Reti Bayesiane II

IALab A.A. 2018/2019

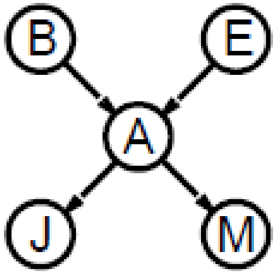
Inferenza Esatta

Inference by enumeration

Slightly intelligent way to sum out variables from the joint without actually constructing its explicit representation

Simple query on the burglary network:

$\mathbf{P}(B|j,m)$
$= \mathbf{P}(B,j,m)/P(j,m)$
$= \alpha\mathbf{P}(B,j,m)$
$= \alpha \sum_e \sum_a \mathbf{P}(B,e,a,j,m)$



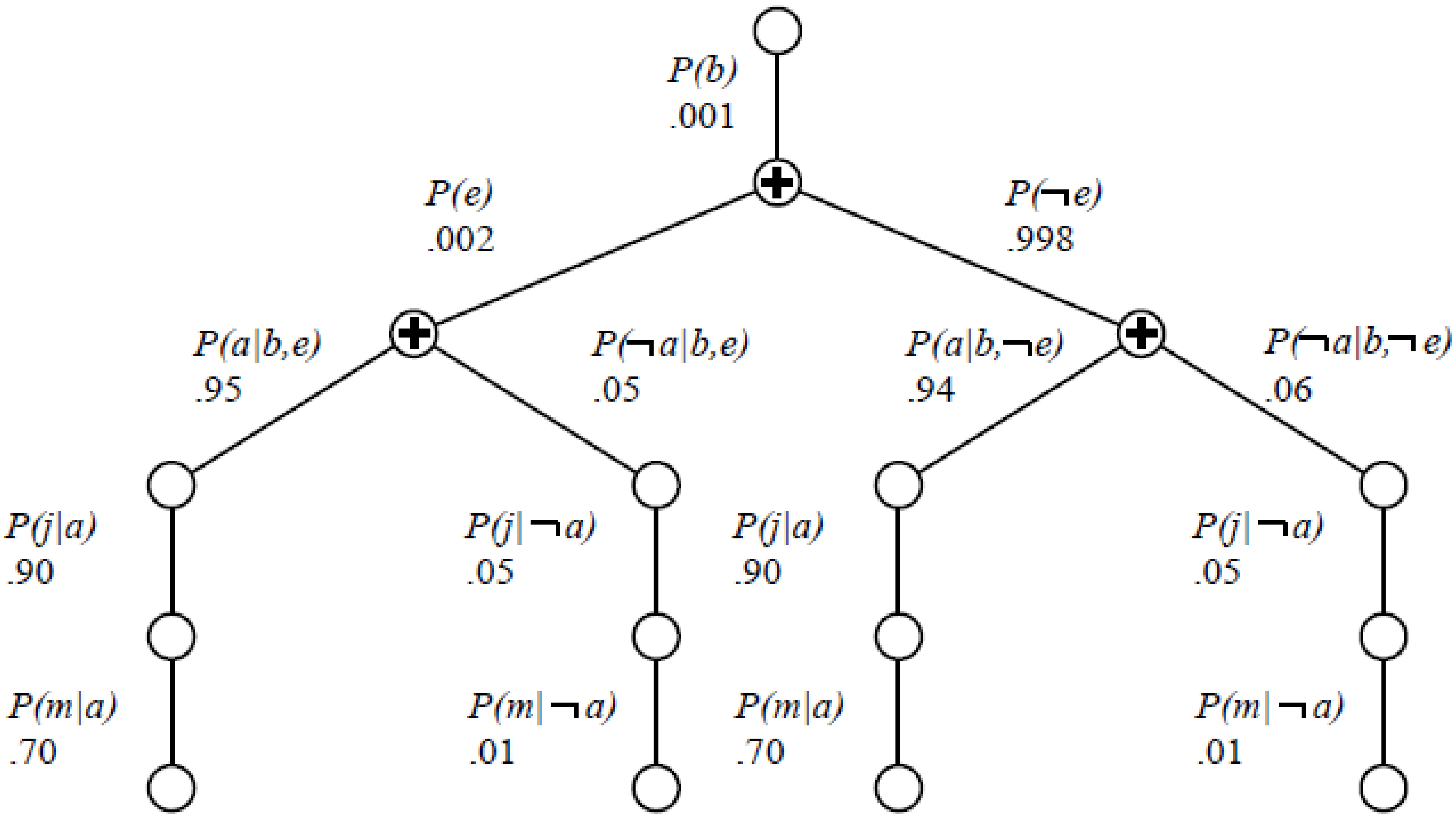
Rewrite full joint entries using product of CPT entries:

$\mathbf{P}(B|j,m)$
$= \alpha \sum_e \sum_a \mathbf{P}(B)P(e)\mathbf{P}(a|B,e)P(j|a)P(m|a)$
$= \alpha\mathbf{P}(B) \sum_e P(e) \sum_a \mathbf{P}(a|B,e)P(j|a)P(m|a)$

Enumerazione sulla Distribuzione Congiunta

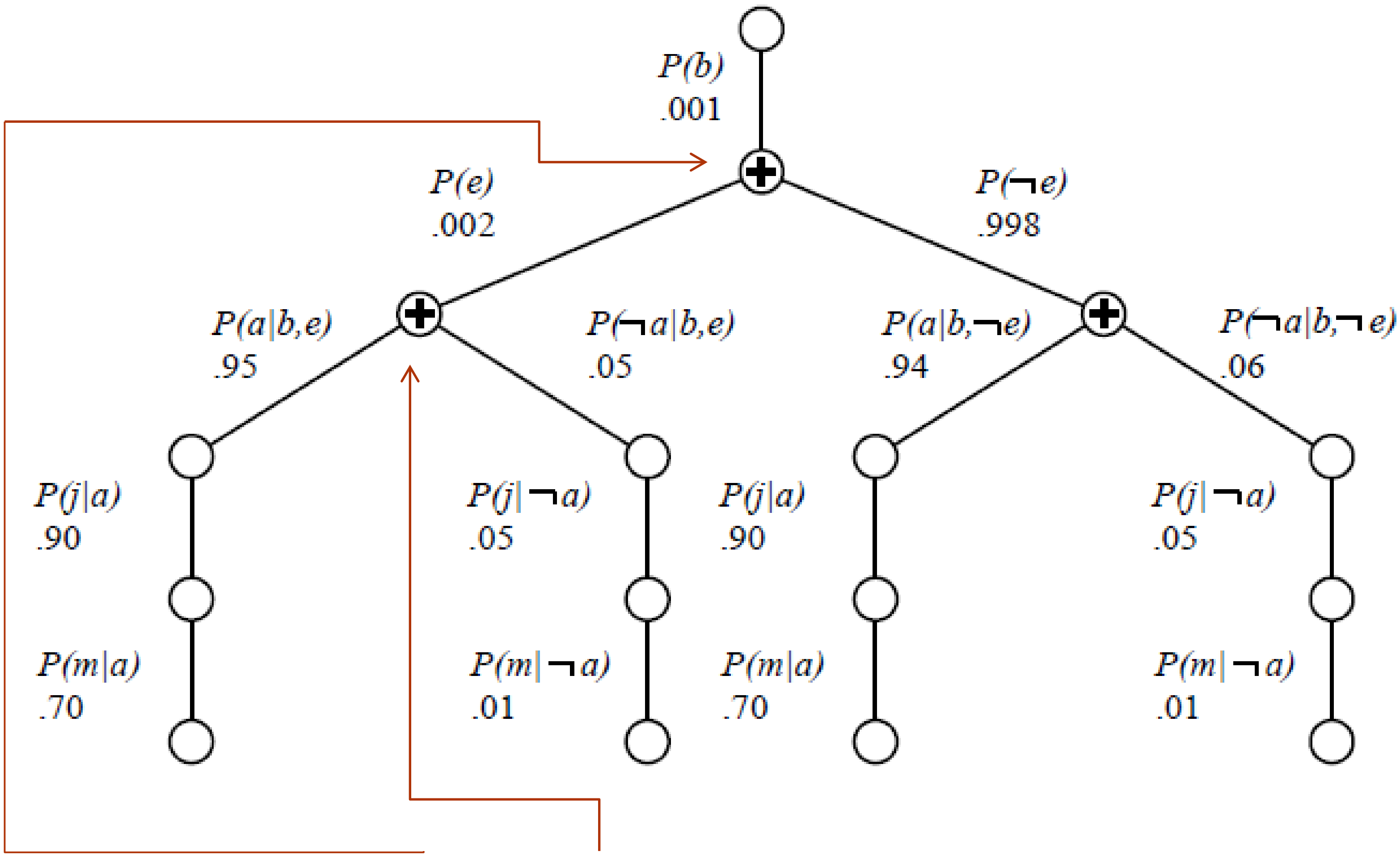
Ordine Topologico da sinistra a destra

Evaluation tree



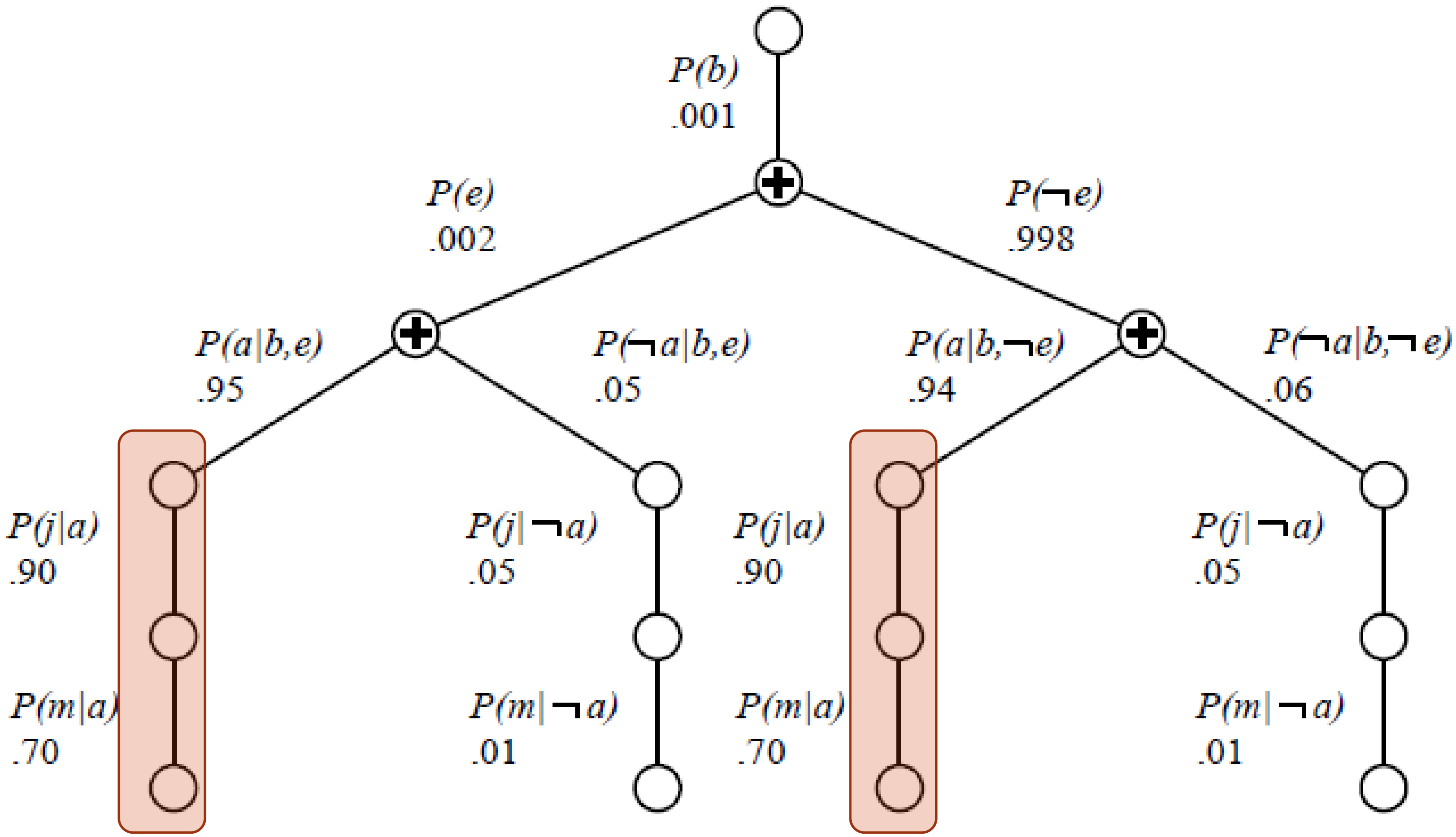
$$\alpha \mathbf{P}(B) \sum_e P(e) \sum_a \mathbf{P}(a|B,e)P(j|a)P(m|a)$$

Evaluation tree



$$\alpha \mathbf{P}(B) \sum_e P(e) \sum_a \mathbf{P}(a|B,e) P(j|a) P(m|a)$$

Evaluation tree



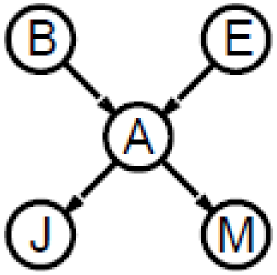
Enumeration is inefficient: repeated computation

e.g., computes $P(j|a)P(m|a)$ for each value of e

Inference by enumeration

Slightly intelligent way to sum out variables from the joint without actually constructing its explicit representation

Simple query on the burglary network:



$\mathbf{P}(B|j,m)$
$= \mathbf{P}(B,j,m)/P(j,m)$
$= \alpha\mathbf{P}(B,j,m)$
$= \alpha\ \sum_e \sum_a \mathbf{P}(B,e,a,j,m)$

Rewrite full joint entries using product of CPT entries:
$\mathbf{P}(B|j,m)$
$= \alpha\ \sum_e \sum_a \mathbf{P}(B)P(e)\mathbf{P}(a|B,e)P(j|a)P(m|a)$
$= \alpha\mathbf{P}(B)\ \sum_e P(e)\ \sum_a \mathbf{P}(a|B,e)P(j|a)P(m|a)$

Recursive depth-first enumeration: $O(n)$ space, $O(d^n)$ time

Variable Elimination

Fare le somme "bottom-up" memorizzando i risultati intermedi (**factors**) per evitare computazioni ripetute.

Dynamic Programming

Esempio di simple query **P**(B|j,m)

1. riscrivo in fattori secondo Rete Bayesiana:

$$\mathbf{P}(B|j,m) = \alpha\,\mathbf{P}(B)\sum_{e} P(e)\sum_{a} P(a|B,e)P(j|a)P(m|a)$$

2. associo ciascuna Distribuzione di Probabilità nella formula alla variabile di cui è la CPT:

$$\mathbf{P}(B|j,m) = \alpha\,\underbrace{\mathbf{P}(B)}_{\mathbf{B}}\sum_{e} \underbrace{P(e)}_{\mathbf{E}}\sum_{a} \underbrace{P(a|B,e)}_{\mathbf{A}}\underbrace{P(j|a)}_{\mathbf{J}}\underbrace{P(m|a)}_{\mathbf{M}}$$

Variable Elimination

$$\mathbf{P}(B|j,m) = \alpha\, \underbrace{\mathbf{P}(B)}_{\mathbf{B}} \sum_{e} \underbrace{P(e)}_{\mathbf{E}} \sum_{a} \underbrace{P(a|B,e)}_{\mathbf{A}} \underbrace{P(j|a)}_{\mathbf{J}} \underbrace{P(m|a)}_{\mathbf{M}}$$

3. al posto delle CPT scrivo delle funzioni il cui ***pedice*** è la variabile associata, e i cui ***parametri*** sono le Variabili non assegnate che compaiono nella CPT:

$$\alpha f_B(b) \sum_{e} f_E(e) \sum_{a} f_A(a,b,e) f_J(a) f_M(a)$$

Variable Elimination

$$\alpha f_B(b) \sum_e f_E(e) \sum_a f_A(a,b,e) f_J(a) f_M(a)$$

4. il fattore ***più a destra*** $f_M(a)$ sarà:

a	$f_M(a)$
T	0,70
F	0,01

5. il fattore seguente (verso ***sinistra***) $f_J(a)$ sarà:

a	$f_J(a)$
T	0,9
F	0,05

Variable Elimination

$$\alpha f_B(b) \sum_e f_E(e) \sum_a f_A(a,b,e) f_J(a) f_M(a)$$

6. il fattore $f_A(a,b,e)$ sarà:

a	b	e	f_A(a,b,e)
T	T	T	0,95
T	T	F	0,94
T	F	T	0,29
T	F	F	0,001
F	T	T	0,05
F	T	F	0,06
F	F	T	0,71
F	F	F	0,999

Variable Elimination

$$\alpha f_B(b) \sum_e f_E(e) \sum_a f_A(a,b,e) f_J(a) f_M(a)$$

7. adesso c'è una sommatoria; ***moltiplichiamo*** i tre fattori ottenendo:

$$f_{AJM}(a,b,e) = f_A(a,b,e) \times f_J(a) \times f_M(a)$$

vedi dopo

8. e poi ***sommiamo*** sulla variabile A ottenendo:

vedi dopo

$$f_{\bar{A}JM}(b,e) = \sum_a f_{AJM}(a,b,e)$$

9. continuiamo nello stesso modo con la formula ottenuta:

$$\alpha f_B(b) \sum_e f_E(e)\, f_{\bar{A}JM}(b,e)$$

Multiplication

La ***moltiplicazione*** di due fattori si fa tra **righe consistenti** sulle variabili condivise:

non è moltiplicazione di matrici

A	B	f1(A,B)
T	**T**	**0,3**
T	F	0,7
F	T	0,9
F	F	0,1

B	C	f2(B,C)
T	**T**	**0,2**
T	F	0,8
F	T	0,6
F	F	0,4

A	B	C	f3(A,B,C)
T	**T**	**T**	**0,3 x 0,2**
T	T	F	0,3 x 0,8
T	F	T	0,7 x 0,6
T	F	F	0,7 x 0,4
F	T	T	0,9 x 0,2
F	T	F	0,9 x 0,8
F	F	T	0,1 x 0,6
F	F	F	0,1 x 0,4

Multiplication

La ***moltiplicazione*** di due fattori si fa tra **righe consistenti** sulle variabili condivise:

A	B	f1(A,B)
T	**T**	**0,3**
T	F	0,7
F	T	0,9
F	F	0,1

B	C	f2(B,C)
T	**T**	**0,2**
T	F	0,8
F	T	0,6
F	F	0,4

A	B	C	f3(A,B,C)
T	**T**	**T**	**0,3 x 0,2**
T	T	F	0,3 x 0,8
T	F	T	0,7 x 0,6
T	F	F	0,7 x 0,4
F	T	T	0,9 x 0,2
F	T	F	0,9 x 0,8
F	F	T	0,1 x 0,6
F	F	F	0,1 x 0,4

Multiplication

La ***moltiplicazione*** di due fattori si fa tra **righe consistenti** sulle variabili condivise:

A	B	f1(A,B)
T	**T**	**0,3**
T	F	0,7
F	T	0,9
F	F	0,1

B	C	f2(B,C)
T	T	0,2
T	**F**	**0,8**
F	T	0,6
F	F	0,4

A	B	C	f3(A,B,C)
T	**T**	**T**	**0,3 x 0,2**
T	**T**	**F**	**0,3 x 0,8**
T	F	T	0,7 x 0,6
T	F	F	0,7 x 0,4
F	T	T	0,9 x 0,2
F	T	F	0,9 x 0,8
F	F	T	0,1 x 0,6
F	F	F	0,1 x 0,4

Summation

La ***sommatoria*** su una variabile *A* si fa sommando le righe corrispondenti con *A* che varia tra i valori possibili *T/F*:

A	B	C	f3(A,B,C)
T	T	T	0,06
T	T	F	0,24
T	F	T	0,42
T	F	F	0,28
F	T	T	0,18
F	T	F	0,72
F	F	T	0,06
F	F	F	0,04

$\Rightarrow$

B	C	f4(B,C)
T	T	0,06+0,18
T	F	0,24+0,72
F	T	0,42+0,06
F	F	0,28+0,04

Summation

La ***sommatoria*** su una variabile *A* si fa sommando le righe corrispondenti con *A* che varia tra i valori possibili *T*/*F*:

A	B	C	f3(A,B,C)
T	**T**	**T**	**0,06**
T	T	F	0,24
T	F	T	0,42
T	F	F	0,28
F	**T**	**T**	**0,18**
F	T	F	0,72
F	F	T	0,06
F	F	F	0,04

⇨

B	C	f4(B,C)
T	**T**	**0,06+0,18**
T	F	0,24+0,72
F	T	0,42+0,06
F	F	0,28+0,04

Summation

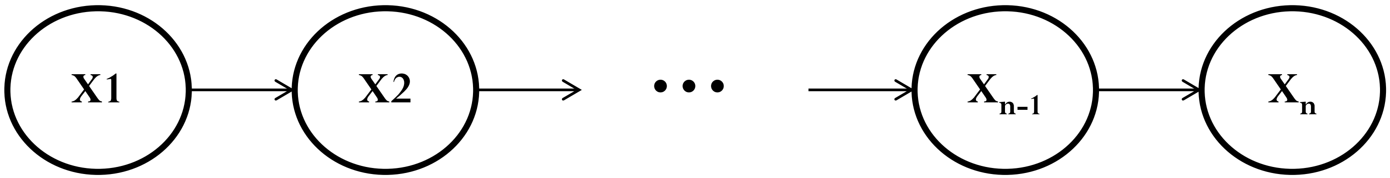
La ***sommatoria*** su una variabile *A* si fa sommando le righe corrispondenti con *A* che varia tra i valori possibili *T/F*:

A	B	C	f3(A,B,C)
T	T	T	0,06
T	**T**	**F**	**0,24**
T	F	T	0,42
T	F	F	0,28
F	T	T	0,18
F	**T**	**F**	**0,72**
F	F	T	0,06
F	F	F	0,04

B	C	f4(B,C)
T	T	0,06+0,18
T	**F**	**0,24+0,72**
F	T	0,42+0,06
F	F	0,28+0,04

Enumeration VS Variable Elimination

Consideriamo una BN speciale, una **catena** (**chain**):

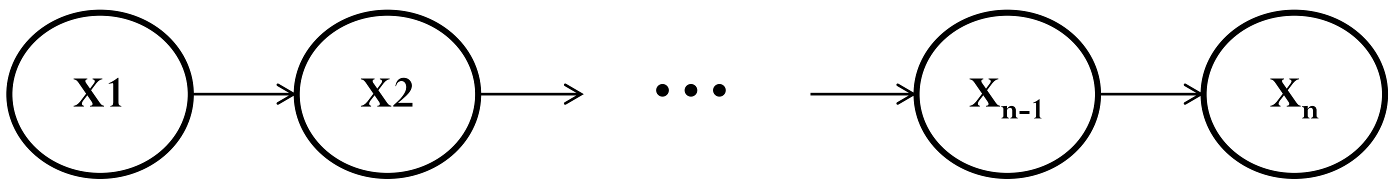


E la simple query $\mathbf{P}(X_1 \mid x_n)$.

Con ***enumerazione*** la complessità è $\mathbf{O(2^n)}$.

Enumeration VS Variable Elimination

Consideriamo una BN speciale, una **catena** (**chain**):



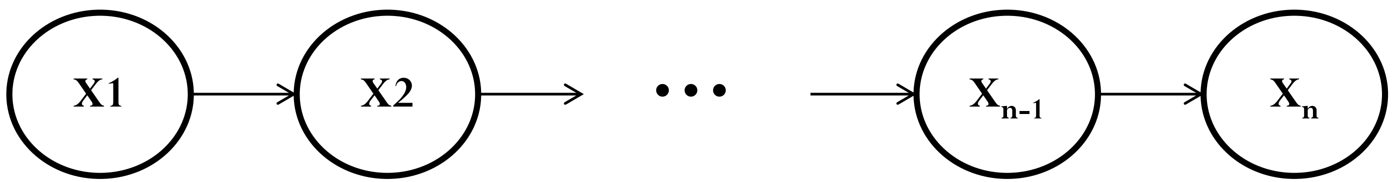
Con ***eliminazione***:

$$P(X1|xn) = P(X1)\sum_{x2} P(x2|X1) \sum_{x3} P(x3|x2) \ldots$$

$$\ldots \sum_{x_{n-2}} P(x_{n-2}|x_{n-3}) \sum_{x_{n-1}} P(x_{n-1}|x_{n-2}) P(x_n|x_{n-1})$$

Enumeration VS Variable Elimination

Consideriamo una BN speciale, una **catena** (**chain**):



Con ***eliminazione***:

$$P(X1|xn) = P(X1)\sum_{x2} P(x2|X1)\sum_{x3} P(x3|x2) \ldots$$

$$\ldots \sum_{x_{n-2}} P(x_{n-2}|x_{n-3}) \sum_{x_{n-1}} P(x_{n-1}|x_{n-2})P(x_n|x_{n-1})$$

Partendo da destra, i fattori non crescono mai oltre 2 variabili (4 valori), con complessità **O(n)**

Complexity of exact inference

Singly connected networks (or polytrees):
- any two nodes are connected by at most one (undirected) path
- time and space cost of variable elimination are $\boxed{O(d^k n)}$

max k padri

Multiply connected networks:
- can reduce 3SAT to exact inference $\Rightarrow$ NP-hard
- equivalent to counting 3SAT models $\Rightarrow$ #P-complete

1. A v B v C
2. C v D v ¬A
3. B v C v ¬D