

UNIVERSITA' DEGLI STUDI DI FIRENZE

Facoltà di Ingegneria  
Corso di Laurea in Ingegneria Informatica

***Esercitazioni per l'esame di  
"Intelligenza Artificiale"  
Prof. G. Soda***

# ***Bayesian Networks***

di Sauro Menchetti

***A.A. 1998-99***

# Costruzione della Rete Bayesiana

Le reti bayesiane sono uno strumento utile in quelle situazioni in cui non è possibile agire con certezza. Un problema della logica del primo ordine è che non si ha quasi mai accesso all'intera verità: un fatto non è vero o falso in assoluto. Nella maggioranza dei casi, anche in situazioni semplici, ci saranno domande importanti a cui non si potrà dare una risposta categorica. Si deve quindi agire in presenza di incertezza. L'incertezza è causata principalmente da dati inaffidabili, mancanti o imprecisi.

## 1.1 Situazione da modellare

Qualsiasi studente iscritto ad una facoltà, per poter giungere alla laurea, deve superare tutti gli esami previsti dal proprio piano di studi. Lo studente può scegliere di frequentare o meno un determinato corso a seconda del tempo che ha a disposizione e dell'interesse verso la materia insegnata: se decide di frequentare il corso, sarà più facile superare l'esame. Quando lo studente decide di sostenere l'esame, si suppone che abbia una sufficiente padronanza della materia: il tempo a sua disposizione e la voglia di studiare incidono sulla preparazione globale e quest'ultima, insieme alla fortuna (che non guasta mai), possono determinare l'esito dell'esame.

## 1.2 Analisi del dominio

Si analizzano le caratteristiche salienti del dominio in esame che porteranno alla costruzione della rete bayesiana.

### 1.2.1 Variabili di dominio

Da un'analisi della precedente situazione, si possono ricavare le seguenti variabili aleatorie con i rispettivi stati:

- $E$  = superamento dell'esame (no =  $e_0$ , si =  $e_1$ );
- $F$  = frequenza del corso (scarsa =  $f_0$ , assidua =  $f_1$ );
- $T$  = tempo a disposizione (insufficiente =  $t_0$ , sufficiente =  $t_1$ );
- $I$  = interesse per la materia (scarso =  $i_0$ , sufficiente =  $i_1$ , molto =  $i_2$ );
- $P$  = preparazione (insufficiente =  $p_0$ , sufficiente =  $p_1$ , buona =  $p_2$ );

$V$  = voglia di studiare (scarsa =  $v_0$ , media =  $v_1$ , molta =  $v_2$ );  
 $C$  = fortuna (no =  $c_0$ , si =  $c_1$ ).

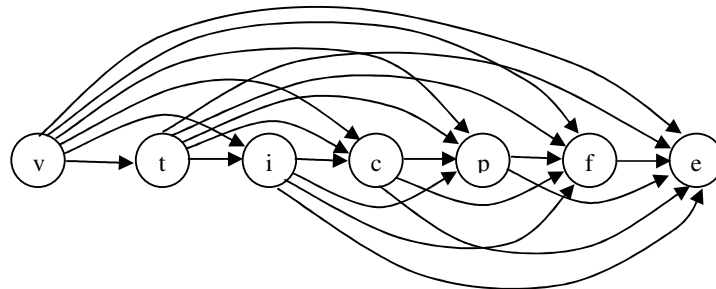
Gli indici assegnati ai vari stati servono a sostituire il nome dello stato della variabile aleatoria nelle tabelle che seguiranno.

### 1.2.2 Ordinamento delle variabili

Per costruire la rete bayesiana corrispondente alla precedente situazione, occorre dare un ordinamento alle variabili: l'ordinamento è cruciale, perché può indurre a ricercare condizioni di dipendenza non semplicemente esplicitabili. Tuttavia qualunque ordinamento garantisce la rappresentazione del dominio. La rete finale dipende comunque dall'ordinamento delle variabili. L'ordinamento scelto è il seguente:

$V, T, I, C, P, F, E.$

Come si può osservare, le variabili che non presentano dipendenze dirette con altre variabili vengono prima nell'ordinamento, seguite poi dalle rimanenti. La precedente sequenza di variabili corrisponde ad un ordinamento topologico del grafo diretto aciclico.



In un ordinamento topologico tutti i discendenti sono alla destra del nodo padre; inoltre ogni nodo è collegato con tutti i suoi discendenti in modo opportuno. Tale ordinamento topologico deve essere semplificato con le condizioni rilevate dal dominio.

### 1.2.3 Condizioni rilevate dal dominio

Esaminando il dominio da modellare, si possono rilevare le seguenti condizioni di dipendenza e di indipendenza:

$$V \perp \{T, I, C\} \mid \emptyset \text{ ed } E \perp \{V, T, I\} \mid \{P, F\}$$

cioè le variabili  $V, T, I, C$  sono indipendenti, mentre  $E$  dipende solo da  $V, T, I$  date  $P, F$ .

### 1.2.4 Legami tra le variabili

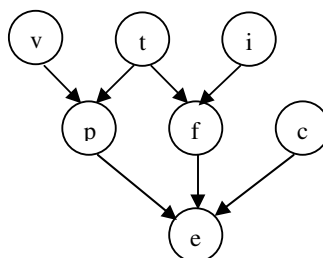
Rispettando l'ordine imposto alle variabili, si possono determinare quali variabili influenzano direttamente un'altra variabile:

variabili	dipendenza
$V$	$P(V)$ ;
$T$	$P(T   V) = P(T)$ ;
$I$	$P(I   T, V) = P(I)$ ;
$C$	$P(C   I, T, V) = P(C)$ ;
$P$	$P(P   C, I, T, V) = P(P   T, V)$ ;
$F$	$P(F   P, C, I, T, V) = P(F   I, T)$ ;
$E$	$P(E   F, P, C, I, T, V) = P(E   F, P, C)$ .

Come si vede, non tutti i legami dell'ordinamento topologico sono effettivamente presenti.

### 1.3 La Rete Bayesiana

Dalla precedente analisi, eliminando gli archi superflui dall'ordinamento topologico, si giunge alla seguente rete bayesiana dove le frecce indicano delle dipendenze dirette tra le variabili:



Le variabili  $V$  e  $T$  influenzano la variabile  $P$ , così come le variabili  $T$  ed  $I$  influenzano  $P$ . Inoltre la variabile  $E$  è influenzata direttamente da  $P$ ,  $F$  e  $C$ .

#### 1.3.1 Le tabelle di probabilità

Si riportano di seguito le tabelle di probabilità associate alla rete di Bayes. Vediamo prima le tabelle di probabilità a priori delle variabili  $V$ ,  $I$ ,  $T$  e  $C$ .

Voglia di studiare	$P(V)$
Scarsa	0.1
Media	0.2
Molta	0.7

Interesse	$P(I)$
Scarso	0.1
Sufficiente	0.3
Molto	0.6

Tempo a disposizione	$P(T)$
Insufficiente	0.3
Sufficiente	0.7

Fortuna	$P(C)$
No	0.5
Si	0.5

Ecco infine le probabilità condizionate delle variabili  $P$ ,  $F$  ed  $E$ .

Preparazione		$P(P   T, V)$		
Tempo	Voglia	Insufficiente	Sufficiente	Buona
Insufficiente	Scarsa	0.85	0.1	0.05
Insufficiente	Media	0.7	0.2	0.1
Insufficiente	Molta	0.5	0.3	0.2
Sufficiente	Scarsa	0.5	0.4	0.1
Sufficiente	Media	0.2	0.6	0.2
Sufficiente	Molta	0.02	0.08	0.9

Frequenza del corso		$P(F   I, T)$	
Interesse	Tempo	Scarsa	Assidua
Scarso	Insufficiente	0.8	0.2
Scarso	Sufficiente	0.65	0.35
Sufficiente	Insufficiente	0.7	0.3
Sufficiente	Sufficiente	0.35	0.65
Molto	Insufficiente	0.55	0.45
Molto	Sufficiente	0.1	0.9

Superamento dell'esame			$P(E   F, P, C)$	
Frequenza	Preparazione	Fortuna	No	Si
Scarsa	Insufficiente	No	0.95	0.05
Scarsa	Insufficiente	Si	0.9	0.1
Scarsa	Sufficiente	No	0.3	0.7
Scarsa	Sufficiente	Si	0.25	0.75
Scarsa	Buona	No	0.15	0.85
Scarsa	Buona	Si	0.1	0.9
Assidua	Insufficiente	No	0.9	0.1
Assidua	Insufficiente	Si	0.8	0.2
Assidua	Sufficiente	No	0.2	0.8
Assidua	Sufficiente	Si	0.1	0.9
Assidua	Buona	No	0.05	0.95
Assidua	Buona	Si	0.03	0.97

# Costruzione del Junction Tree

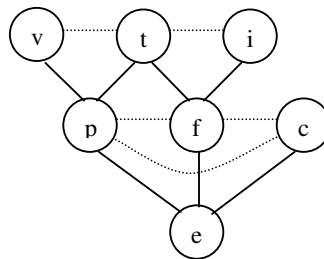
Le reti bayesiane non sono adatte per fare inferenza sul dominio di interesse: le tabelle di probabilità crescono in modo esponenziale e diventano presto intrattabili. Bisogna quindi trovare un tipo di rappresentazione più conveniente: il junction tree risolve il problema.

## 2.1 Il Junction Tree

Ci sono vari passi che portano alla costruzione del junction tree: esaminiamoli di seguito uno ad uno.

### 2.1.1 Costruzione del grafo moralizzato

Il grafo moralizzato viene costruito a partire dalla rete di Bayes: per ogni variabile  $A$  si crea un collegamento tra tutte le variabili in  $pr(A) \cup \{A\}$ . Si ottiene il seguente grafo non diretto:



Gli archi tratteggiati indicano i link di moralizzazione.

**Osservazione:** la rete di Bayes originaria è un grafo a connessione multipla. Tuttavia non occorre triangolarizzare il grafo moralizzato in quanto i link di moralizzazione svolgono tale compito.

### 2.1.2 Individuazione delle cricche

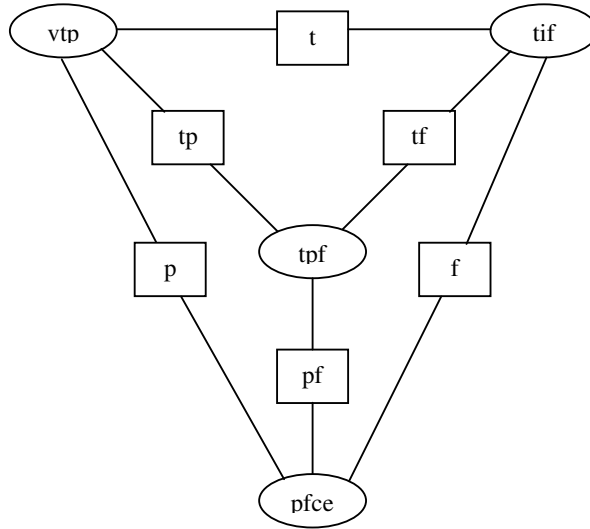
In questa fase si cerca di individuare tutte le cricche del grafo moralizzato che andranno a costituire i nodi del junction tree. Si identificano le seguenti cricche:

## VTP TIF TPF PFCE

che serviranno per il passo successivo.

### 2.1.3 Costruzione del Junction Graph

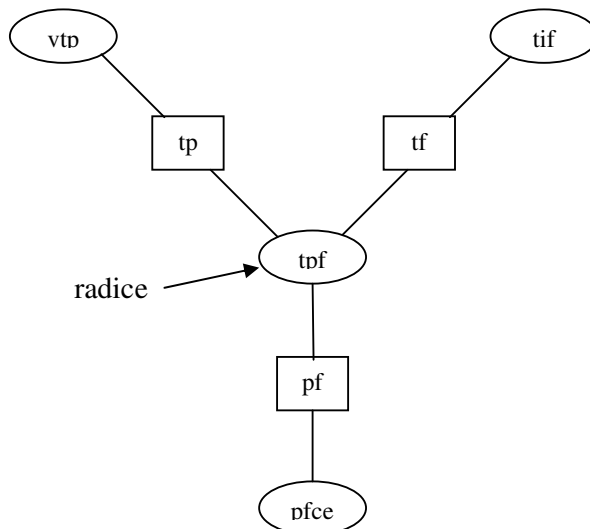
I nodi del junction graph sono le cricche precedentemente individuate: tra ogni coppia di nodi che hanno almeno una variabile in comune, si mette un link con un separatore che contiene l'intersezione delle variabili. Ecco cosa si ottiene:



I nodi sono indicati con delle ellissi, mentre i separatori con dei rettangoli.

### 2.1.4 Costruzione del Junction Tree

Dato il junction graph, si devono eliminare i link che costituiscono dei cicli in modo da ottenere un albero. Nell'eliminazione, vengono mantenuti i link a massima intersezione. Si ottiene il seguente albero:



Come si può vedere, per ogni coppia di nodi  $V, W$  tutti i nodi del cammino che li congiunge contengono l'intersezione di  $V \cap W$ .

## 2.2 Assegnazione delle tabelle al Junction Tree

Il passo successivo assegna le tabelle di probabilità ad ogni nodo ed ad ogni separatore del junction tree. È suddiviso in due fasi: una prima in cui si inizializzano le tabelle ed una seconda in cui modificano in modo iterativo i valori contenuti in esse.

### 2.2.1 Inizializzazione delle tabelle

Per prima cosa si inizializzano tutte le tabelle dei nodi e dei separatori ad uno:

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = 1$	$t_{TP} = 1$
$t_{TIF} = 1$	$t_{TF} = 1$
$t_{TPF} = 1$	$t_{PF} = 1$
$t_{PFCE} = 1$	

Gli indici indicano il nodo o il separatore a cui si riferisce la tabella.

### 2.2.2 Assegnazione delle tabelle

Per ciascuna variabile  $A$ , si sceglie un nodo  $V$  del junction tree contenente  $pr(A) \cup \{A\}$  (un tale nodo esiste per come è stato costruito l'albero) e si moltiplica  $P(A | pr(A))$  per la tavola di  $V$ : questa tabella aggiornata diventa la nuova tabella di  $V$  ( $pr$  è l'abbreviazione di parent e sta ad indicare i genitori di un nodo). Le variabili vengono scelte nel seguente ordine:  $V, T, I, C, P, F, E$ . Esaminiamo cosa succede passo per passo.

#### Variabile $V$

parent: nessuno

nodo: VTP

$$t_{VTP} = P(V) * 1$$

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(V)$	$t_{TP} = 1$
$t_{TIF} = 1$	$t_{TF} = 1$
$t_{TPF} = 1$	$t_{PF} = 1$
$t_{PFCE} = 1$	

#### Variabile $T$

parent: nessuno

nodo: TPF

$$t_{TPF} = P(T) * 1$$



tabelle dei nodi

$$\begin{aligned} t_{VTP} &= P(V) \\ t_{TIF} &= 1 \\ t_{TPF} &= P(T) \\ t_{PFCE} &= 1 \end{aligned}$$

tabelle dei separatori

$$\begin{aligned} t_{TP} &= 1 \\ t_{TF} &= 1 \\ t_{PF} &= 1 \end{aligned}$$

**Variabile I**

parent: nessuno

nodo: TIF

$$t_{TIF} = P(I) * 1$$

tabelle dei nodi

$$\begin{aligned} t_{VTP} &= P(V) \\ t_{TIF} &= P(I) \\ t_{TPF} &= P(T) \\ t_{PFCE} &= 1 \end{aligned}$$

tabelle dei separatori

$$\begin{aligned} t_{TP} &= 1 \\ t_{TF} &= 1 \\ t_{PF} &= 1 \end{aligned}$$

**Variabile C**

parent: nessuno

nodo: PFCE

$$t_{PFCE} = P(C) * 1$$

tabelle dei nodi

$$\begin{aligned} t_{VTP} &= P(V) \\ t_{TIF} &= P(I) \\ t_{TPF} &= P(T) \\ t_{PFCE} &= P(C) \end{aligned}$$

tabelle dei separatori

$$\begin{aligned} t_{TP} &= 1 \\ t_{TF} &= 1 \\ t_{PF} &= 1 \end{aligned}$$

**Variabile P**

parent: VT

nodo: VTP

$$t_{VTP} = P(P | V, T) P(V) = P(P, V | T)$$

tabelle dei nodi

$$\begin{aligned} t_{VTP} &= P(P, V | T) \\ t_{TIF} &= P(I) \\ t_{TPF} &= P(T) \\ t_{PFCE} &= P(C) \end{aligned}$$

tabelle dei separatori

$$\begin{aligned} t_{TP} &= 1 \\ t_{TF} &= 1 \\ t_{PF} &= 1 \end{aligned}$$

	$P(P, V   T)$								
$T$	$P_0 V_0$	$P_0 V_1$	$P_0 V_2$	$P_1 V_0$	$P_1 V_1$	$P_1 V_2$	$P_2 V_0$	$P_2 V_1$	$P_2 V_2$
$t_0$	0.085	0.14	0.35	0.01	0.04	0.21	0.005	0.02	0.14
$t_1$	0.05	0.04	0.014	0.04	0.12	0.056	0.01	0.04	0.63

La somma delle due righe della tabella fa uno.

**Variabile F**

parent: TI

nodo: TIF

$$t_{TIF} = P(F | T, I) P(I) = P(F, I | T)$$

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(P, V   T)$	$t_{TP} = 1$
$t_{TIF} = P(F, I   T)$	$t_{TF} = 1$
$t_{TPF} = P(T)$	$t_{PF} = 1$
$t_{PFCE} = P(C)$	

		$P(F, I   T)$					
$T$		$f_0 i_0$	$f_0 i_1$	$f_0 i_2$	$f_1 i_0$	$f_1 i_1$	$f_1 i_2$
$t_0$		0.08	0.21	0.33	0.02	0.09	0.27
$t_1$		0.065	0.105	0.06	0.035	0.195	0.54

La somma delle due righe della tabella fa uno.

### Variabile E

parent: PFC

nodo: PFCE

$$t_{PFCE} = P(E | P, F, C) P(C) = P(E, C | P, F)$$

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(P, V   T)$	$t_{TP} = 1$
$t_{TIF} = P(F, I   T)$	$t_{TF} = 1$
$t_{TPF} = P(T)$	$t_{PF} = 1$
$t_{PFCE} = P(E, C   P, F)$	

		$P(E, C   P, F)$			
P	F	$e_0 c_0$	$e_0 c_1$	$e_1 c_0$	$e_1 c_1$
$p_0$	$f_0$	0.475	0.45	0.025	0.05
$p_0$	$f_1$	0.45	0.4	0.05	0.1
$p_1$	$f_0$	0.15	0.125	0.35	0.375
$p_1$	$f_1$	0.1	0.05	0.4	0.45
$p_2$	$f_0$	0.075	0.05	0.425	0.45
$p_2$	$f_1$	0.025	0.015	0.475	0.485

La somma delle sei righe della tabella fa uno.

Come si vede, tutte le tabelle dei separatori sono rimaste inizializzate ad uno: questo perché, quando iniziamo a muovere informazione sul junction tree, il prodotto di tutte le tabelle dei nodi diviso il prodotto di tutte le tabelle dei separatori è invariante. Infatti:

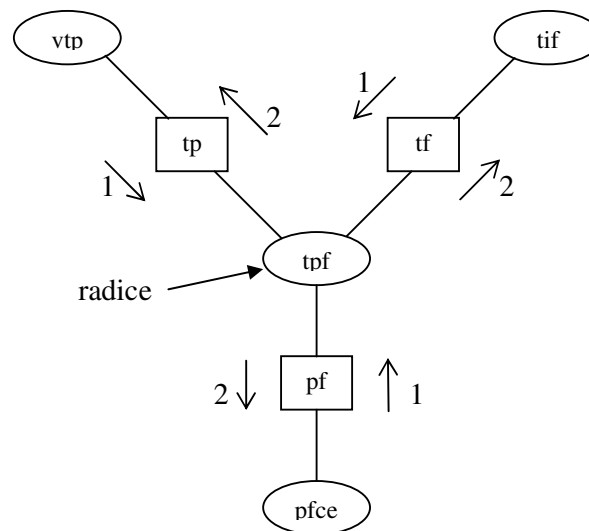
$$\begin{aligned}
 P(U) &= P(V, T, I, C, P, F, E) \\
 &= (t_{VTP} t_{TIF} t_{TPF} t_{PFCE}) / (t_{TP} t_{TF} t_{PF}) \\
 &= P(V) P(T) P(I) P(C) P(P | V, T) P(F | T, I) P(E | P, F, C) \\
 &= P(P, V | T) P(T) P(F, I | T) P(E, C | P, F).
 \end{aligned}$$

## 2.3 Consistenza nel Junction Tree

Dato il precedente junction tree con tutte le tabelle di probabilità dei nodi e dei separatori, il passo successivo consiste nel renderlo globalmente consistente. In un junction tree la consistenza locale implica la consistenza globale. Un junction tree si dice localmente consistente se per ogni coppia di nodi contigui  $V, W$  si verifica che:

$$\sum_{V \setminus S} t_V = t_S = \sum_{W \setminus S} t_W$$

dove  $S$  rappresenta il nodo separatore contenente l'intersezione  $V \cap W$ ; ciò significa che  $V$  e  $W$  contengono la stessa informazione su  $S$  e il link tra i due nodi è detto consistente. Per rendere consistente un link, occorre far transitare un messaggio in entrambe le direzioni. Si usa il seguente schema di trasmissione di messaggi:



con le seguenti tabelle di probabilità dei nodi e dei separatori:

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(P, V   T)$	$t_{TP} = 1$
$t_{TIF} = P(F, I   T)$	$t_{TF} = 1$
$t_{TPF} = P(T)$	$t_{PF} = 1$
$t_{PFCE} = P(E, C   P, F)$	

Come si vede, ogni link è percorso in entrambe le direzioni e questo ne garantisce la consistenza.

### 2.3.1 Trasmissione del messaggio da VTP a TPF

- 1)  $t_{TP}^* = \sum_V P(P, V | T) = P(P | T)$
- 2)  $t_{TPF}^* = P(T) P(P | T) = P(T, P)$

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(P, V   T)$	$t_{TP} = P(P   T)$
$t_{TIF} = P(F, I   T)$	$t_{TF} = 1$
$t_{TPF} = P(T, P)$	$t_{PF} = 1$
$t_{PFCE} = P(E, C   P, F)$	

		$P(P   T)$		
$T$		$P_0$	$P_1$	$P_2$
$t_0$		0.575	0.26	0.165
$t_1$		0.104	0.216	0.68

		$P(T, P)$		
$T$		$P_0$	$P_1$	$P_2$
$t_0$		0.1725	0.078	0.0495
$t_1$		0.0728	0.1512	0.476

### 2.3.2 Trasmissione del messaggio da TIF a TPF

$$1) t_{TF}^* = \sum_I P(F, I | T) = P(F | T)$$

$$2) t_{TPF}^* = P(T, P) P(F | T) = P(P | T) P(T) P(F | T) = P(T, P, F)$$

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(P, V   T)$	$t_{TP} = P(P   T)$
$t_{TIF} = P(F, I   T)$	$t_{TF} = P(F   T)$
$t_{TPF} = P(T, P, F)$	$t_{PF} = 1$
$t_{PFCE} = P(E, C   P, F)$	

		$P(F   T)$	
$T$		$f_0$	$f_1$
$t_0$		0.62	0.38
$t_1$		0.23	0.77

		$P(T, P, F)$		
$T$	$F$	$P_0$	$P_1$	$P_2$
$t_0$	$f_0$	0.10695	0.04836	0.03069
$t_0$	$f_1$	0.06555	0.02964	0.01881
$t_1$	$f_0$	0.016744	0.034776	0.10948
$t_1$	$f_1$	0.056056	0.116424	0.36652

### 2.3.3 Trasmissione del messaggio da PFCE a TPF

$$1) t_{PF}^* = \sum_{EC} P(E, C | P, F) = 1$$

$$2) t_{TPF}^* = P(T, P, F)$$

non cambia nessuna tabella

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(P, V   T)$	$t_{TP} = P(P   T)$
$t_{TIF} = P(F, I   T)$	$t_{TF} = P(F   T)$
$t_{TPF} = P(T, P, F)$	$t_{PF} = 1$
$t_{PFCE} = P(E, C   P, F)$	

### 2.3.4 Trasmissione del messaggio da TPF a VTP

$$1) t_{TP}^* = \sum_F P(T, P, F) = P(T, P)$$

$$2) t_{VTP}^* = P(P, V | T) P(T, P) / P(P | T) = P(P, V | T) P(P | T) P(T) / P(P | T) = P(P, V | T) P(T) = P(V, T, P)$$

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(V, T, P)$	$t_{TP} = P(T, P)$
$t_{TIF} = P(F, I   T)$	$t_{TF} = P(F   T)$
$t_{TPF} = P(T, P, F)$	$t_{PF} = 1$
$t_{PFCE} = P(E, C   P, F)$	

	$P(T, P)$		
$T$	$P_0$	$P_1$	$P_2$
$t_0$	0.1725	0.078	0.0495
$t_1$	0.0728	0.1512	0.476

		$P(V, T, P)$		
$T$	$V$	$P_0$	$P_1$	$P_2$
$t_0$	$v_0$	0.0255	0.003	0.0015
$t_0$	$v_1$	0.042	0.012	0.006
$t_0$	$v_2$	0.105	0.063	0.042
$t_1$	$v_0$	0.035	0.028	0.007
$t_1$	$v_1$	0.028	0.084	0.028
$t_1$	$v_2$	0.0098	0.0392	0.441

### 2.3.5 Trasmissione del messaggio da TPF a TIF

$$1) t_{TF}^* = \sum_P P(T, P, F) = P(T, F)$$

$$2) t_{TIF}^* = P(F, I | T) P(T, F) P(F | T) = P(F, I | T) P(F | T) P(T) / P(F | T) = P(F, I | T) P(T) = P(T, I, F)$$

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(V, T, P)$	$t_{TP} = P(T, P)$
$t_{TIF} = P(T, I, F)$	$t_{TF} = P(T, F)$
$t_{TPF} = P(T, P, F)$	$t_{PF} = 1$
$t_{PFCE} = P(E, C   P, F)$	

$P(T, F)$		
$T$	$f_0$	$f_1$
$t_0$	0.186	0.114
$t_1$	0.161	0.539

$P(T, I, F)$				
$T$	$F$	$i_0$	$i_1$	$i_2$
$t_0$	$f_0$	0.024	0.063	0.099
$t_0$	$f_1$	0.006	0.027	0.081
$t_1$	$f_0$	0.0455	0.0735	0.042
$t_1$	$f_1$	0.0245	0.1365	0.378

### 2.3.6 Trasmissione del messaggio da TPF a PFCE

$$1) t_{PF}^* = \sum_T P(T, P, F) = P(P, F)$$

$$2) t_{PFCE}^* = P(E, C | P, F) P(P, F) = P(P, F, C, E)$$

tabelle dei nodi		tabelle dei separatori	
$t_{VTP}$	$= P(V, T, P)$	$t_{TP}$	$= P(T, P)$
$t_{TIF}$	$= P(T, I, F)$	$t_{TF}$	$= P(T, F)$
$t_{TPF}$	$= P(T, P, F)$	$t_{PF}$	$= P(P, F)$
$t_{PFCE}$	$= P(P, F, C, E)$		

$P(P, F)$		
$P$	$f_0$	$f_1$
$p_0$	0.123694	0.121606
$p_1$	0.083136	0.146064
$p_2$	0.14017	0.38533

$P(P, F, C, E)$					
$P$	$F$	$e_0 c_0$	$e_0 c_1$	$e_1 c_0$	$e_1 c_1$
$p_0$	$f_0$	0.05875465	0.0556623	0.00309235	0.0061847
$p_0$	$f_1$	0.0547227	0.0486424	0.0060803	0.0121606
$p_1$	$f_0$	0.0124704	0.010392	0.0290976	0.031176
$p_1$	$f_1$	0.0146064	0.0073032	0.0584256	0.0657288
$p_2$	$f_0$	0.01051275	0.0070085	0.05957225	0.0630765
$p_2$	$f_1$	0.00963325	0.00577995	0.18303175	0.18688505

**Osservazione:** ogni tabella di ogni nodo e di ogni separatore contiene la probabilità congiunta delle variabili che si trovano nel nodo e nel separatore.

# Inferenza

Fare inferenza su una rete bayesiana significa dare evidenza ad alcune variabili ed osservare come si modificano le rimanenti. Si possono fare tre tipi di inferenza: il ragionamento causale in cui, data la causa, si cerca la probabilità dell'effetto, il ragionamento diagnostico in cui, dato un effetto e sapendo che può essere provocato da una certa causa, si cerca di determinare la probabilità di tale causa e l'explaining away in cui, dato un certo effetto e sapendo che può essere provocato da due cause, il manifestarsi di una rende meno probabile l'altra. In generale la procedura di inferenza è esponenziale: se  $m$  è il numero di stati e  $k$  il numero di variabili, la complessità è  $O(m^k)$ .

## 3.1 Inferenza senza evidenza

Dato il precedente junction tree consistente senza alcuna evidenza, si calcolano le seguenti probabilità non presenti tra le probabilità assegnate alla rete bayesiana.

$$P(P) = \sum_T P(T, P) = (p_0 = 0.2453 \quad p_1 = 0.2292 \quad p_2 = 0.5255)$$

$$P(F) = \sum_T P(T, F) = (f_0 = 0.347 \quad f_1 = 0.653)$$

$$P(E) = \sum_{PFC} P(P, F, C, E) = (e_0 = 0.2954885 \quad e_1 = 0.7045115)$$

Dopo aver calcolato queste tre probabilità, si hanno tutte le probabilità delle sette variabili: è interessante vedere come variano tali valori quando viene inserita l'evidenza su alcune variabili.

## 3.2 Inserimento dell'evidenza

Dato il precedente junction tree consistente, si dà evidenza ad alcune variabili e si vuole vedere come variano le probabilità delle altre. Innanzitutto premettiamo la seguente definizione:

**Definizione:** Un finding su una variabile aleatoria  $A$  con  $n$  stati è un vettore  $n$ -dimensionale di zero e di uno: indica quali stati di una variabile sono impossibili.

C'è un teorema che gestisce l'inserimento dell'evidenza:

**Teorema:** Sia BN una rete bayesiana rappresentante  $P(U)$  e sia T il junction tree corrispondente a BN. Sia  $e = \{f_1, \dots, f_m\}$  il finding sulle variabili  $\{A_1, \dots, A_m\}$ . Per ogni  $i$  trovare un nodo contenente  $A_i$  e si moltiplichi la sua tabella per  $f_i$ . Dopo un trasferimento completo di messaggi si ha:

$$t_V = P(V, e) \quad t_S = P(S, e) \quad P(e) = \sum_V t_V \quad P(W | e) = P(W, e) / P(e)$$

per ogni nodo V e per ogni separatore S (W indica V oppure S).

### 3.2.1 Assegnazione dell'evidenza

Si assegna evidenza alla variabile aleatoria  $E$  che indica superamento o meno dell'esame: tale variabile contiene i due stati no =  $e_0$  e si =  $e_1$ . Il finding su tale variabile è  $e = (0, 1)$  con  $P(e) = 0.7045115$ . L'unico nodo del junction tree che contiene la variabile  $E$  è PFCE, quindi si moltiplica la tabella di PFCE per  $e$ , ottenendo la seguente tabella:

		$P(P, F, C, E, e)$			
P	F	$e_0 c_0$	$e_0 c_1$	$e_1 c_0$	$e_1 c_1$
$p_0$	$f_0$	0	0	0.00309235	0.0061847
$p_0$	$f_1$	0	0	0.0060803	0.0121606
$p_1$	$f_0$	0	0	0.0290976	0.031176
$p_1$	$f_1$	0	0	0.0584256	0.0657288
$p_2$	$f_0$	0	0	0.05957225	0.0630765
$p_2$	$f_1$	0	0	0.18303175	0.18688505

Se i valori della tabella vengono divisi per  $P(e)$ , la somma fa di nuovo uno.

### 3.3 Propagazione dell'evidenza

Dopo aver dato evidenza ai nodi del junction tree, tale evidenza deve essere propagata agli altri nodi. Per fare questo si usa lo schema di propagazione di Hugin che consiste nei seguenti passi:

- 1) Si inserisce l'evidenza all'interno di un nodo a seconda della variabile scelta
- 2) Si seleziona un nodo radice R
- 3) Si esegue Collect Evidence a partire dalla radice R
- 4) Si esegue Distribute Evidence a partire dalla radice R
- 5) Si effettua una normalizzazione delle tabelle dividendo per  $P(e)$ .

#### 3.3.1 Trasmissione del messaggio da VTP a TPF

$$1) t_{TP}^* = \sum_V P(V, T, P) = P(T, P)$$

$$2) t_{TPF}^* = P(T, P, F)$$



non cambia nessuna tabella

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(V, T, P)$	$t_{TP} = P(T, P)$
$t_{TIF} = P(T, I, F)$	$t_{TF} = P(T, F)$
$t_{TPF} = P(T, P, F)$	$t_{PF} = P(P, F)$
$t_{PFCE} = P(P, F, C, E, e)$	

### 3.3.2 Trasmissione del messaggio da TIF a TPF

$$1) t_{TF}^* = \sum_I P(T, I, F) = P(T, F)$$

$$2) t_{TPF}^* = P(T, P, F)$$

non cambia nessuna tabella

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(V, T, P)$	$t_{TP} = P(T, P)$
$t_{TIF} = P(T, I, F)$	$t_{TF} = P(T, F)$
$t_{TPF} = P(T, P, F)$	$t_{PF} = P(P, F)$
$t_{PFCE} = P(P, F, C, E, e)$	

### 3.3.3 Trasmissione del messaggio da PFCE a TPF

$$1) t_{PF}^* = \sum_{EC} P(P, F, C, E, e) = P(P, F, e)$$

$$2) t_{TPF}^* = P(T, P, F) P(P, F, e) / P(P, F) = P(T \setminus P, F) P(P, F) P(P, F, e) / P(P, F) \\ = P(T \setminus P, F) P(P, F, e) = P(T, P, F, e)$$

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(V, T, P)$	$t_{TP} = P(T, P)$
$t_{TIF} = P(T, I, F)$	$t_{TF} = P(T, F)$
$t_{TPF} = P(T, P, F, e)$	$t_{PF} = P(P, F, e)$
$t_{PFCE} = P(P, F, C, E, e)$	

	$P(P, F, e)$	
$P$	$f_0$	$f_1$
$p_0$	0.00927705	0.0182409
$p_1$	0.0602736	0.1241544
$p_2$	0.12264875	0.3699168

		$P(T, P, F, e)$		
$T$	$F$	$p_0$	$p_1$	$p_2$
$t_0$	$f_0$	0.00802125	0.035061	0.02685375
$t_0$	$f_1$	0.0098325	0.025194	0.0180576
$t_1$	$f_0$	0.0012558	0.0252126	0.095795
$t_1$	$f_1$	0.0084084	0.0989604	0.3518592

### 3.3.4 Trasmissione del messaggio da TPF a VTP

$$1) t_{TP}^* = \sum_F P(T, P, F, e) = P(T, P, e)$$

$$2) t_{VTP}^* = P(V, T, P) P(T, P, e) / P(T, P) = P(V | T, P) P(T, P) P(T, P, e) / P(T, P) \\ = P(V | T, P) P(T, P, e) = P(V, T, P, e)$$

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(V, T, P, e)$	$t_{TP} = P(T, P, e)$
$t_{TIF} = P(T, I, F)$	$t_{TF} = P(T, F)$
$t_{TPF} = P(T, P, F, e)$	$t_{PF} = P(P, F, e)$
$t_{PFCE} = P(P, F, C, E, e)$	

$P(T, P, e)$			
$T$	$P_0$	$P_1$	$P_2$
$t_0$	0.01785375	0.060255	0.04491135
$t_1$	0.0096642	0.124173	0.4476542

$P(V, T, P, e)$				
$T$	$V$	$P_0$	$P_1$	$P_2$
$t_0$	$v_0$	0.00263925	0.0023175	0.00136095
$t_0$	$v_1$	0.004347	0.00927	0.0054438
$t_0$	$v_2$	0.0108675	0.0486675	0.0381066
$t_1$	$v_0$	0.00464625	0.022995	0.00658315
$t_1$	$v_1$	0.003717	0.068985	0.0263326
$t_1$	$v_2$	0.00130095	0.032193	0.41473845

### 3.3.5 Trasmissione del messaggio da TPF a TIF

$$1) t_{TF}^* = \sum_P P(T, P, F, e) = P(T, F, e)$$

$$2) t_{TIF}^* = P(T, I, F) P(T, F, e) / P(T, F) = P(I | T, F) P(T, F) P(T, F, e) / P(T, F) = \\ = P(I | T, F) P(T, F, e) = P(T, I, F, e)$$

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(V, T, P, e)$	$t_{TP} = P(T, P, e)$
$t_{TIF} = P(T, I, F, e)$	$t_{TF} = P(T, F, e)$
$t_{TPF} = P(T, P, F, e)$	$t_{PF} = P(P, F, e)$
$t_{PFCE} = P(P, F, C, E, e)$	

$P(T, F, e)$		
$T$	$f_0$	$f_1$
$t_0$	0.069936	0.0530841
$t_1$	0.1222634	0.459228

		$P(T, I, F, e)$		
$T$	$F$	$i_0$	$i_1$	$i_2$
$t_0$	$f_0$	0.009024	0.023688	0.037224
$t_0$	$f_1$	0.0027939	0.01257255	0.03771765
$t_1$	$f_0$	0.0345527	0.0558159	0.0318948
$t_1$	$f_1$	0.020874	0.116298	0.322056

### 3.3.6 Trasmissione del messaggio da TPF a PFCE

$$1) t_{PF}^* = \sum_T P(T, P, F, e) = P(P, F, e)$$

$$2) t_{PFCE}^* = P(P, F, C, E, e)$$

non cambia nessuna tabella

tabelle dei nodi	tabelle dei separatori
$t_{VTP} = P(V, T, P, e)$	$t_{TP} = P(T, P, e)$
$t_{TIF} = P(T, I, F, e)$	$t_{TF} = P(T, F, e)$
$t_{TPF} = P(T, P, F, e)$	$t_{PF} = P(P, F, e)$
$t_{PFCE} = P(P, F, C, E, e)$	

### 3.4 Inferenza con evidenza

Si calcolano adesso le probabilità calcolate precedentemente per vedere quali variazioni hanno subito in seguito all'inserimento dell'evidenza.

$$P(P, e) = \sum_T P(T, P, e) = (p_0 = 0.02751795 \quad p_1 = 0.184428 \quad p_2 = 0.49256555)$$

$$P(P | e) = P(P, e) / P(e) = (p_0 = 0.039059618 \quad p_1 = 0.26178139 \quad p_2 = 0.699158992)$$

$$P(F, e) = \sum_T P(T, F, e) = (f_0 = 0.1921994 \quad f_1 = 0.5123121)$$

$$P(F | e) = P(F, e) / P(e) = (f_0 = 0.272812296 \quad f_1 = 0.727187704)$$

$$P(E, e) = \sum_{PFC} P(P, F, C, E, e) = (e_0 = 0 \quad e_1 = 0.7045115)$$

$$P(E | e) = P(E, e) / P(e) = (e_0 = 0 \quad e_1 = 1).$$

Come si può vedere, le probabilità di essere preparato e di aver frequentato il corso, supposto di aver superato l'esame, sono aumentate. Infatti le probabilità senza evidenza erano le seguenti:

$$P(P) = (p_0 = 0.2453 \quad p_1 = 0.2292 \quad p_2 = 0.5255)$$

$$P(F) = (f_0 = 0.347 \quad f_1 = 0.653)$$

$$P(E) = (e_0 = 0.2954885 \quad e_1 = 0.7045115).$$

Vediamo anche come si sono modificate le probabilità a priori delle variabili  $V$ ,  $T$ ,  $I$  e  $C$ :

$$P(V, e) = \sum_{TP} P(V, T, P, e) = (v_0 = 0.0405421 \ v_1 = 0.1180954 \ v_2 = 0.545874)$$

$$P(V | e) = P(V, e) / P(e) = (v_0 = 0.0575464 \ v_1 = 0.167627356 \ v_2 = 0.774826244)$$

$$P(T, e) = \sum_F P(T, F, e) = (t_0 = 0.1230201 \ t_1 = 0.5814914)$$

$$P(T | e) = P(T, e) / P(e) = (t_0 = 0.17461759 \ t_1 = 0.82538241)$$

$$P(I, e) = \sum_{TF} P(T, I, F, e) = (i_0 = 0.0672446 \ i_1 = 0.20837445 \ i_2 = 0.42889245)$$

$$P(I | e) = P(I, e) / P(e) = (i_0 = 0.095448549 \ i_1 = 0.295771538 \ i_2 = 0.608779913)$$

$$P(C, e) = \sum_{PFE} P(P, F, C, E, e) = (c_0 = 0.33929985 \ c_1 = 0.36521165)$$

$$P(C | e) = P(C, e) / P(e) = (c_0 = 0.481610094 \ c_1 = 0.518389906)$$

Si nota che i valori si sono modificati nel senso che si poteva facilmente prevedere. Le probabilità a priori erano infatti le seguenti:

$$P(V) = (v_0 = 0.1 \ v_1 = 0.2 \ v_2 = 0.7)$$

$$P(T) = (t_0 = 0.3 \ t_1 = 0.7)$$

$$P(I) = (i_0 = 0.1 \ i_1 = 0.3 \ i_2 = 0.6)$$

$$P(C) = (c_0 = 0.5 \ c_1 = 0.5).$$