

Laurea Specialistica in Informatica
a.a. 2005-2006

Interazione Uomo-Macchina II:

Interfacce Intelligenti

Fiorella de Rosis

1

Introduzione

Prima parte: Formalizzazione e Ragionamento

1.1. Ragionamento logico:

- Formalizzazione
- Risoluzione

1.2. **Ragionamento incerto**

- **Reti Causali Probabilistiche**
- Reti dinamiche
- Apprendimento di Reti

Seconda parte: Modelli di Utente

2.1. Modelli logici

2.2. Modelli con incertezza

Terza parte: Interazione in linguaggio naturale

3.1. Generazione di messaggi

- Introduzione
- Teorie
- Metodi

3.2. Comprensione di messaggi

Quarta parte: Simulazione di dialoghi

Programma
del Corso

Consideriamo di nuovo l'esempio Es 2.5

"Mangiare ad orari fissi **aiuta ad** evitare di saltare i pasti ..."

Nella Unità 2, avavamo formalizzato questa frase come:

$\forall x (Person(x) \wedge EatAtFixedTime(x)) \rightarrow AvoidJumpMeal(x)$

Ma la formalizzazione con linguaggio logico radicalizza l'affermazione

(non necessariamente si evita di saltare i pasti se si mangia in modo regolare).

E' necessario, in questo caso, 'rilassare' l'implicazione introducendo un elemento d'incertezza.

$\forall x (Person(x) \wedge EatAtFixedTime(x)) \rightarrow? AvoidJumpMeal(x)$

L'implicazione avrà, in questo modo, un *grado di forza*.

3

Consideriamo ora l'esempio Es 2.4

Nella Unità 3, abbiamo definito come 'certi' i dati:

Implies(a,g)

Likes(x,g)

CanDo(x,a)

mentre potrebbero avere un 'grado di certezza', così come l'implicazione:

$\forall x \forall a \forall g$

$((Implies(a,g) \wedge Likes(x,g) \wedge CanDo(x,a)) \rightarrow (ShouldDo(x,a)))$

Vediamo come si può formalizzare la conoscenza e come si può ragionare in condizioni di incertezza.

*Sceglieremo, in particolare, le misure di incertezza che si basano sulla **teoria della probabilità**.*

Cominciamo quindi con qualche chiarimento sul concetto di probabilità.

4

Diversi Concetti di Probabilità

1. Definizione 'classica' (Laplace, 1820)

Probabilità come rapporto fra il numero di casi favorevoli all'evento di cui si vuole misurare la probabilità e il n totale di casi possibili.

Es: probabilità che esca un picche in un mazzo di 52 carte è .25 (1/4)

5

Diversi Concetti di Probabilità

2. Idea 'frequentista' (Von Mises, 1928)

Chiamiamo 'collettivo' una sequenza di *eventi* E uniformi che differiscono per alcuni *attributi* osservabili. Effettuiamo un *esperimento* estraendo dal collettivo un individuo e osservando il valore dell'attributo.

Se l'esperimento è ripetuto n volte e $S_n(E)$ è il numero di volte in cui l'evento E si verifica, allora:

$$P(E) = \lim_{n \rightarrow \infty} S_n(E) / n$$

Al crescere di n (ripetizioni dell'esperimento), le frequenze si stabilizzano fornendo una stima sempre meno approssimata del valore di probabilità.

6

Diversi Concetti di Probabilità

2. Idea 'frequentista' (segue)

Tuttavia, in esperimenti reali non è possibile procedere all'infinito.

Von Mises sostiene che *la sua definizione è applicabile ogni volta che la frequenza dell'attributo osservato tenderebbe ad un limite ben definito se l'esperimento potesse continuare.*

Esempio:

Collettivo: insieme degli studenti della Laurea Specialistica in Informatica

Evento: agli studenti piace la musica di Bob Marley.

7

Limiti delle concezioni 'classiche'

Le due teorie viste finora sostengono che affermazioni come "Probabilmente, il Milan non vincerà il prossimo campionato" non rientrano nel dominio dei problemi trattati dalla teoria della probabilità.

Infatti: non si può calcolare il n. di casi favorevoli su quelli possibili; e non si può effettuare un esperimento ripetuto nelle stesse condizioni.

Tuttavia, il mondo delle scommesse assegna comunemente un valore a queste affermazioni.

Questo valore viene stabilito sulla base del *grado di fiducia* (degree of belief) sul verificarsi dell'evento in gioco:

metodo di misura tipicamente soggettivo!!

8

Un'idea (relativamente) nuova di probabilità

3. Probabilità soggettiva o 'neo-bayesiana' (De Finetti, 1960)

"Dato un numero aleatorio x , Tu sei obbligato a scegliere un valore x^ nell'intesa che, dopo tale scelta, sarai impegnato ad accettare qualunque scommessa con guadagno $c(x - x^*)$, con c qualunque (positivo o negativo) a scelta di un competitore. $P(x)$, previsione secondo la tua opinione, è per definizione il valore x^* che a tal fine avrai scelto."*

Si suppone che Tu non desideri stipulare scommesse che ti diano certamente una perdita.

Pertanto, un tuo insieme di previsioni si dice 'coerente' se, tra le combinazioni di scommesse che ti sei impegnato ad accettare, non ve ne sono che diano guadagni tutti uniformemente negativi"

9

Vediamo subito un esempio

Qual è la probabilità che la CdL vinca le elezioni?

Io penso che sia = .48

Questo significa che sono disponibile a pagare 48 euro, accettando di avere 100 euro se la CdL perde e 0 se vince.

Ma significa anche che sarei disponibile (per coerenza) a pagare 52 euro, accettando di avere 100 euro se la CdL vince e 0 se perde.

Complessivamente, avremo raccolto 100 euro.

Quali sono i guadagni?

Se la CdL vince, chi aveva pagato 52 euro ne guadagna 48; se perde, chi aveva pagato 48 euro ne guadagna 52.

10

La probabilità: chi è costei? B. de Finetti, (dall'Enciclopedia Einaudi)

Prima di rispondere a tale domanda è certamente opportuno chiedersi: ma davvero 'esiste' la probabilità? e cosa mai sarebbe? Io risponderei di no, che non esiste. Qualcuno, cui diedi questa risposta mi chiese ironicamente perché mai, allora, me ne occupo. Mah! Potrei anche dire, viceversa e senza contraddizione, che la probabilità regna ovunque, che è, o almeno dovrebbe essere, la nostra 'guida nel pensare e nell'agire', e che perciò mi interessa. Soltanto, mi sembra improprio, e perciò mi urta, vederla concretizzata in un sostantivo, 'probabilità', mentre riterrei meglio accettabile e più appropriato che si usasse soltanto l'aggettivo, 'probabile', o, meglio ancora, soltanto l'avverbio, 'probabilmente'. Dire che la probabilità di una certa asserzione vale 40 per cento appare- purtroppo!- come espressione concreta di una verità apodittica. Non pretendo né desidero che tale modo di esprimersi vada bandito, ma certo è che l'asserzione apparirebbe assai più appropriatamente formulata se la si ammorbidisse dicendo, invece, che quel fatto lo si giudica 'probabile al 40 per cento', o, meglio ancora (a parte che suona male), che ci si attende 'al 40 per cento-probabilmente' che sia o che risulti vero. Il guaio è che il realismo (come accuratamente osservò Jeffreys) ha il vantaggio che il linguaggio è stato creato da realisti, e per di più da realisti molto primitivi, ed è perciò che 'noi abbiamo larghissime possibilità di descrivere le proprietà attribuite agli oggetti, ma scarsissime di descrivere quelle direttamente conosciute come sensazioni'. Da ciò la mania (che forse per altri è invece indizio di saggezza, serietà, accuratezza) di assolutizzare, di concretizzare, di oggettivare perfino quelle che sono soltanto proprietà dei nostri atteggiamenti soggettivi. Non altrimenti si spiegherebbe lo sforzo di fare della Probabilità qualcosa di nobler than it is (sempre parole di Jeffreys), nascondendone la natura soggettiva e gabelandola per oggettiva. Secondo la spiritosa fantasia di Hans Freudenthal si tratterebbe di uno strano pudore per impedire di farci vedere la Probabilità 'come Dio l'ha fatta': occorre una 'foglia di fico', e spesso la si riveste tutta di foglie di fico rendendola addirittura invisibile o irrisconscibile.

Bruno de Finetti: Definizione soggettiva e operativa di probabilità.

The classic exposition of his distinctive theory is the 1937 "La prévision: ses lois logiques, ses sources subjectives," Annales de l'Institut Henri Poincaré, 7, 1-68, which discussed *probability founded on the coherence of betting odds and the consequences of exchangeability*.

De Finetti ha proposto il seguente esperimento:

You must set the price of a promise to pay \$1 if there was life on Mars 1 billion years ago, and \$0 if there was not, and tomorrow the answer will be revealed.

You know that your opponent will be able to choose either to buy such a promise from you at the price you have set, or require you to buy such a promise from your opponent, still at the same price. In other words: you set the odds, but your opponent decides which side of the bet will be yours.

The price you set is the "operational subjective probability" that you assign to the proposition on which you are betting. This price has to obey the probability axioms if you are not to face certain loss, as you would if you set a price above \$1 (or a negative price). Prices, or equivalently odds, that do not expose you to certain loss are called *coherent*.

Vedi anche a

<http://www.answers.com/topic/bruno-de-finetti> e inoltre a

<http://www.controluce.it/giornali/a13n09/20-cultura-brunodefinetti3.htm>

12

L'idea Soggettivista di probabilità rispetta le regole del Calcolo della Probabilità

Poiché l'ipotesi è che la persona sia *disponibile a stare da ciascuna delle due parti che partecipano alla scommessa*, dovrà accettare di essere:

quello che paga la somma $P(E)$ e
che riceve 1 o 0 a seconda che E si verifichi o no, o viceversa.

Questo vincolo di coerenza implica che, se vi sono n eventi E_i esaustivi e mutuamente esclusivi ed una persona assegna una probabilità $P(E_i)$ ad ognuno di essi, debba essere d'accordo sullo scambiare

$\sum_{i=1, \dots, n} P(E_i)$ unità contro 1 unità.

Un insieme di valori di probabilità che non consente scommesse certamente perdenti si dice coerente.

13

Frequentismo e soggettivismo 'non radicali'

Un frequentista 'non radicale' si rende conto di dover prendere decisioni anche in condizioni di incertezza non misurabili attraverso frequenze.

In questi casi, piuttosto che inventare argomenti frequentisti difficili da sostenere, riconoscerà che le sue stime di probabilità corrispondono semplicemente a dei 'degree of belief' e cercherà di rispettare il vincolo di coerenza dei soggettivisti, per utilizzare le regole del calcolo delle probabilità.

D'altra parte, anche un 'soggettivista non radicale' è convinto che assegnare all'uscita (al lotto) del 53 una prob di 1/90 sia più corretto che assegnarle il valore di 1/80 (a meno che non abbia ragioni diverse per crederlo: "*Me lo ha detto mio nonno in sogno, e mio nonno mi ha sempre dato numeri buoni...*").

14

Vediamo allora come rappresentare la conoscenza e come ragionare secondo la teoria bayesiana della probabilità.

Considereremo, fra i vari metodi proposti,
le *Reti Causali Probabilistiche*,
che sono oggi il formalismo prevalente nel campo delle
interfacce intelligenti.

(il loro uso è stato introdotto per la prima volta, in
questo dominio, da un Centro Ricerche della Microsoft,
con il '*Progetto Lumière*')

15

Belief Network (o Bayesian Network, o Reti Causali Probabilistiche, RCP): Definizione

Grafi orientati senza cicli i cui nodi rappresentano variabili a più valori.
Gli archi rappresentano l'esistenza di *influenze causali* fra le variabili che collegano.

La forza di queste influenze è misurata in termini di *probabilità condizionate*.

Se il grafo contiene le variabili $x_1, x_2, \dots, x_i, \dots, x_n$,
e G_i è l'insieme dei *genitori* di x_i ,

una quantificazione completa e consistente dell'influenza fra le variabili
può essere ottenuta specificando,

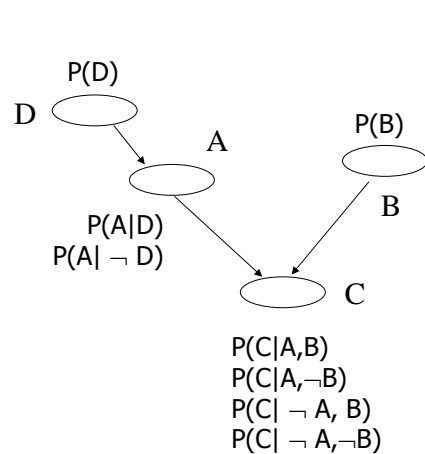
$\forall x_i$, la *Prob*($x_i | G_i$).

In questo caso, la *distribuzione congiunta di probabilità* $\text{Prob}(x_1, x_2, \dots, x_i, \dots, x_n)$ potrà essere approssimata con il prodotto delle probabilità condizionate:

$$\text{Prob}(x_1, x_2, \dots, x_i, \dots, x_n) = \prod_i \text{Prob}(x_i | G_i).$$

16

Un esempio semplice di RCP



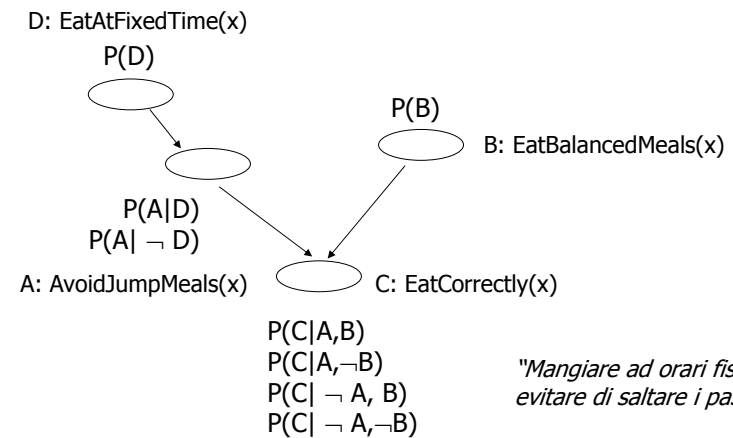
Il grafo rappresenta ipotesi di 'indipendenza condizionale' tra le variabili associate ai suoi nodi:

A è indipendente da B e da C, condizionatamente a D:
 $P(A|B,C,D) = P(A|D)$

C è indipendente da D, condizionatamente ad A:
 $P(C|A,B,D) = P(C|A,B)$

17

Cosa potrebbe rappresentare la RCP che abbiamo appena visto



"Mangiare ad orari fissi aiuta ad evitare di saltare i pasti ..."

"Evitare di saltare i pasti e mangiare pasti bilanciati contribuiscono a mangiare correttamente"

18

Fasi nella Costruzione di una RCP

1. Stabilire la differenza fra variabili 'nascoste' (hidden) e 'osservabili'
2. Definire la struttura del grafo:
 - semantica dei nodi
 - semantica delle relazioni
 - attenzione alle ipotesi di indipendenza condizionale!
3. Assegnare i parametri

19

Quali parametri occorre specificare nella RCP

- Le probabilità 'a priori' dei nodi-radice (quelli che non hanno genitori)
- Le probabilità condizionate di tutti i nodi intermedi e dei nodi-foglia (quelli che non hanno figli).

Assegnare un valore a questi parametri significa, implicitamente, assegnare valori di probabilità a priori a ciascuna delle variabili.

Nell'esempio precedente:

$$P(A) = P(A, D) + P(A, \neg D) = P(A | D) * P(D) + P(A | \neg D) * P(\neg D)$$

... e tutti questi sono parametri di input...

... analogamente per C.

20

Come assegnare i parametri alla RCP?

Due metodi, che riflettono l'idea soggettivista e l'idea frequentista:

- Assegnare le prob a priori e condizionate sulla base della conoscenza soggettiva del dominio, e quindi del proprio 'grado di fiducia' sull'associazione fra le variabili;
- 'Calcolare' questi parametri da una misura di frequenze su un set di dati, applicando metodi di apprendimento.

Nota:

Il metodo a. richiede l'applicazione di procedure per la verifica della *coerenza* delle valutazioni.

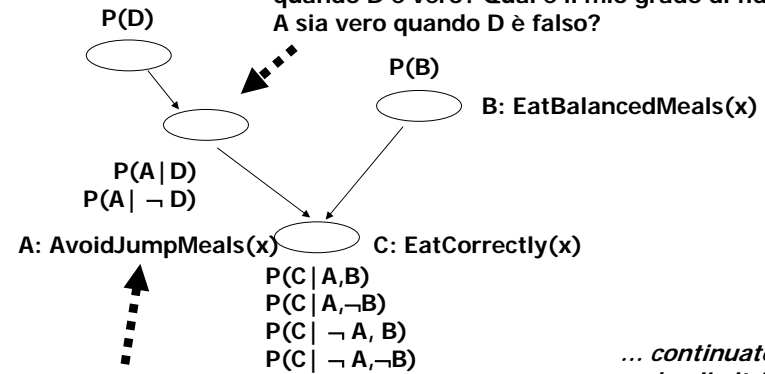
Il metodo b. richiede la realizzazione di un *esperimento* su un numero sufficientemente elevato di casi, rappresentativi della popolazione a cui la RCP dovrà poi essere applicata.

Come vedremo, i due metodi possono essere integrati con una procedura di *apprendimento interattivo della rete*.

21

Vediamo allora, ad esempio, come assegnare parametri alla nostra rete (Metodo a)

D: EatAtFixedTime(x) Qual è il mio grado di fiducia sul fatto che A sia vero quando D è vero? Qual è il mio grado di fiducia sul fatto A sia vero quando D è falso?



Quanto è probabile che si eviti di saltare i pasti se si mangia ad orari fissi? .9?

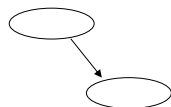
Quanto è probabile che questo avvenga se *non* si mangia ad orari fissi? .7?

22

... Continuiamo l'esempio

D: EatAtFixedTime(x)

$$P(D) = .6$$



$$P(B) = .4$$

B: EatBalancedMeals(x)

$$P(A|D) = .9$$

$$P(A|\neg D) = .7$$

A: AvoidJumpMeals(x)

C: EatCorrectly(x)

$$P(C|A,B) = .99$$

$$P(C|A,\neg B) = .7$$

$$P(C|\neg A,B) = .6$$

$$P(C|\neg A,\neg B) = .4$$

Prova a calcolare P(C) e P(A)

23

Verifiche sulle RCP

- Calibrare i parametri:
 - calcolare la prob a priori dei nodi intermedi e dei nodi-foglia
 - valutarne la plausibilità
 - testare la sensibilità del modello a variazioni più o meno piccole nei parametri e
 - rivedere i parametri critici
- Verificare la struttura:
 - inserire/cancellare archi rivedendo le ipotesi di indipendenza condizionale

Vediamo un algoritmo per la manipolazione dei parametri nelle RCP (D. Spiegelhalter)

24

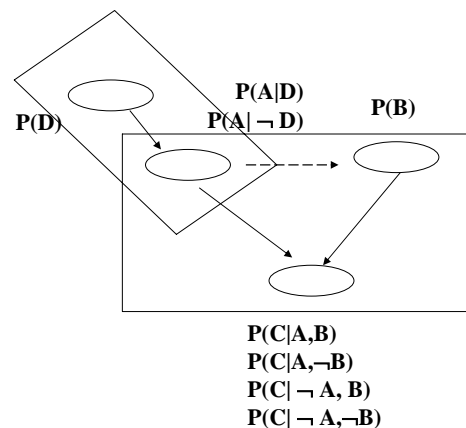
Propagazione dell'Evidenza in un BN

Il problema della propagazione corretta dell'evidenza in una RCP è NP-hard *

Un buon algoritmo di propagazione approssimata dell'evidenza (di Lauritzen e Spiegelhalter) opera sull'albero delle cricche (*junction tree*) anziché sul grafo iniziale.

* G Cooper: "The computational complexity of probabilistic inference using bayesian belief networks", *Artificial Intelligence*, 1990

Outline dell'Algoritmo di Spiegelhalter



Il calcolo delle probabilità a priori dei nodi si può effettuare con un algoritmo di:

- moralizzazione del grafo,
- decomposizione del grafo in 'cricche'
- calcolo delle prob 'marginali' delle cricche.

Le prob a priori di ogni nodo sono quindi calcolabili in funzione delle prob delle cricche e delle loro intersezioni.

Le prob marginali del grafo sono funzione del rapporto fra:

- prodotto delle prob delle cricche e
- prodotto delle prob delle loro intersezioni

Vediamo meglio:

Introduciamo il concetto di *probabilità marginale*

La probabilità marginale di un grafo è la *probabilità congiunta* dei valori assunti dai suoi nodi.

Nel nostro esempio: $P(A,B,C,D)$, oppure $P(-A,B,C,D)$, $P(A,-B,C,D)$, $P(-A,-B,C,D)$, ... ecc

Sappiamo che

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

Ma, per le ipotesi di indipendenza condizionale introdotte implicitamente nel grafo, possiamo semplificare:

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

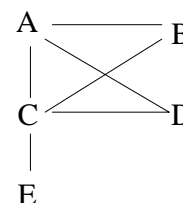
Cosa è una 'cricca'

Dato un grafo non orientato G con un insieme V di nodi.

Un sottoinsieme W di V si dice *cricca* di G se W è un *insieme completo massimo*, cioè

un grafo in cui tutte le coppie di nodi sono adiacenti e non c'è un altro insieme completo che lo contiene.

Esempio:



ABC e CDE non sono cricche

ABCD è una cricca

CE è una cricca

C è l'intersezione fra le cricche

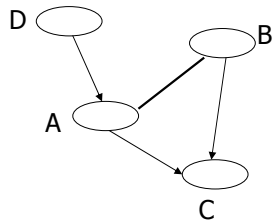
ABCD e CE

La distribuzione congiunta di probabilità su G può essere espressa in funzione delle distribuzioni congiunte delle sue cricche, in un particolare grafo non orientato

Moralizzazione di un grafo

La 'moralizzazione' di un grafo consiste nel trasformare un grafo orientato in un grafo *non* orientato equivalente introducendo nuovi archi fra i nodi-genitori dello stesso nodo e lasciando i parametri invariati.

Esempio: moralizziamo il grafo del nostro esempio:



In questo nuovo grafo, ABC e DA sono cricche. A è l'intersezione fra di esse.

29

Probabilità a priori dei nodi

Prima abbiamo visto come calcolarle in funzione dei parametri inseriti nella rete:

$$P(A) = P(A, D) + P(A, \neg D) =$$

$$P(A | D) * P(D) + P(A | \neg D) * P(\neg D)$$

Ma (A,D) è una cricca, così come (A,B,C).

Quindi, in generale, le probabilità a priori dei nodi si possono calcolare

a partire dalle probabilità delle cricche:

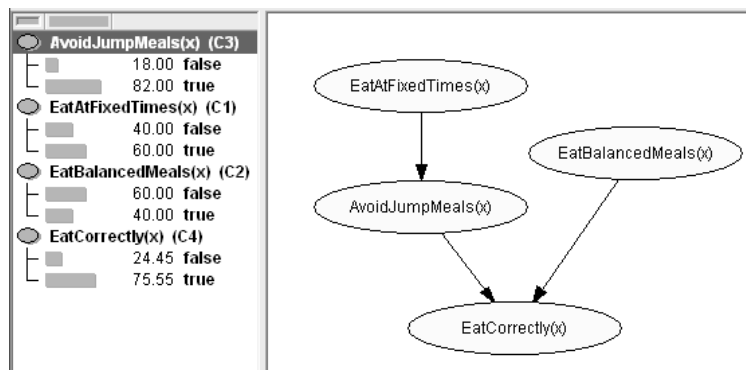
$$P(A) = P(A,B,C) + P(A,\neg B,C) + P(A,B,\neg C) + P(A,\neg B,\neg C)$$

$$P(C) = P(A,B,C) + P(\neg A,B,C) + P(A,\neg B,C) + P(\neg A,\neg B,C)$$

... ECC

30

Probabilità a priori dei nodi nel nostro esempio



$$P(\text{AvoidJumpMeals}) = P(\text{AvoidJumpMeals} | \text{EatAtFixedTimes}) * P(\text{EatAtFixedTimes}) + P(\text{AvoidJumpMeals} | \neg \text{EatAtFixedTimes}) * P(\neg \text{EatAtFixedTimes}) = .9 * .6 + .7 * .4 = .82$$

31

Tracing dell'algoritmo in Hugin

Marriages (bigamy allowed):

Marrying C2 and C3

Triangulation by minimum fill-in weight heuristic:

Fill-in links and node numbering:

4 C4

3 C2

2 C3

1 C1

Cliques:

Clique 1, 2 members (C1, C3), table size = 4 ← (A,D)

Clique 2, 3 members (C3, C2, C4), table size = 8 ← (A,B,C)

Total clique table size: 12

The junction forest:

Creating junction tree with clique 1 as root ...

Cliques 2 and 1 linked, separated by {C3} (table size = 2)

Checking tables for all nodes ...

↑
A

32

Assignment of potentials to cliques:

Node C4 assigned to clique 2
 Node C3 assigned to clique 1
 Node C2 assigned to clique 2
 Node C1 assigned to clique 1

Checking tables for all nodes ...
 Mon Jan 30 13:53:09 2006

Beliefs for chance nodes:

C4 --		
false	0.244480	
true	0.755520	
C3 --		
false	0.180000	
true	0.820000	
C2 --		
false	0.600000	
true	0.400000	
C1 --		
false	0.400000	
true	0.600000	

Calcolo delle 'probabilità marginali'

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

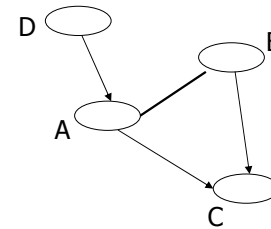
Semplificando:

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

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

$$P(A|D) * P(D) = P(A,D)$$

$$P(A,B,C,D) = (P(A,B,C) * P(A,D)) / P(A)$$



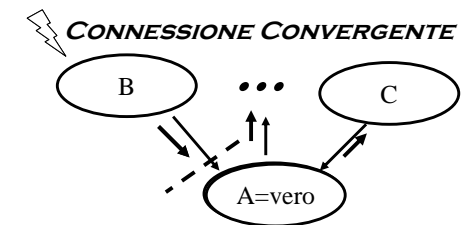
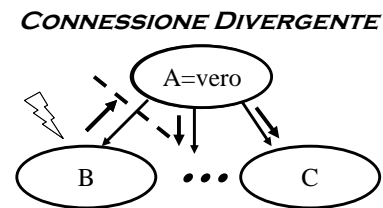
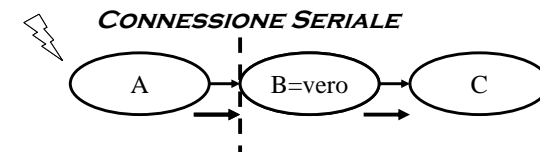
Distribuzione congiunta di probabilità = rapporto fra: prodotto delle probabilità delle cricche e prodotto delle probabilità delle loro intersezioni

Vediamo ora come utilizzare una RCP

Data una 'evidenza' (valore di verità per alcune delle variabili rappresentate nel grafo),
 Valutare come questa evidenza influisce sulla probabilità delle altre variabili rappresentate nel grafo.

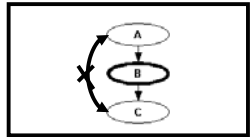
Introduciamo, prima di tutto, il concetto di 'd-separazione'.

I tre modi in cui una evidenza può essere trasmessa in una RCP

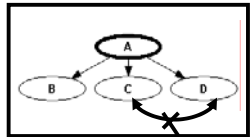


Proprietà di d-separazione nelle RCP

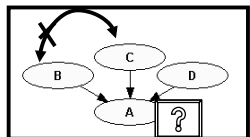
Due variabili sono d-separate se esiste una variabile intermedia tale che la connessione sia:



Seriale e lo stato della variabile intermedia è noto

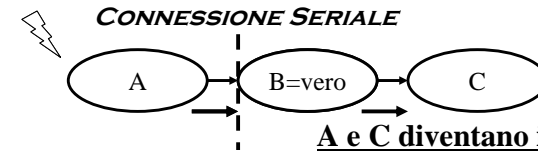


Divergente e lo stato della variabile intermedia è noto

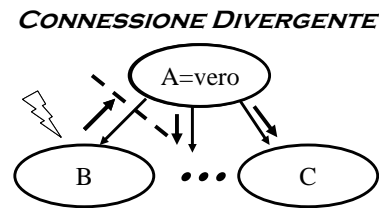


Convergente e lo stato della variabile intermedia è non noto

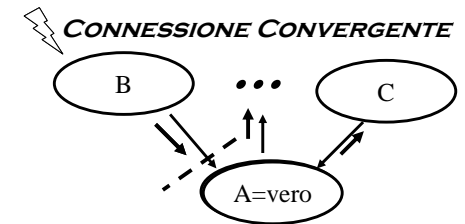
Se due variabili sono d-separate, sono condizionalmente indipendenti (nessuna evidenza 'passa' dall'una all'altra).



A e C diventano indipendenti (d-separati) dato B

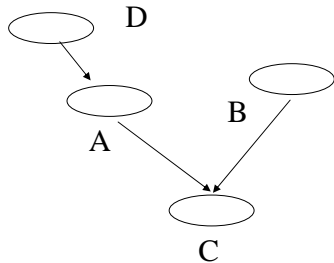


B...C diventano indipendenti (d-separati) dato A



B...C diventano dipendenti (d-connessi) dato A

Aggiornamento della probabilità delle cricche



Supponiamo di avere un'evidenza sul nodo C; Ad esempio, $P(C) = 1$.
Vogliamo valutare come questa conoscenza influisce sulla probabilità dell'evento associato al nodo A: $P^*(A) = P(A|C)$.
A è l'intersezione fra le due cricche AD e ABC.

Consideriamo la cricca AD.

$$\begin{aligned} P^*(A,D) &= P(A,D|C) = P(A,D,C) / P(C) = \\ &= P(D|A,C) * P(A|C) = \\ &= P(D|A) * P(A|C) = \\ &= P(A,D) / P(A) * P(A|C) \\ &= P(A,D) * (P^*(A) / P(A)) \end{aligned}$$

Cioè:
La probabilità a posteriori della cricca AD è = alla sua probabilità a priori, per il rapporto fra prob a posteriori e a priori dell'intersezione fra le due cricche AD e ABC.

Probabilità a posteriori dei nodi

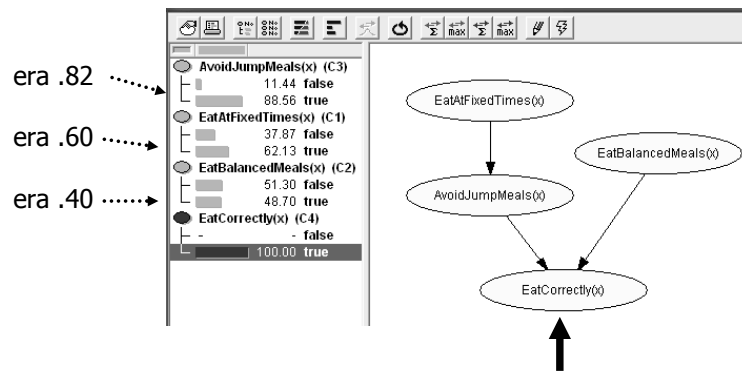
Si può calcolare dalla probabilità a posteriori delle cricche:

$$P^*(A) = P^*(A,B,C) + P^*(A,-B,C) + P^*(A,B,-C) + P^*(A,-B,-C)$$

$$P^*(C) = P^*(A,B,C) + P^*(-A,B,C) + P^*(A,-B,C) + P^*(-A,-B,C)$$

... ecc

Nel nostro esempio



So che l'individuo considerato 'mangia correttamente':
 $(P(\text{EatCorrectly}(x) = 1))$.
 Qual è, ora la probabilità che x mangi ad orari fissi?
 Qual è la probabilità che segua una dieta corretta?

41

In che ordine esplorare l'albero delle cricche?

Viene definito un ordinamento dei nodi, secondo il criterio seguente:

- Il nodo di cui si conosce il valore è quello di partenza;
- Ad ogni passo, si sceglie il nodo collegato al massimo numero di nodi già scelti.

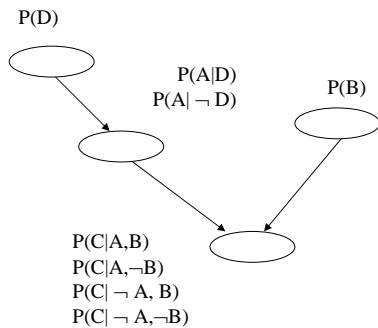
Questo ordinamento dei nodi crea un *sequenza di cricche*.

La propagazione dell'evidenza viene effettuata sull'albero delle cricche anziché sul grafo orientato originale (e l'algoritmo è meno complesso).

Nota: per evitare ciclicità negli aggiornamenti, ogni volta che il nodo diventa 'noto' viene eliminato dal grafo.

42

Forme di ragionamento

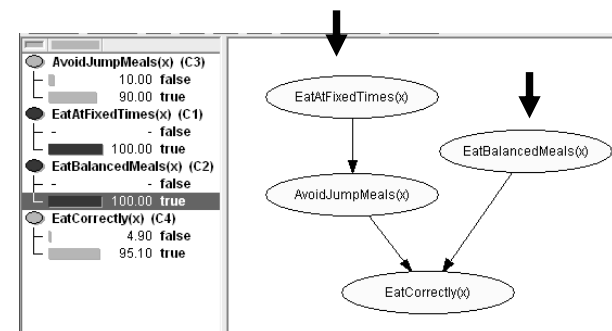


Si possono valutare le conseguenze, sull'intero grafo, di 'evidenza' acquisita su uno o più nodi:

- ragionando *dalle cause agli effetti* (rag 'prognostico'): noto il valore di D, valutare la prob di A,B,C
- ragionando *dagli effetti alle cause* (rag 'diagnostico'): noto il valore di C, valutare la prob di A, B, D (come abbiamo visto prima)

43

Un esempio di ragionamento prognostico



So che l'individuo considerato 'mangia ad orari fissi' e 'segue una dieta bilanciata':
 $(P(\text{EatAtFixedTimes}(x) = 1); P(\text{EatBalancedMeals}(x) = 1))$.
 Qual è, la probabilità che x mangi correttamente??

44

Riprendiamo ora l'esempio della persuasione (Es 2.4)

- **Task model: le strategie persuasive di Walton**
 - Appeal to positive consequences*
"Se ritieni che compiere una determinata azione comporti conseguenze importanti per te e sei in grado di compierla, dovresti farla".
 - Appeal to negative consequences*
"Se ritieni che compiere una determinata azione comporti conseguenze che desideri evitare e puoi evitarle di compierla, dovresti farlo".
- **Domain model**
"Fare sport fa bene alla salute e alla forma fisica. Fare una vita sedentaria aumenta il rischio di ingrassare e di perdere tono muscolare".
"Il running è un particolare tipo di sport". Tutti i giovani senza particolari problemi di salute sono in grado di fare running.
- **User model**
Young(G)
Healthy(G)
Likes(G,InShape)

45

Come avevamo formalizzato il problema

- **Task model: le strategie persuasive di Walton**
 $\forall x \forall a \forall g ((\text{Implies}(a,g) \wedge \text{Likes}(x,g) \wedge \text{CanDo}(x,a)) \rightarrow (\text{ShouldDo}(x,a)))$
 $\forall x \forall a \forall g ((\text{Implies}(a,g) \wedge \neg \text{Likes}(x,g) \wedge \text{CanAvoid}(x,a)) \rightarrow \neg(\text{ShouldDo}(x,a)))$
- **Domain model**
 $\forall s \forall x (\text{Sport}(s) \rightarrow \text{Implies}(s, \text{InShape}))$
 $\forall s \forall x (\text{Sport}(s) \rightarrow \text{Implies}(s, \text{GoodHealth}))$
 $\forall s \forall x ((\text{Sport}(s) \wedge \text{Young}(x) \wedge \text{Healthy}(x)) \rightarrow \text{CanDo}(x,s))$
 $\text{Sport}(R)$
- **User model**
 $\text{Young}(G)$
 $\text{Healthy}(G)$
 $\text{Likes}(G, \text{InShape})$

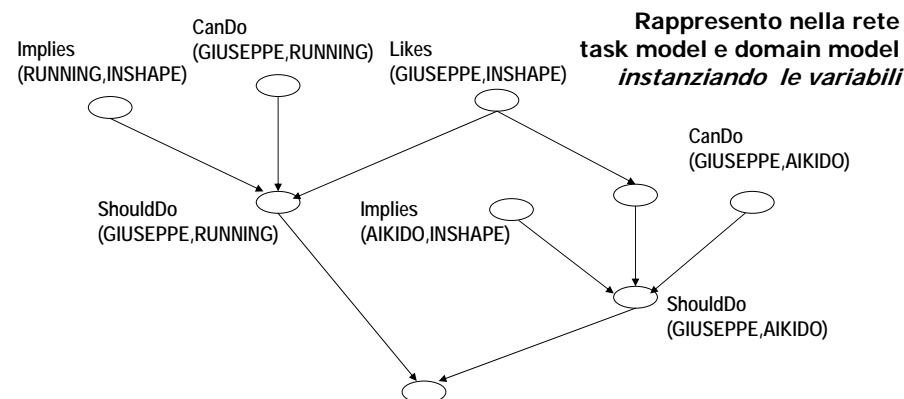
46

Inseriamo l'incertezza

- **Task model: le strategie persuasive di Walton**
 - Appeal to positive consequences*
"Se sei abbastanza convinto che compiere una determinata azione comporti conseguenze importanti per te e *pensi che potresti* compierla, allora *probabilmente* dovresti farla".
 - Appeal to negative consequences*
"Se sei abbastanza convinto che compiere una determinata azione comporti conseguenze che desideri evitare e *pensi che potresti* evitare di compierla, allora *probabilmente* dovresti farlo".
- **Domain model**
"Fare sport fa, *in genere*, bene alla salute e alla forma fisica. Fare una vita sedentaria aumenta, *in genere*, il rischio di ingrassare e di perdere tono muscolare".
"Il running è un particolare tipo di sport". *La maggior parte dei* giovani senza particolari problemi di salute sono in grado di fare running. ⁴⁷

47

Come formalizzare il problema con una RCP



Ho perso in generalità ma...

ShouldDo (GIUSEPPE,s)

...posso scegliere l'argomento più forte, data la conoscenza sul dominio e su Giuseppe, con il grado di incertezza di ogni elemento!! ⁴⁸

48

Come cambia l'algoritmo di ricerca della strategia 'piu' promettente?

- Introduco nella rete l'evidenza (conoscenze che ho sul dominio e su Giuseppe)
- Propago
- Osservo i genitori del nodo 'ShouldDo(Giuseppe,s)'
- Cerco quello con probabilità massima: questo corrisponde alla strategia 'ottimale'.

Che messaggi posso generare?

- *Quale sport* deve fare Giuseppe
- *Perché*: osservando i genitori di questo nodo e traducendo in linguaggio naturale sia il loro significato che il loro valore di probabilità

Esercizio

Completate l'esempio della persuasione, estendendo la struttura della rete, inserendo i parametri nel modello, simulando diverse situazioni possibili e ragionando sui risultati

50

Riferimenti

- Introduzione ai BN: articolo di Spiegelhalter reperibile sul sito
- Descrizione approfondita dei metodi: libro di Judea Pearl
- Argomenti specifici:
 - Il Progetto Lumiere della Microsoft:
research.microsoft.com/~horvitz/lumiere.HTM
 - applicazione alla gestione di modelli di Studente:
articoli di Kurt VanLehn
 - applicazione alla costruzione di modelli dell' Interlocutore nel dialogo: Ingrid Zukerman (Università di Monash, Australia)
 - applicazione alla traduzione di stati mentali in 'espressioni':
Gene Ball (Microsoft)
 - applicazione alla modellizzazione cognitiva di stati emotivi: articoli nostri reperibili sul sito

51