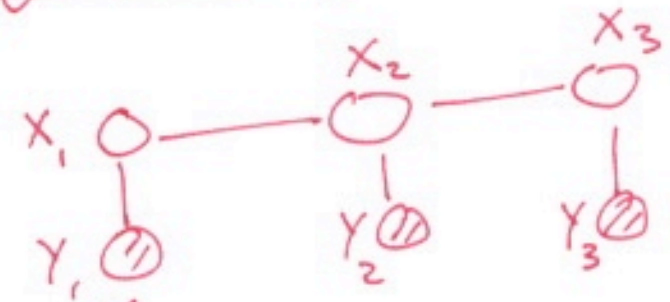
Message passing:



Shaded, cause we know the observation

$$P(X_1, X_2, X_3, Y_1 = y_1, \cdots)$$

$$= \frac{1}{Z} \left[\psi_1(X_1)\, \psi_2(X_2)\, \psi_3(X_3) \;\; \varphi_1(X_1, X_2)\, \varphi_2(X_2, X_3) \right]$$

$\psi_3(x_3) = \varphi_\cdot(x_3, Y_3 = \cdot)$

normalizing const

we want

$$P(X_1) = \sum_{X_2, X_3} P(X_1, X_2, X_3)$$

this is a marginal

computing by summing is clunky, cause exponential in # vars

But

$$\sum_{x_2 x_3}\left[p(x_1, x_2, x_3)\right]$$

$$= \frac{1}{Z}\sum_{x_2 x_3}\left[\begin{array}{l}\psi_1(x_1)\psi_2(x_2)\,\psi_3(x_3)\\ \varphi_1(x_1,x_2)\,\varphi_2(x_2,x_3)\end{array}\right]$$

$$= \frac{1}{Z}\,\psi_1(x_1)\underbrace{\sum_{x_2}\left[\psi_2(x_2)\,\varphi_1(x_1,x_2)\underbrace{\left\{\sum_{x_3}\psi_3(x_3)\,\varphi(x_2,x_3)\right\}}_{\text{"message" from } 3\to 2}\right]}_{\text{"message" from } 2\to 1}$$

$$p(x_1) = \frac{1}{Z}\,\psi_1(x_1)\,.\,m_{21}(x_1)$$

$$m_{21}(x_1) = \sum_{x_2}\left[\psi_2(x_2)\,\varphi_1(x_1,x_2)\,m_{32}(x_2)\right]$$

$$m_{32}(x_2) = \sum_{x_3}\left[\psi_3(x_3)\,\varphi_2(x_2,x_3)\right]$$

Notice

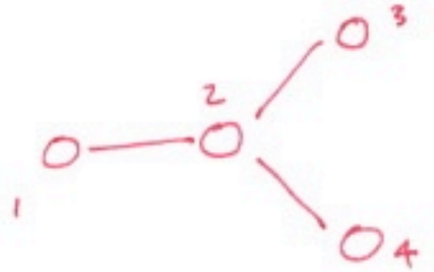
$$Z = \sum_{x_1}\psi_1(x_1)\,.\,m_{21}(x_1)$$

So we have the normalizing const, too

$$x_1 \text{ --- } x_2 \text{ --- } x_3$$

want
$$P(x_2) = \frac{1}{Z} \sum_{x_1, x_3} \left[\psi_1 \, \varphi_1 \, \psi_2 \, \varphi_2 \, \psi_3 \right]$$

$$= \frac{1}{Z} \left[\underbrace{\left[\sum_{x_1} \psi_1 \varphi_1 \right]}_{\substack{\text{"message" from} \\ 1 \to 2}} \psi_2 \underbrace{\left[\sum_{x_3} \varphi_2 \psi_3 \right]}_{\substack{\text{"message" from} \\ 3 \to 2}} \right]$$



$$p(x_2) = \frac{1}{Z} \sum_{x_1, x_3, x_4} \left[\psi_1 \varphi_1 \psi_2 \varphi_{23} \psi_3 \varphi_{24} \psi_4 \right]$$

$$= \frac{1}{Z} \left[\underbrace{\left[\sum_{x_1} \psi_1 \varphi_1 \right]}_{m_{1 \to 2}} \psi_2 \underbrace{\left[\sum_{x_3} \varphi_{23} \psi_3 \right]}_{m_{3 \to 2}} \underbrace{\left[\sum_{x_4} \varphi_{24} \psi_4 \right]}_{m_{4 \to 2}} \right].$$

We ~~know~~ now have a <u>general</u> procedure for finding marginals on trees

$$p(x_i) \propto \psi_i \prod_{j \in N(i)} m_{j \to i}(x_i)$$

$$m_{j \to i}(x_i) = \sum_{x_j} \psi_j \, \psi_{ij} \left[\prod_{k \in N(j) - i} m_{kj}(x_j) \right]$$

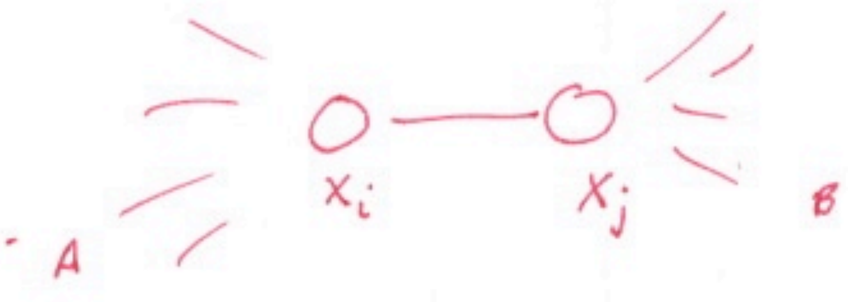
<u>notice this is a recursion</u>.

assume we want all marginals:

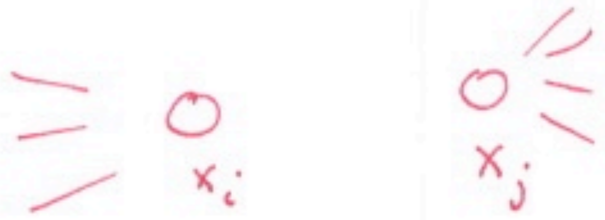
- choose a root
- orient all edges toward root
- compute rootward messages
  leaf → root
- now leafward messages
- <u>DONE</u>: you have all msgs, so can compute any marginal

Computing pairwise marginals

cases:



I



II

we've done II: x_i and x_j are conditionally indep given all others if there isn't an edge (why?)

I:

$$p(x_i, x_j) = \sum_A \sum_B p(\overbrace{\cdots}^{A \text{ vars}} x_i, x_j \overbrace{\cdots}^{B \text{ vars}})$$

$$= \frac{1}{Z} \left[\prod_{u \in N(i)-j} m_{u \to i} \right] \psi_i \, \psi_{ij} \, \psi_j \left[\prod_{v \in N(j)-i} m_{v \to j} \right]$$

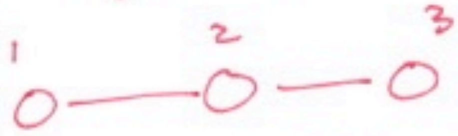
and Z is easy (PDF!)

MAP inference

want

$$\max_{x_1 \cdots x_N} P(X_1, \ldots X_N, Y_1 = y_i \cdots)$$

we can message pass



$$\max_{x_1 x_2 x_3} p = \max_{x_1 x_2 x_3} \frac{1}{Z} \left[\Psi_1 \, \varphi_1 \, \Psi_2 \, \varphi_2 \, \Psi_3 \right]$$

$$= \frac{1}{Z} \max_{x_1, x_2} \left[\Psi_1 \, \varphi_1 \, \Psi_2 \right] \underbrace{\left[\max_{x_3} \varphi_2 \, \Psi_3 \right]}$$

↳ cost to go function

- <u>like</u> a message but now max not sum.

$$x_i^{MAP} = \arg\max \left[\psi_i \prod_{u \in N(i)} m_{u \to i}^{MAP} \right]$$

$$m_{j \to i}^{MAP} = \max_{x_j} \left[\psi_j \, \varphi_{ij} \prod_{k \in N(j)-i} m_{k \to j}^{MAP} \right]$$

works for forest

- pass from leaves to root to get m^{MAP}'s
- from root to leaves to get x's

Idea :

- for non-trees, compute messages anyhow
  iterate until some kind of convergence

  "

  loopy belief propagation

- often rather successful, but what does
  it mean?