ESEMPIO DIAGNOSI

REGOLE "DIAGNOSTICHE"

$$\forall p \; Sintomo(p, MalDiDenti) \Rightarrow Malattia(p, Carie)$$

(SBAGLIATA!)

$$\forall p \; Sintomo(p, MalDiDenti) \Rightarrow Malattia(p, Carie) \lor Malattia(p, Gengive) \lor Malattia(p, Piorrea) \lor \ldots$$

(TROPPO LUNGA!)

REGOLA CAUSALE

$$\forall p \; Malattia(p, carie) \Rightarrow Sintomo(p, MalDiDenti)$$

(SBAGLIATA)

$\Rightarrow$ LA LOGICA DEL I ORDINE

1 NON È SUFFICIENTE (ADEGUATA)

PER ESPRIMERE INCERTEZZA

DISTRIBUZIONE DI PROB. CONGIUNTA

SPECIFICA COMPLETAMENTE LE ASSEGNAZIONI DI PROBABILITÀ PER TUTTE LE PROPOSIZIONI (VALORI DELLE VAR CASUALI) NEL DOMINIO DI UN AGENTE

EVENTO ATOMICO: È UNA ASSEGNAZIONE DI VALORI PARTICOLARI A TUTTE LE VARIABILI DEL DOMINIO:

$$\{X_1, X_2, \dots, X_n\}$$

$$Dom(X_1) = Dom(X_2) = \dots = Dom(X_n) = \{a, b, c\}$$

$$e_1 = \begin{array}{l} X_1 = a \\ X_2 = b \\ X_3 = a \\ \vdots \\ X_n = c \end{array}$$

e_1 = EVENTO ATOMICO

LA DISTRIBUZIONE DI PROBABILITÀ CONGIUNTA $\bar{P}(X_1, X_2, \dots, X_n)$ ASSEGNA PROBABILITÀ A TUTTI GLI EVENTI:

NELL'ESEMPIO $\bar{P}(X_1, X_2, \dots, X_n)$ è una matrice $3 \times 3 \times \dots \times 3$ [n volte]

"MARGINALIZZAZIONE"

SIA $B_1, B_2, \dots, B_n$ UN INSIEME DI PROPOSIZIONI <u>ESAUSTIVE</u> E <u>MUTUAMENTE ESCLUSIVE</u>,

$$P(A) = \sum_{i=1}^{n} P(A\ B_i) =$$

$$= \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

<u>NOTAZIONE</u> : $P(A \wedge B) = P(A,B) = P(A\,B)$

INDIPENDENZA

DUE EVENTI A e B SI DICONO **INDIPENDENTI** $(A \perp B)$ SE:

$$P(A|B) = P(A)$$

O, EQUIVALENTEMENTE, SE

$$P(A,B) = P(A) \cdot P(B)$$

DUE EVENTI A e B SI DICONO **INDIPENDENTI DATO IL CONTESTO C** $(A \perp B | C)$, SE

$$P(A|B,C) = P(A|C)$$

B è irrilevante a predire A dato C

QUESTA FORMA DI INDIPENDENZA SI CHIAMA **INDIPENDENZA CONDIZIONALE**

TEOREMA DI BAYES

PER LA REGOLA DEL PRODOTTO:

$$P(A,B) = P(A|B)\cdot P(B) = P(B|A)\cdot P(A)$$

$\Downarrow$

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

TEOREMA DI BAYES

estensione a due evidenze ("prove")

$$P(A|BC) = \frac{P(B|AC)\cdot P(A|C)}{P(B|C)}$$

TEOREMA DI BAYES PER VARIABILI CAUSALI

$$\bar{P}(x|y) = \frac{\bar{P}(y|x)\cdot \bar{P}(x)}{\bar{P}(y)}$$

$$\bar{P}(x/zy) = \frac{\bar{P}(z/xy)\cdot \bar{P}(x/y)}{\bar{P}(z/y)}$$

NORMALIZZAZIONE

$$P(M|S) \overset{\text{BAYES}}{=} \frac{P(S|M) \cdot P(M)}{P(S)}$$

$$P(\neg M|S) \overset{\text{BAYES}}{=} \frac{P(S|\neg M)\ P(\neg M)}{P(S)}$$

$$P(M|S) + P(\neg M|S) = 1 \qquad (*)$$

$\Rightarrow$

$$P(S) = P(S|M) \cdot P(M) + P(S|\neg M) \cdot P(\neg M)$$

$P(S)$ PUÒ ESSERE CALCOLATO USANDO $P(S|M)$, $P(S|\neg M)$, $P(M)$ e $P(\neg M)$

$P(S)$ È UNA COSTANTE DI NORMALIZZAZIONE

$$P(x|y) = \alpha\, P(y|x) \cdot P(x)$$

COMBINAZIONE DI EVIDENZE (PROVE)

ESEMPIO

C = CARIATO

E_1 = evidenza maldidenti $\quad E_2$ = evidenza dolore se colpito

$$P(C|E_1E_2) = \alpha \, P(\underline{E_1E_2|C}) \cdot P(C) \quad (0)$$

DIFFICOLTÀ DI ACQUISIRE I DATI NECESSARI

$$P(C|E_1) = \alpha_1 \, P(C) \cdot P(E_1|C) \quad (1)$$

$$P(C|E_2E_1) = \alpha_2 \, P(C|E_1) \cdot P(E_2|CE_1) \quad (2)$$

<u>AGGIORNAMENTO BAYESIANO</u>:

$P(C|E_2E_1)$ È CALCOLATO INCREMENTALMENTE CONSIDERANDO E_1 e E_2 IN SEQUENZA

E_1 È IL "CONTESTO" ENTRO CUI VALUTARE

$P(C|E_2E_1)$ (probabilità di C se osservo E_2 dopo aver già osservato E_1)

COMBINAZIONE DI EVIDENZE (CONT.)

SE HO UNA SEQUENZA DI EVIDENZE $E_1 E_2 \cdots E_n$:

$$P(H \mid E E_1 E_2 \cdots E_n) = \alpha P(E \mid H E_1 \cdots E_n) \cdot P(H \mid E_1 E_2 \cdots E_n) \qquad (4)$$

(4) PUÒ ESSERE MOLTO PIÙ CONVENIENTE DI ([illegible]) PERCHÉ CONSENTE DI SFRUTTARE IPOTESI DI <u>INDIPENDENZA CONDIZIONALE</u> TRA LE EVIDENZE.

SE $P(E \mid H E_1 E_2 \cdots E_n) = P(E \mid H)$ ALLORA

$$P(H \mid E E_1 \cdots E_n) = \alpha\, P(E \mid H) \cdot P(H \mid E_1 \cdots E_n)$$

$\Rightarrow$ DRASTICA RIDUZIONE DEL NUMERO DI PROBABILITÀ CONDIZIONALI RICHIESTE PER APPLICARE IL TEOREMA DI BAYES

9) ES:

$$P(H \mid E_3 E_2 E_1) = \alpha P(E_3 \mid H E_2 E_1) \cdot P(H \mid E_2 E_1) =$$
$$P(E_3 \mid H) \cdot P(H \mid E_2 E_1) = \alpha' P(E_3 \mid H) \cdot P(E_2 \mid H E_1) \cdot P(H \mid E_1)$$
$$= \alpha' P(E_3 \mid H) \cdot P(E_2 \mid H) \cdot P(H \mid E_1) = \alpha'' P(E_3 \mid H) \cdot P(E_2 \mid H) \cdot P(E_1 \mid H) \cdot P$$

ESEMPIO USO IND. COND.

$$P(H|E_3E_2E_1) \overset{B}{=} \alpha\, P(E_3|HE_2E_1)\cdot P(H|E_2E_1) =$$

$$\overset{I}{=} \alpha\, \underline{P(E_3|H)}\cdot P(H|E_2E_1) =$$

$$\overset{B}{=} \alpha'\, P(E_3|H)\cdot \underline{P(E_2|HE_1)\cdot P(H|E_1)} =$$

$$\overset{I}{=} \alpha'\, P(E_3|H)\cdot \underline{P(E_2|H)}\cdot P(H|E_1) =$$

$$\overset{B}{=} \alpha''\, P(E_3|H)\cdot P(E_2|H)\cdot \underline{P(E_1|H)\cdot P(H)}.$$

B = APPLICAZIONE TEOREMA DI BAYES

I = APPLICAZIONE ASSUNZIONE INDIPENDENZA CONDIZIONALE DELLE EVIDENZE DATA LA CAUSA

ESEMPIO CARIE

$$P(E_2|CE_1) = P(E_2|C)$$
$$P(E_1|CE_2) = P(E_1|C)$$

$$P(C|E_2E_1) = \alpha_{12}\, P(C)\cdot P(E_2|C)\cdot P(E_1|C)$$

REGOLE CAUSALI $\rightleftharpoons$ PROBABILITÀ CONDIZIONALI

$P(E_i \mid H_j) \cong$ SE LA CAUSA H_j È VERA, ALLORA SI OSSERVERÀ E_i CON UNA CERTA CONFIDENZA, A PATTO CHE "TUTTO IL RESTO" SIA IRRILEVANTE

TEOREMA DI BAYES $\cong$ "MOTORE DI INFERENZA" PROBABILISTICO CHE CI DÀ INFORMAZIONI SUL VERSO OPPOSTO DELLE REGOLE CAUSALI: <u>CI DÀ INFORMAZIONI SULLE CAUSE A PARTIRE DAGLI EFFETTI (EVIDENZE)</u>

"RAGIONAMENTO EVIDENZIALE" (DIAGNOSTICO)

$\left.\begin{array}{l}\text{CAUSE} \cong \text{IPOTESI} \\ \text{EFFETTI} \cong \text{EVIDENZE}\end{array}\right\}$ VARIABILI CASUALI

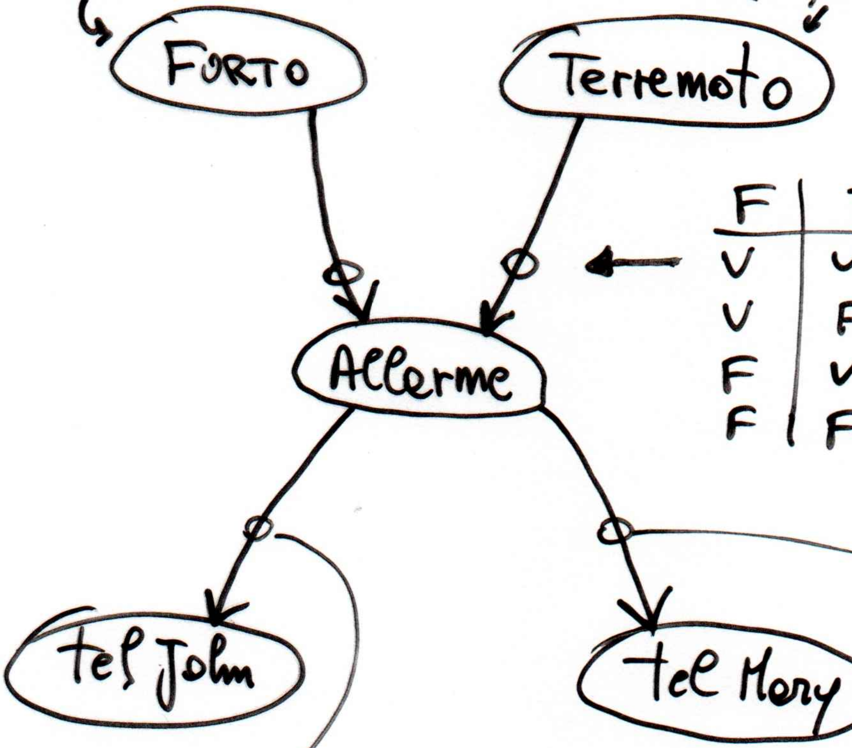
RETI BAYESIANE : ESEMPIO ALLARME

FURTO   TERREMOTO   ALLARME
TEL JOHN   TEL MARY

} VARIABILI CAUSALI / CASUAL DEL DOMINIO

$P(F) = 0.001$

$P(T) = 0.002$



F	T	P(A|FT)	P(¬A|FT)
V	V	0.95	1 − P(A|FT)
V	F	0.94	
F	V	0.29	
F	F	0.001	

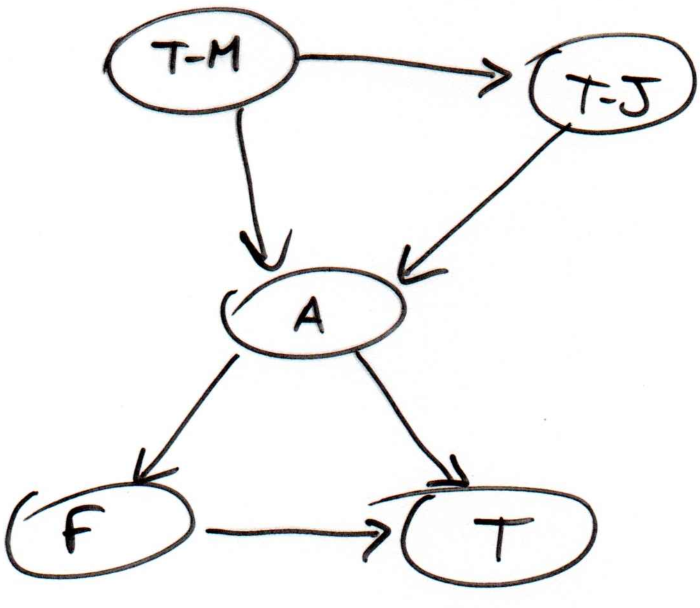
A	P(J|A)
V	0.90
F	0.10

A	P(M|A)
V	0.70
F	0.30

ORDINAMENTO DELLE VARIABILI

LA STRUTTURA TOPOLOGICA DELLA RETE (E QUINDI IL NUMERO DI VALORI RICHIESTI PER LE PROBABILITÀ CONDIZIONALI) DIPENDE DALL'ORDINE SCELTO

ES: Tel-Mary, Tel-John, Allarme, Furto, Terr.



2 ARCHI IN PIÙ
3 NUMERI IN PIÙ DA SPECIFICARE

NEL CASO + SFORTUNATO LA RETE È "COMPLETAMENTE CONNESSA"

<u>REGOLA EURISTICA</u>: SCEGLIERE L'ORDINE IN MODO CHE LE VARIABILI "CAUSA" SIANO PRIMA DELLE RELATIVE VARIABILI "EFFETTO".

COMPATTEZZA DI UNA RB

n VARIABILI

K = numero max di variabili che possono influenzare una var. ($\leq n$)

$\Downarrow$

numero di probabilità da specificare: $O(n \cdot 2^k)$

(se le variabili sono booleane)

PUÒ ESSERE UN "RISPARMIO" NOTEVOLE

$n = 20 \qquad K = 5 \qquad 2^n \Rightarrow 10^6 \qquad n * 2^K = 640$!!

SEMANTICA RETI BAYESIANE

UNA RB CON VARIABILI $V_1, V_2, V_3 \dots V_n$
RAPPRESENTA $P(V_1 V_2 \dots V_n)$

$$P(V_1 = x_1, V_2 = x_2, \dots, V_n = x_n) = P(x_1, x_2 \dots, x_n)$$

$$P(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} P(x_i \mid \{x_j \mid V_j \text{ GENITORE DI } V_i\})$$

ES ALLARME

$$P(T\text{-}J, T\text{-}M, A, \neg F, \neg T) =$$

$$P(T\text{-}J \mid A) \cdot P(T\text{-}M \mid A) \cdot P(A \mid \neg F, \neg T) \cdot P(\neg F) \cdot P(\neg T)$$

$$0.9 \qquad 0.7 \qquad 0.001 \qquad 0.999 \qquad 0.998$$

$= 0.00062$

LE MATRICI CONDIZIONALI DI UNA RB FORNISCONO UNA RAPPRESENTAZIONE DECOMPOSTA DELLA DISTRIBUZIONE DI PROBABILITÀ CONGIUNTA

LA FORMULA PRECEDENTE VALE PERCHÉ

$$P(x_1 \cdots x_n) = P(x_n | x_{n-1} \cdots x_1) \cdot P(x_{n-1} \cdots x_1) =$$

$$= P(x_n | x_{n-1} \cdots x_1) \cdot P(x_{n-1} | x_{n-2} \cdots x_1) \cdot$$

$$\cdots P(x_2 | x_1) \cdot P(x_1) =$$

$$= \prod_{i=1}^{n} P(x_i | x_{i-1} \cdots x_1) =$$

$$= \prod_{i=1}^{n} P\left(x_i \,\middle|\, \left\{ x_j \,\middle|\, \begin{array}{l} V_j \text{ È UN NODO} \\ \text{GENITORE DI } V_i \end{array} \right\}\right)$$

VALE A PATTO CHE LE ASSUNZIONI DI INDIPENDENZA CONDIZIONALE FATTE NELLA <u>STRUTTURA</u> DELLA RETE SONO EFFETTIVAMENTE VALIDE.

(DATE LE VARIABILI "PADRE" DI X_i, X_i È INDIPENDENTE DALLE <u>ALTRE</u> VARIABILI CHE LA PRECEDONO $[x_1 \cdots x_{i-1}]$)

COSTRUZIONE DI UNA RETE

(1) SCEGLIERE UN <u>INSIEME V</u> DI VARIABILI RILEVANTI DEL DOMINIO

(2) SCEGLIERE UN <u>ORDINE</u> PER LE VARIABILI IN V

(3) Finché ci sono Var. in S $\left(\begin{array}{l} S = V \\ \text{inizialmente} \end{array} \right.$

$\boxed{A}$ PRENDI E TOGLI DA S UNA VAR V_i

$\boxed{B}$ FORMA UN NODO PER V_i NELLA RETE

$\boxed{C}$ CONNETTI GLI ALTRI NODI DELLA RETE CHE <u>INFLUENZANO DIRETTAMENTE</u> V_i A V_i

$\boxed{D}$ DEFINISCI LA TABELLA DI <u>PROBABILITÀ CONDIZIONALI</u> PER V_i DATE LE ALTRE VARIABILI CHE INFLUENZANO V_i

LA PROCEDURA GARANTISCE CHE NON SI FORMINO CICLI

<u>ESEMPIO ALLARME</u> $\left\{ \begin{array}{l} \text{FURTO, TERREMOTO, ALLARME,} \\ \text{TEL-JOHN, TEL-MARY} \end{array} \right\}$

INDIPENDENZA CONDIZIONALE

RICONOSCERE AUTOMATICAMENTE CHE UN CERTO INSIEME X DI VAR. È INDIPENDENTE DA UN ALTRO INSIEME Y DATO UN TERZO INSIEME S È IMPORTANTE PER

(1) COSTRUIRE ALGORITMI PIÙ EFFICIENTI DI INFERENZA PROBABILISTICA

(2) REALIZZARE UN SISTEMA DI RAGIONAMENTO QUALITATIVO SULLE INFLUENZE DELLE VARIABILI (PER RILEVANZA)

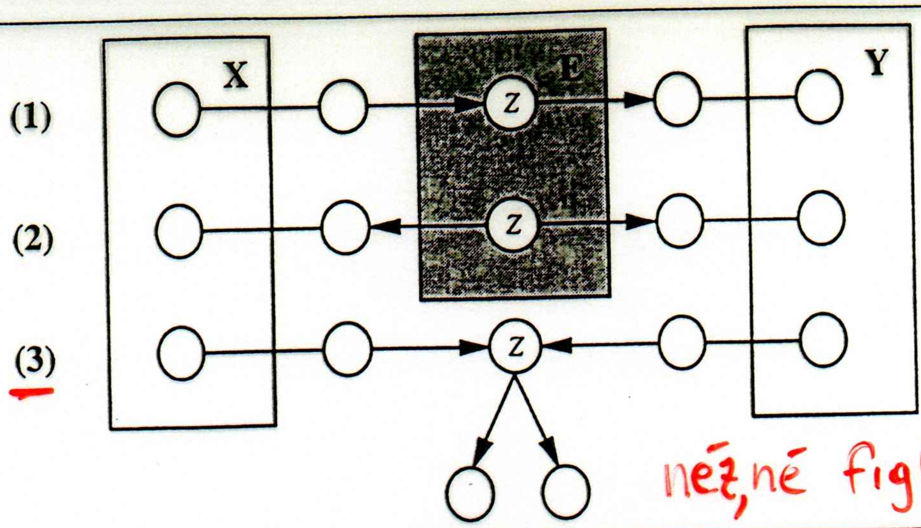
<u>TEOREMA</u>: SE OGNI CAMMINO <u>NON</u>-DIRETTO DA UN NODO IN X AD UN NODO IN Y È D-SEPARATO DA E, ALLORA X E Y SONO CONDIZIONALMENTE INDIPENDENTI DATO E

NOTA

S SPESSO INDICATO CON E, PERCHÈ RAPPRESENTI UN INSIEME DI <u>evidenze</u>.

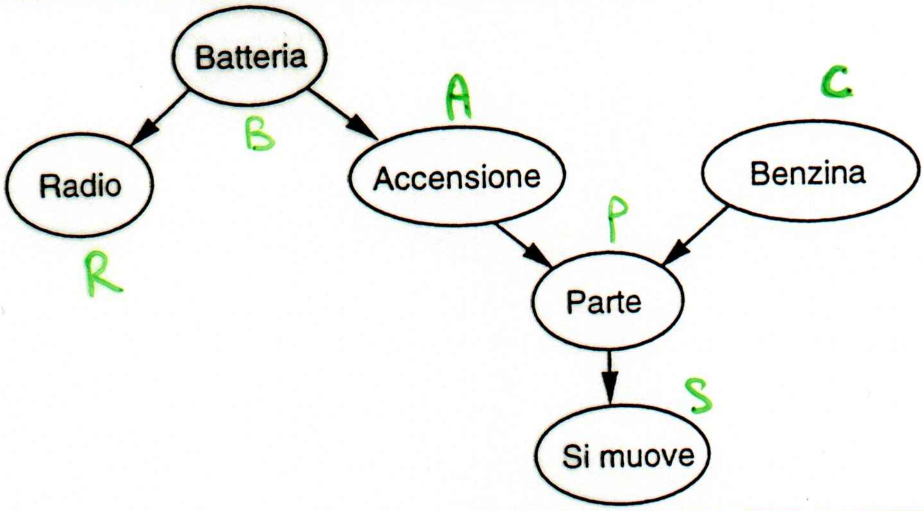
UN INSIEME DI NODI E D-SEPARA DUE INSIEMI X E Y SE OGNI CAMMINO NON DIRETTO DA UN NODO DI X (Y) AD UN NODO DI Y (X) È BLOCCATO DA E.

UN CAMMINO È BLOCCATO DA UN INSIEME DI NODI E SE ∃ UN NODO Z SUL CAMMINO PER IL QUALE VALE UNA DELLE SEGUENTI PROPRIETÀ:



né z, né figli(z) sono in E

Figura 15.4 Tre modi in cui è possibile bloccare un percorso da X a Y, data la prova E. Se ogni percorso da X a Y è bloccato, allora diciamo che E d-separa X e Y.



1) $C \perp R, A$

2) $C \perp R, B$

3) $C \perp R, \{\}$

C e R DIP SE CONOSCO P O S:

es: P = Falso

R = Vero

↓ C FALSO

Figura 15.5 Una rete di credenze che descrive alcune caratteristiche del sistema elettrico di una macchina e del suo motore.

15.3 INFERENZE NELLE RETI DI CREDENZE

Il compito fondamentale per ogni sistema di inferenza probabilistico è quello di calcolare la distribuzione di probabilità a posteriori per un insieme di **variabili di interrogazione**, dati i valori esatti per alcune **variabili di prova**. Ovverosia il sistema calcola $\mathbf{P}(Interrogazione|Prova)$. Nell'esempio dell'allarme, $Intruso$ è una variabile di interrogazione ovvia, $JohnTelefona$ e $MaryTelefona$ possono servire come variabili di prova. Le reti di credenze sono abbastanza flessibili in modo tale che ogni nodo può servire sia come variabile di prova che di interrogazione. Niente ci può impedire di chiedere $\mathbf{P}(Allarme|JohnTelefona, Terremoto)$, anche se sarebbe

$P(\neg Benzina \mid \neg parte) = \alpha$

$P(\neg Benzina \mid RadioOK, \neg parte) = \alpha'$ $\qquad \alpha' > \alpha$

D