Modelli Probabilistici per le Decisioni

Enza Messina

Reti Bayesiane

DISCo

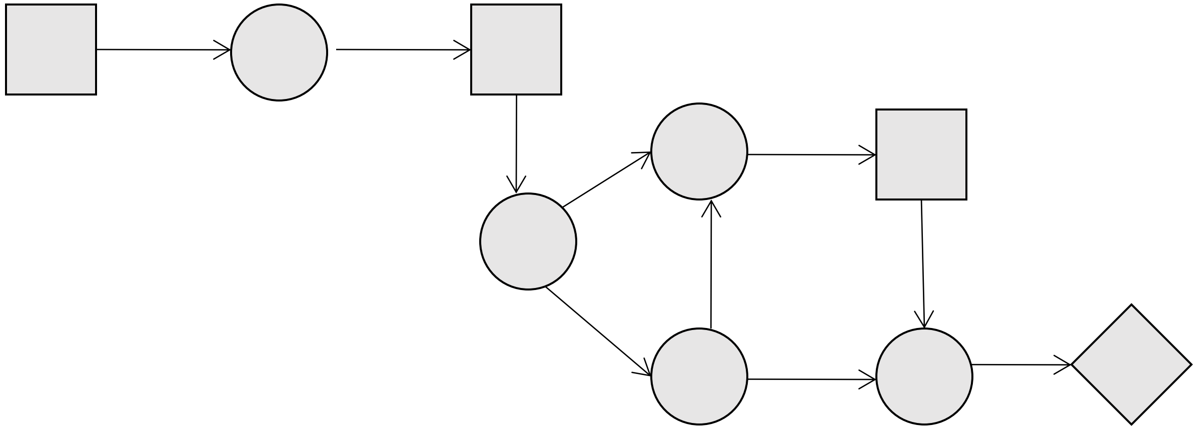
Università degli Studi di Milano-

Bicocca

Viale Sarca, 336

20126 Milano

messina@disco.unimib.it



Semantica delle reti Bayesiane

Semantica delle Reti Bayesiane: *relazioni di indipendenza condizionale nelle Reti Bayesiane*

Abbiamo appena proposto una *semantica numerica* per le *Reti Bayesiane* in termini di rappresentazione di una distribuzione congiunta di probabilità

$$P(x_1,\ldots,x_n) = \prod_{i=1}^{n} P(x_i \mid parents\,(X_i))$$

L'utilizzo di tale semantica ha consentito di derivare un metodo per la costruzione di *Reti Bayesiane* che ha posto in evidenza come *ogni nodo* risulti essere *condizionalmente indipendente dai* suoi *predecessori data la conoscenza dello stato dei suoi nodi genitore:*

$$\mathbf{P}\,(X_{(i)} \mid\; = X_{(1)},\ldots,X_{(i-1)}) = \mathbf{P}(X_i \mid Parents\,(X_{(i)}))$$

Dato un ordinamento di variabili $\{X_{(1)},\ldots, X_{(n)}\}$, per ogni variabile (nodo) X_i i suoi *predecessori* o *non discendenti* sono i nodi $\{X_{(1)},\ldots, X_{(i-1)}\}$ mentre i suoi *successori* o *discendenti* sono i nodi $\{X_{(i+1)},\ldots, X_{(n)}\}$.

È possibile procedere anche secondo la direzione opposta, ovvero partire da una *semantica topologica* che specifica le *relazioni di indipendenza condizionale* codificate dalla *componente strutturale* della *Rete Bayesiana* e giungere a ricavarne una *semantica numerica.*

Semantica delle reti Bayesiane

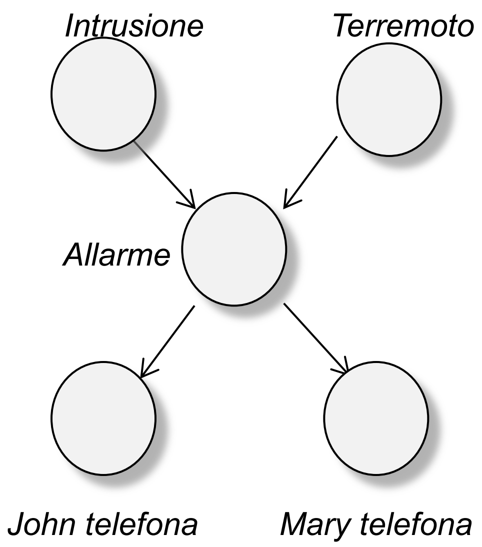
La *semantica topologica* viene fornita per mezzo di una delle due specificazioni equivalenti:

1. Un *nodo* è *condizionalmente indipendente dai* suoi *non-discendenti* dati *i suoi genitori.*

   John telefona è *indipendente* da *Intrusione* e *Terremoto dato Allarme*

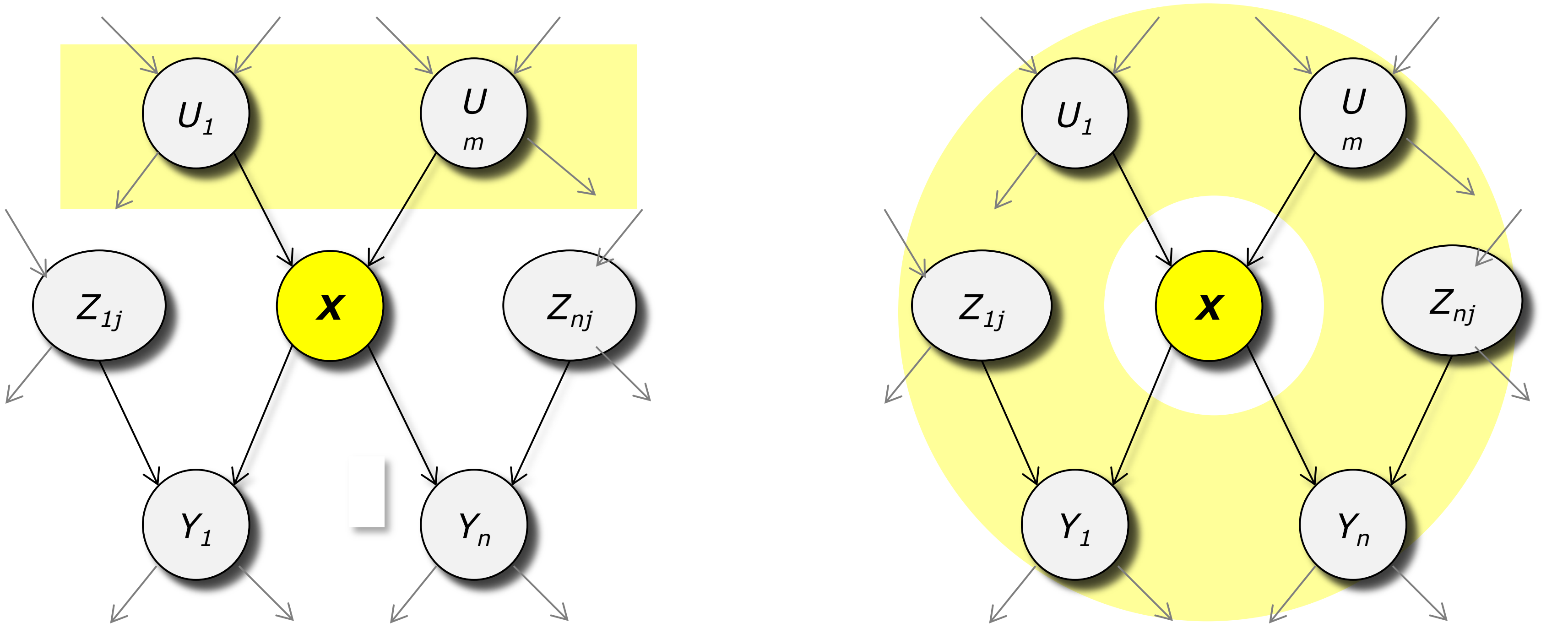
2. Un *nodo* è *condizionalmente indipendente da tutti i nodi restanti* della rete, *data* la *conoscenza* dello stato dei suoi *genitori*, dei suoi *figli* e dei *genitori dei suoi figli*, ovvero l'insieme di nodi che è noto con il nome di *Markov Blanket.*

   Intrusione è *indipendente* da *John telefona* e *Mary telefona dati Allarme* e *Terremoto*



Semantica delle reti Bayesiane

Le due specificazioni presentate sono illustrate nelle figure riportate sotto.



Il nodo ***X*** è *condizionalmente indipendente dai* suoi *non-discendenti* (***Z***$_{ij}$) *data* ***la*** *conoscenza* dello *stato* dei suoi *genitori* (***U***$_i$)

Il nodo ***X*** è *condizionalmente indipendente* da tutti i *nodi restanti* della rete *data* la *conoscenza* dello *stato* del suo *Markov Blanket* (***U***$_i$, ***Z***$_{ij}$, ***Y***$_j$).

D-separazione delle variabili

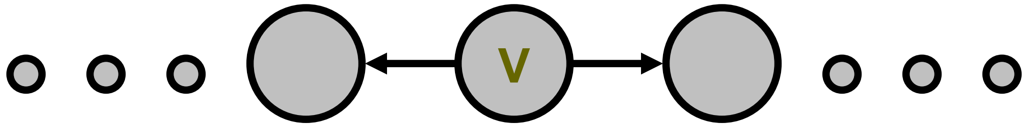
- Fortunatamente c'è un metodo relativamente semplice per determinare se in una rete Bayesiana due variabili sono condizionalmente indipendenti

- Definizione: *X* e *Z* sono *d-separate* da un insieme E di variabili con evidenza (osservazioni) se e solo se ogni cammino non orientato da X a Z è "bloccato"

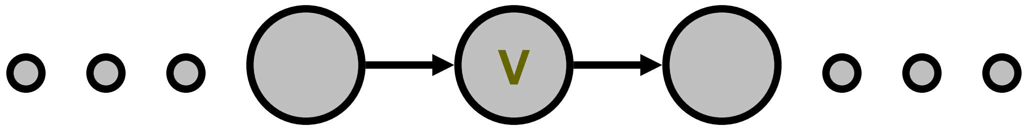
- Un cammino è "bloccato" se e solo se vale almeno una delle seguenti condizioni

Cammini “bloccati”

- Lungo il cammino esiste una variabile V tale che:
  - Appartiene all’insieme E delle variabili con evidenza
  - Gli archi che collegano V al cammino sono “*tail to tail*”



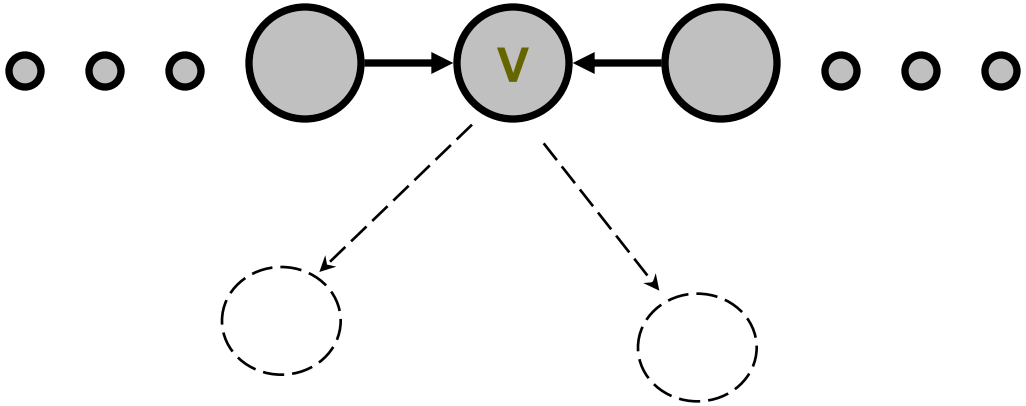
- Oppure, esiste una variabile V lungo il cammino tale che
  - Appartiene all’insieme E delle variabili con evidenza
  - Gli archi che collegano V al cammino sono *“tail-to-head”*



- Oppure, ...

Cammini “bloccati”

- … Oppure, esiste una variabile V lungo il cammino tale che
  - NON appartiene all’insieme E delle variabili con evidenza
  - Nessuno dei suoi discendenti appartiene all’insieme E delle variabili con evidenza
  - Gli archi che collegano V al cammino sono *“head-to-head”*



D-Separazione e independenza

- Teorema [Verma & Pearl, 1998]:
  - Se in una rete Bayesiana un insieme E di variabili con evidenza D-separa X e Z, allora X e Z sono indipendenti.
- La *d*-separazione può essere calcolata in tempo lineare.
- Quindi abbiamo a diposizione un algoritmo efficiente per inferire automaticamente se apprendere il valore di una variabile può fornirci delle informazioni aggiuntive su alter variabili, date le informazioni a disposizione.