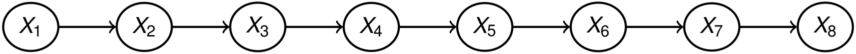
Belief Propagation in Bayesian Networks

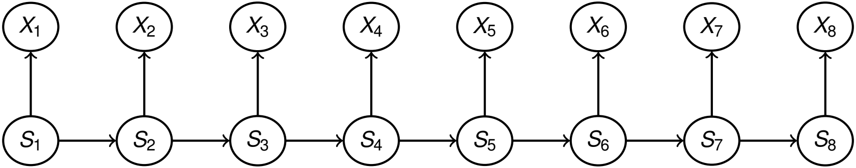
G. Nuel

INSMI, CNRS, Stochastics and Biology Group (PSB),
LPMA, UPMC, Sorbonne Universités, Paris, France

January 2017
Master Introduction Meeting
UPMC, Paris

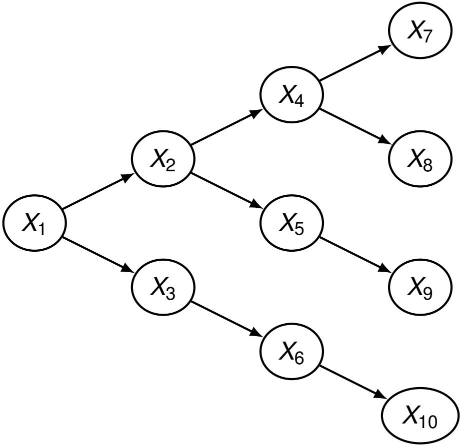


$$\mathbb{P}(X_1 \ldots X_8) = \mathbb{P}(X_1) \prod_{i=2}^{8} \mathbb{P}(X_i | X_{i-1})$$



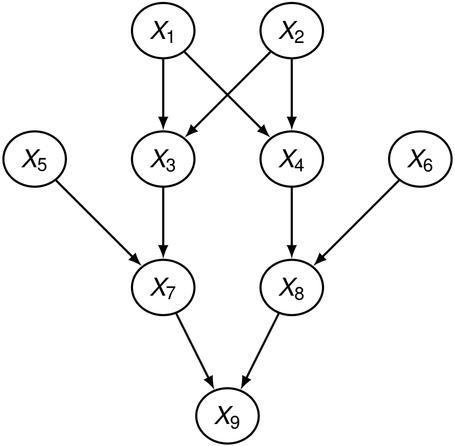
$$\mathbb{P}(X_1 \ldots X_8, S_1 \ldots S_8) = \mathbb{P}(S_1) \prod_{i=2}^{8} \mathbb{P}(S_i | S_{i-1}) \prod_{i=1}^{8} \mathbb{P}(X_i | S_i)$$

Markov Tree



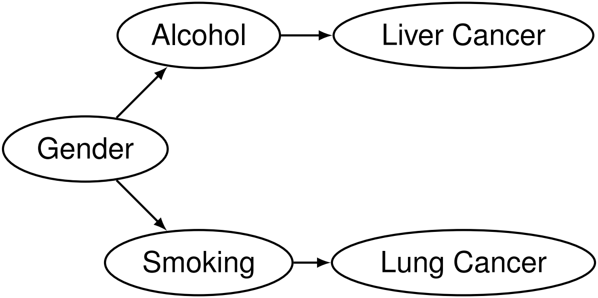
$$\mathbb{P}(X_1 \ldots X_{10}) = \mathbb{P}(X_1)\mathbb{P}(X_2|X_1)\mathbb{P}(X_3|X_1)\mathbb{P}(X_4|\ X_2)\mathbb{P}(X_5|\ X_2)\ldots$$

First Cousins Pedigree

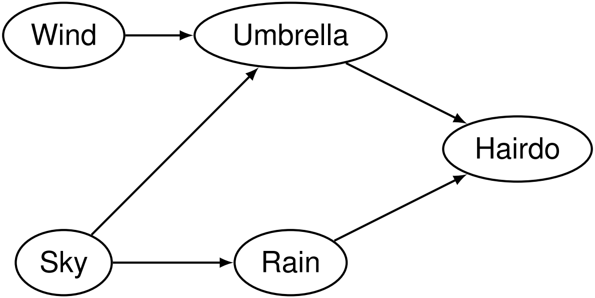


$$\mathbb{P}(X_1 \ldots X_9) = \mathbb{P}(X_1)\mathbb{P}(X_2)\mathbb{P}(X_3|X_{1,2})\mathbb{P}(X_4|\ X_{1,2})\mathbb{P}(X_5)\ldots$$

Cancer



$$\mathbb{P}(A, G, S, LivC, LunC) = \mathbb{P}(G)\mathbb{P}(A|\ G)\mathbb{P}(S|\ G)\mathbb{P}(LivC|\ A)\mathbb{P}(LunC|\ S)$$



$$\mathbb{P}(W,S,U,R,H)=\mathbb{P}(W)\mathbb{P}(S)\mathbb{P}(U|W,S)\mathbb{P}(R|\,S)\mathbb{P}(H|\,U,R)$$

General Bayesian Network

Definition (Bayesian Network): $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$ such that:

$$\mathbb{P}(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} \mathbb{P}\left(X_i | X_{\text{pa}(i)}\right)$$

with $\text{pa}(i) \subset \{1, \ldots, i-1\}$ (for a certain ordering).

Bayesian Networks $\equiv$ PGFs[1] on DAGs[2]

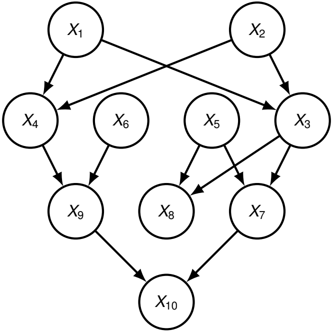
Purpose of this course: performing probabilistic computations

$$\mathbb{P}(\mathrm{ev}) = \sum_X \mathbb{P}(X, \mathrm{ev}) = \sum_X \prod_{i=1}^{n} \mathbb{P}\left(X_i, \mathrm{ev} | X_{\text{pa}(i)}\right)$$

for any *evidence* ev. NB: we hence also get: $\mathbb{P}(X|\mathrm{ev})$.

[1] Probabilistic Graphical Models
[2] Directed Acyclic Graphs

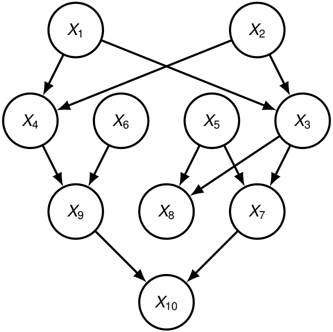
First cousins pedigree



$$\mathbb{P}(X_{1:10}) = \mathbb{P}(X_1)\mathbb{P}(X_2)\mathbb{P}(X_3|X_{1:2})\mathbb{P}(X_4|X_{1:2})\mathbb{P}(X_5)\mathbb{P}(X_6)$$
$$\mathbb{P}(X_7|X_{3,5})\mathbb{P}(X_8|X_{3,5})\mathbb{P}(X_9|X_{4,6})\mathbb{P}(X_{10}|X_{7,9})$$

$\Rightarrow X_i \in \{\mathtt{0}, \mathtt{1p}, \mathtt{1m}, \mathtt{2}\}$ gives $4^{10} = 1,048,576$ configurations

First cousins pedigree



$$\begin{aligned}\mathbb{P}(\text{ev}) &= \sum_X \prod_{i=1}^{10} \left\{ \sum_{Y_i} \mathbb{P}(Y_i, \text{ev} | X_i) \mathbb{P}(X_i, \text{ev} | X_{\text{pat}_i}, X_{\text{mat}_i}) \right\} \\ &= \sum_X \prod_{i=1}^{10} K_i(X_i, X_{\text{pat}_i}, X_{\text{mat}_i})\end{aligned}$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_5}\sum_{X_6}\sum_{X_7}\sum_{X_8}\sum_{X_9}\sum_{X_{10}} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})K_5(X_5)K_6(X_6) K_7(X_{3,5,7})K_8(X_{3,5,8})K_9(X_{4,6,9})K_{10}(X_{7,9,10})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_5}\sum_{X_6}\sum_{X_7}\sum_{X_8}\sum_{X_9}\sum_{X_{10}} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})K_5(X_5)K_6(X_6) K_7(X_{3,5,7})K_8(X_{3,5,8})K_9(X_{4,6,9})K_{10}(X_{7,9,10})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_5}\sum_{X_6}\sum_{X_7}\sum_{X_9}\sum_{X_{10}} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})K_5(X_5)K_6(X_6) K_7(X_{3,5,7})K_9(X_{4,6,9})K_{10}(X_{7,9,10})\sum_{X_8} K_8(X_{3,5,8})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_5}\sum_{X_6}\sum_{X_7}\sum_{X_9}\sum_{X_{10}} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})K_5(X_5)K_6(X_6) K_7(X_{3,5,7})K_9(X_{4,6,9})K_{10}(X_{7,9,10})M_1(X_{3,5})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_5}\sum_{X_6}\sum_{X_7}\sum_{X_9}\sum_{X_{10}} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})K_5(X_5)K_6(X_6) K_7(X_{3,5,7})K_9(X_{4,6,9})K_{10}(X_{7,9,10})M_1(X_{3,5})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_6}\sum_{X_7}\sum_{X_9}\sum_{X_{10}} K_1(X_1)K_2(X_2)$$

$$K_3(X_{1,2,3})K_4(X_{1,2,4})K_6(X_6)$$

$$K_9(X_{4,6,9})K_{10}(X_{7,9,10}) \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_6}\sum_{X_7}\sum_{X_9}\sum_{X_{10}} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})K_6(X_6) K_9(X_{4,6,9})K_{10}(X_{7,9,10})M_2(X_{3,7})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_6}\sum_{X_7}\sum_{X_9}\sum_{X_{10}} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})K_6(X_6) K_9(X_{4,6,9})K_{10}(X_{7,9,10})M_2(X_{3,7})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_6}\sum_{X_7}\sum_{X_9} K_1(X_1)K_2(X_2)$$
$$K_3(X_{1,2,3})K_4(X_{1,2,4})K_6(X_6)$$
$$K_9(X_{4,6,9})M_2(X_{3,7})\sum_{X_{10}} K_{10}(X_{7,9,10})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_6}\sum_{X_7}\sum_{X_9} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})K_6(X_6)K_9(X_{4,6,9})M_2(X_{3,7})M_3(X_{7,9})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

$$M_3(X_{7,9}) = \sum_{X_{10}} K_{10}(X_{7,9,10})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_6}\sum_{X_7}\sum_{X_9} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})K_6(X_6)K_9(X_{4,6,9})M_2(X_{3,7})M_3(X_{7,9})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

$$M_3(X_{7,9}) = \sum_{X_{10}} K_{10}(X_{7,9,10})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_6}\sum_{X_9} K_1(X_1)K_2(X_2)$$

$$K_3(X_{1,2,3})K_4(X_{1,2,4})K_6(X_6)K_9(X_{4,6,9}) \sum_{X_7} M_2(X_{3,7})M_3(X_{7,9})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

$$M_3(X_{7,9}) = \sum_{X_{10}} K_{10}(X_{7,9,10})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_6}\sum_{X_9} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})K_6(X_6)K_9(X_{4,6,9})M_4(X_{3,9})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

$$M_3(X_{7,9}) = \sum_{X_{10}} K_{10}(X_{7,9,10}) \quad M_4(X_{3,9}) = \sum_{X_7} M_2(X_{3,7})M_3(X_{7,9})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_6}\sum_{X_9} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})K_6(X_6)K_9(X_{4,6,9})M_4(X_{3,9})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

$$M_3(X_{7,9}) = \sum_{X_{10}} K_{10}(X_{7,9,10}) \quad M_4(X_{3,9}) = \sum_{X_7} M_2(X_{3,7})M_3(X_{7,9})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_9} K_1(X_1)K_2(X_2)$$
$$K_3(X_{1,2,3})K_4(X_{1,2,4})M_4(X_{3,9}) \sum_{X_6} K_6(X_6)K_9(X_{4,6,9})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

$$M_3(X_{7,9}) = \sum_{X_{10}} K_{10}(X_{7,9,10}) \quad M_4(X_{3,9}) = \sum_{X_7} M_2(X_{3,7})M_3(X_{7,9})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_9} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})M_4(X_{3,9})M_5(X_{4,9})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

$$M_3(X_{7,9}) = \sum_{X_{10}} K_{10}(X_{7,9,10}) \quad M_4(X_{3,9}) = \sum_{X_7} M_2(X_{3,7})M_3(X_{7,9})$$

$$M_5(X_{4,9}) = \sum_{X_6} K_6(X_6)K_9(X_{4,6,9})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4}\sum_{X_9} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4})M_4(X_{3,9})M_5(X_{4,9})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

$$M_3(X_{7,9}) = \sum_{X_{10}} K_{10}(X_{7,9,10}) \quad M_4(X_{3,9}) = \sum_{X_7} M_2(X_{3,7})M_3(X_{7,9})$$

$$M_5(X_{4,9}) = \sum_{X_6} K_6(X_6)K_9(X_{4,6,9})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4} K_1(X_1)K_2(X_2) K_3(X_{1,2,3})K_4(X_{1,2,4}) \sum_{X_9} M_4(X_{3,9})M_5(X_{4,9})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

$$M_3(X_{7,9}) = \sum_{X_{10}} K_{10}(X_{7,9,10}) \quad M_4(X_{3,9}) = \sum_{X_7} M_2(X_{3,7})M_3(X_{7,9})$$

$$M_5(X_{4,9}) = \sum_{X_6} K_6(X_6)K_9(X_{4,6,9})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4} K_1(X_1)K_2(X_2)K_3(X_{1,2,3})K_4(X_{1,2,4})M_6(X_{3,4})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

$$M_3(X_{7,9}) = \sum_{X_{10}} K_{10}(X_{7,9,10}) \quad M_4(X_{3,9}) = \sum_{X_7} M_2(X_{3,7})M_3(X_{7,9})$$

$$M_5(X_{4,9}) = \sum_{X_6} K_6(X_6)K_9(X_{4,6,9}) \quad M_6(X_{3,4}) = \sum_{X_9} M_4(X_{3,9})M_5(X_{4,9})$$

Variable elimination and messages

$$\mathbb{P}(\text{ev}) = \sum_{X_1}\sum_{X_2}\sum_{X_3}\sum_{X_4} K_1(X_1)K_2(X_2)K_3(X_{1,2,3})K_4(X_{1,2,4})M_6(X_{3,4})$$

$$M_1(X_{3,5}) = \sum_{X_8} K_8(X_{3,5,8}) \quad M_2(X_{3,7}) = \sum_{X_5} K_5(X_5)K_7(X_{3,5,7})M_1(X_{3,5})$$

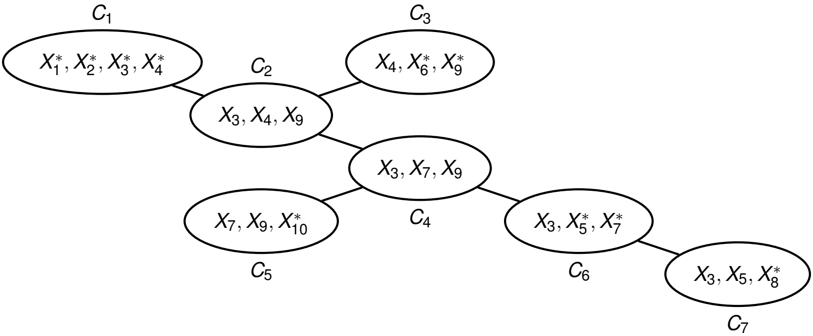
$$M_3(X_{7,9}) = \sum_{X_{10}} K_{10}(X_{7,9,10}) \quad M_4(X_{3,9}) = \sum_{X_7} M_2(X_{3,7})M_3(X_{7,9})$$

$$M_5(X_{4,9}) = \sum_{X_6} K_6(X_6)K_9(X_{4,6,9}) \quad M_6(X_{3,4}) = \sum_{X_9} M_4(X_{3,9})M_5(X_{4,9})$$

$\Rightarrow$ total $6 \times 4^3 + 4^4 = 640$ configurations
(*versus* $4^{10} = 1,048,576$ using brute force)

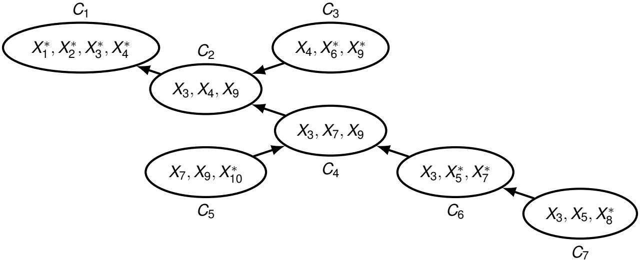
Belief Propagation

a compatible (and minimum) **junction tree**



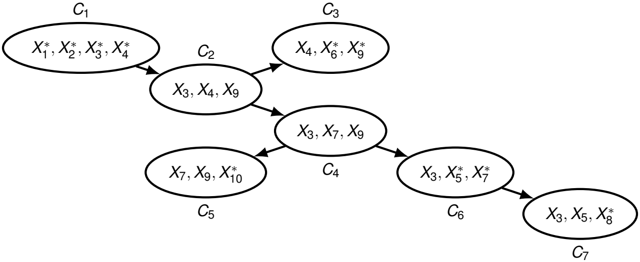
$\Rightarrow 6 \times 4^3 + 4^4 = 640$ configurations
($\textit{versus}$ $4^{10} = 1,048,576$ using brute force)

compute all **inward** messages



- $M_{7\to6}(X_3, X_5) = \sum_{X_8} K_8(X_3, X_5, X_8)$;
- $M_{6\to4}(X_3, X_7) = \sum_{X_5} K_5(X_5)K_7(X_3, X_5, X_7)M_{7\to6}(X_3, X_5)$;
- $M_{5\to4}(X_7, X_9) = \sum_{X_{10}} K_{10}(X_7, X_9, X_{10})$;
- $M_{4\to2}(X_3, X_9) = \sum_{X_7} M_{5\to4}(X_7, X_9)M_{6\to4}(X_3, X_7)$;
- $M_{3\to2}(X_4, X_9) = \sum_{X_6} K_6(X_6)K_9(X_4, X_6, X_9)$;
- $M_{2\to1}(X_3, X_4) = \sum_{X_9} M_{3\to2}(X_4, X_9)M_{4\to2}(X_3, X_9)$.

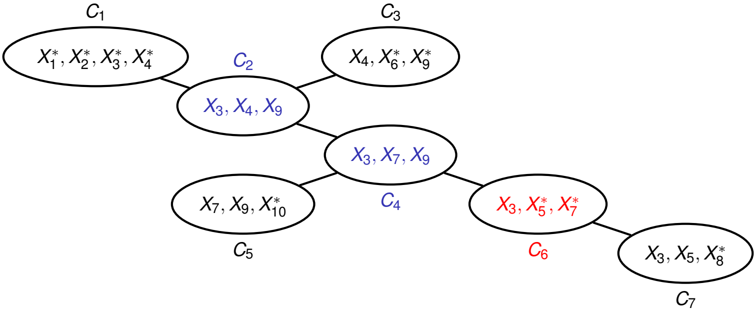
compute all **backward** messages



- $M_{1\to 2}(X_3,X_4) = \sum_{X_1,X_2} K_1(X_1)K_2(X_2)K_3(X_{1,2,3})K_4(X_{1,2,4})$;
- $M_{2\to 3}(X_4,X_9) = \sum_{X_3} M_{1\to 2}(X_3,X_4)\underline{M_{4\to 2}(X_3,X_9)}$;
- $M_{2\to 4}(X_3,X_9) = \sum_{X_4} M_{1\to 2}(X_3,X_4)\underline{M_{3\to 2}(X_4,X_9)}$;
- $M_{4\to 5}(X_7,X_9) = \sum_{X_3} M_{2\to 4}(X_3,X_9)\underline{M_{6\to 4}(X_3,X_7)}$;
- $M_{4\to 6}(X_3,X_7) = \sum_{X_9} M_{2\to 4}(X_3,X_9)\underline{M_{5\to 4}(X_7,X_9)}$;
- $M_{6\to 7}(X_3,X_5) = \sum_{X_7} K_5(X_5)K_7(X_{3,5,7})M_{4\to 6}(X_3,X_7)$.

Belief Propagation

use all computed **messages**



- $C_2 \cap C_4 = \{X_3, X_9\}$, $\mathbb{P}(X_3, X_9, \mathrm{ev}) = M_{2\to 4}(X_3, X_9) M_{4\to 2}(X_3, X_9)$
- $\mathbb{P}(\mathrm{ev}) = \sum_{X_3} \sum_{X_9} \mathbb{P}(X_3, X_9, \mathrm{ev})$
- $\mathbb{P}(X_3, X_5, X_7, \mathrm{ev}) = M_{4\to 6}(X_3, X_7) K_5(X_5) K_7(X_{3,5,7}) M_{7\to 6}(X_3, X_5)$
- $\mathbb{P}(\mathrm{ev}) = \sum_{X_3} \sum_{X_5} \sum_{X_9} \mathbb{P}(X_3, X_5, X_7, \mathrm{ev})$

Summary

Main focus: probabilistic computations such as:

$$\mathbb{P}(\text{ev}) = \sum_X \mathbb{P}(X, \text{ev}) = \sum_X \prod_{i=1}^{n} \mathbb{P}\left(X_i, \text{ev} | X_{\text{pa}(i)}\right)$$

Key features:

- Bayesian network, notion of evidence;
- sum-product algorithm, belief propagation;
- extensions to max-product or to pgf/mgf computations;
- *ad-hoc* proofs in particular cases (ex: HMM);
- illustrations in R with a dedicated library (`BNlib.R`).