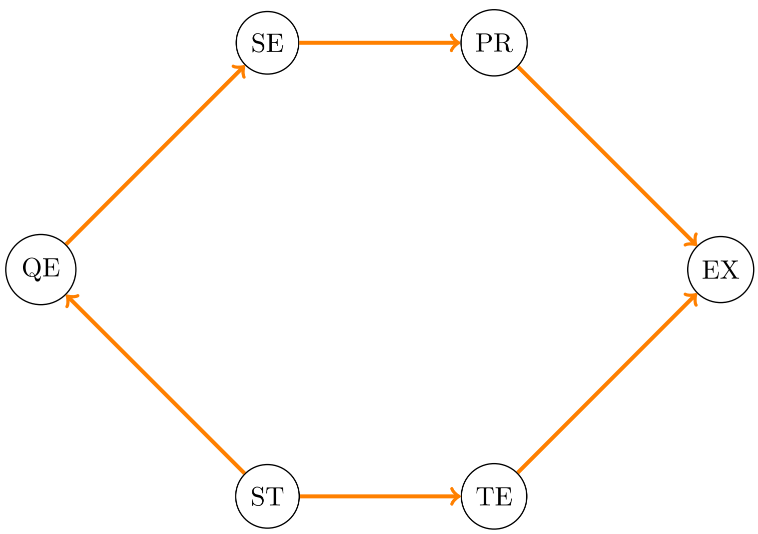
Sappiamo anche che la quantità di esercitazioni seguite (TE) dipende da quanto lo studente sia studioso (ST)



Inioltre sappiamo è che avere fatto un numero sufficente di esercizi (SE) dipende dalla quantità di esercitazioni seguite (QE)



Infine sappiamo che avere svolto un numero sufficiente di esercizi influenza le capacità pratiche dello studente.



Abbiamo ora la struttura topologica, è neccessario avere anche le distribuzione di probabilità a priori o condizionate.

Il nodo (ST) ha una distribuzione di proprità a priori per i suoi possibili valori ($true, false$).

Tutte le altre probabilità sono condizionate da uno o più genitori:

Prendiamo dei dati a caso come esempio:

$P(ST)$	
$true$	$false$
0.7	0.3

Per quando riguarda ad esempio la probabilità TE è condizionata da un solo genitore ST, avremo quindi una situazione di questo tipo:

ST	P(TE |ST)	
	T	F
true	x	1-x
false	y	1-y

Se prendiamo una probabilità con due genitori, come EX, avremo una situazione di questo tipo:

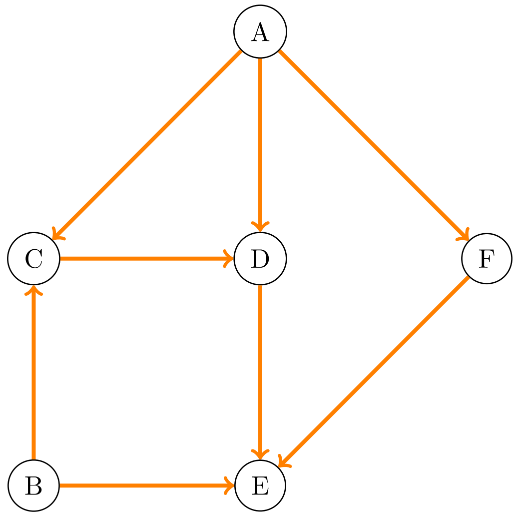
		P(EX |TE, PR)	
TE	PR	true	false
true	true	x	
false	false	y	
false	true	z	
false	false	a	

Proviamo a costruire la tabella della probabilità QE con dei valori plausibili, un esempio potrebbe essere:

ST	P(EX |TE, PR)		
	molte	poche	nessuna
true	0.5	0.3	0.2
false	0	0.1	0.9

1.6 Esercizio 6

Consideriamo questa rete bayesiana:



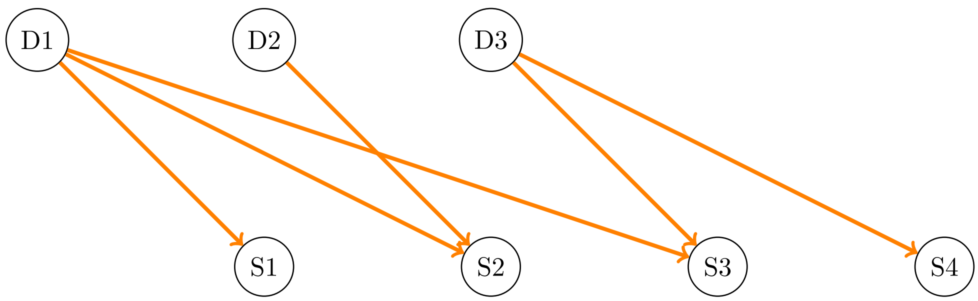
Quali sono le probabilità da definire per avere una rete bayesiana?

- $P(A)$
- $P(B)$
- $P(C|B,A)$
- $P(D|C,A)$
- $P(E|B,D,F)$
- $P(F|A)$

1.7 Esercizio 7

Un paziente si reca dal dottore per sottoporre una patologia, il dottore sospetta 3 possibili malattie come causa della condizione patologica. Le 3 malattie sono D_1, D_2, D_3 le quali sono marginalmente indipendenti tra loro. Ci sono 4 sintomi S_1, S_2, S_3, S_4 di cui il dottore vuole verificare la presenza in modo da trovare la causa più probabile per la condizione patologica. I sintomi sono condizionalmente dipendenti alle 3 malattie come segue: S_1 dipende solamente da D_1, S_2 dipende da D_1 e da D_2 , S_3 dipende da D_1 e da D_3, e S_4 dipende solamente da D_3. Si assuma che tutte le variabili casuali siano booleane.

- Costruire la struttura topologica alla rete bayesiana per il problema descritto.



- Scrivere la distribuzione di probabilità congiunta espressa come prodotto delle probabilità condizionate

$$P(D_1, D_2, D_3, S_1, S_2, S_3, S_4) =$$

$$= P(D_1) \cdot P(D_2) \cdot P(D_3) \cdot$$

$$= P(S_1 \mid D_1) \cdot P(S_2 \mid D_1, D_2) \cdot$$

$$= P(S_3 \mid D_1, D_3) \cdot P(S_4 \mid D_3)$$

- Qual è il numero di parametri indipendenti richiesti per descrivere la distribuzione congiunta?

Iniziamo ad immaginare le CPT (*conditional probability table*) di ogni nodo.
Ad esempio per D_1 avremo due valori nella CPT ($true\ e\ false$), così come D_2 e D_3.
Invece S_1 ha un genitore (D_1) per cui la tabella avrà quattro possibili valori.
Analogamente S_2 ha due (D_1 e D_2) per cui la tabella avrà otto possibili valori.
Seguendo con questo ragionamento arriveremo ad avere

$$2 + 2 + 2 + 4 + 8 + 8 + 4 = 30$$

parametri.

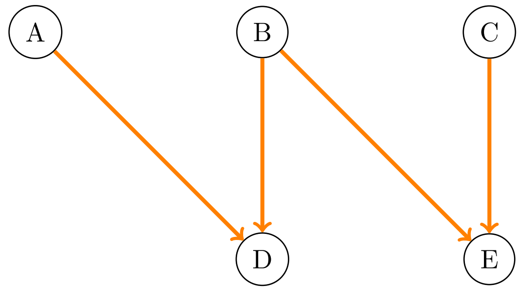
Come possiamo stimare i parametri indipendenti?
Essendo una distribuzione booleana per D_1, D_2 e D_3 serve un parametro per ciascuno.
Per S_1 ne serviranno 2, così come per S_4: per S_2 ed S_3 serviranno 4 variabili, in totale serviranno quindi 15 parametri.

- Assumendo che non ci sia dipendenza condizionale tra le variabili, quanti parametri indipendenti sarebbero dunque richiesti?

$$2^7 - 1$$

1.8 Esercizio 8

Supponiamo di avere la seguente rete Bayesiana su spazio booleano.



Sappiamo che $P(A = T) = 0.2$, $P(B = T) = 0.2$ e che $P(C = T) = 0.8$
Sappiamo inoltre che:

A	B	P(D = T | A,B)
T	T	0.1
T	F	0.5
F	T	0.6
F	F	0.9

e che

C	B	P(E = T | C,B)
T	T	0.3
T	F	0.8
F	T	0.4
F	F	0.2

- Qual è la probabilità che tutte le variabili siano false?

$$P(A = F, B = F, C = F, D = F, E = F) =$$

$$= P(A = F) \cdot P(B = F) \cdot P(C = F) \cdot P(D = F|A = F, B = F) \cdot P(E = F|B = F, C = F)$$

$$= 0.6 \cdot 0.5 \cdot 0.2 \cdot 0.1 \cdot 0.8 = 0.064$$

- Qual è la probabilità di A, avendo la conoscenza che tutte le altre variabili sono vere?

Formalmente la richiesta è

$$P(A|B = T, C = T, D = T, E = T)$$

che corrisponde all'unione di

$$P(A = T|B = T, C = T, D = T, E = T)$$

e

$$P(A = F|B = T, C = T, D = T, E = T)$$

https://elearning.unimib.it/pluginfile.php/1002302/mod_resource/content/2/Esercitazione%202.pdf