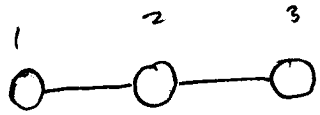


$$p(x_1, x_2, x_3) = \frac{1}{Z}\phi_1(x_1) \cdot \phi_2(x_2)\,\phi_3(x_3) \cdot \psi(x_1, x_2) \cdot \psi(x_2, x_3)$$

Q: What is $p(x_1)$?

A: $\sum_{x_2}\sum_{x_3} p(x_1, x_2, x_3)$ ← clunky, because becomes exponential in # of vars.

<u>But</u>

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

$$= \frac{1}{Z}\sum_{x_2}\sum_{x_3}\left[\phi_1(x_1) \cdot \phi_2(x_2)\,\phi_3(x_3)\,\psi(x_1, x_2)\,\psi(x_2, x_3)\right]$$

$$= \frac{1}{Z}\phi_1(x_1) \cdot \underbrace{\sum_{x_2}\left[\phi_2(x_2) \cdot \psi(x_1, x_2)\underbrace{\left\{\sum_{x_3}\phi_3(x_3) \cdot \psi(x_2, x_3)\right\}}_{\text{"message" from } 3 \to 2}\right]}_{\text{"message" from } 2 \to 1}$$

<u>So</u>:

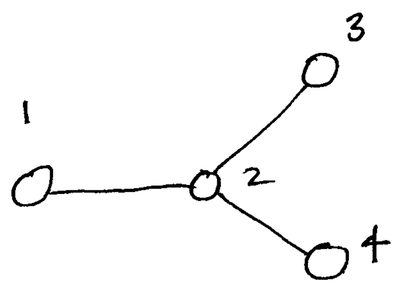
$$p(x_1) = \frac{1}{Z}\,\phi_1(x_1)\cdot m_{21}(x_1)$$

$$m_{21}(x_1) = \sum_{x_2}\left[\phi_2(x_2)\,\psi(x_1,x_2)\,m_{32}(x_2)\right]$$

$$m_{32}(x_2) = \sum_{x_3}\left(\phi_3(x_3)\cdot\psi(x_2,x_3)\right).$$

<u>But</u> this is a marginal, so ~~Z~~ ~~$\sum_{x_1}$~~

So $Z = \sum_{x_1} \phi_1(x_1)\, m_{21}(x_1)$



$$p(x_1) \cancel{\ne \frac{1}{Z}\sum} = \frac{1}{Z}\phi(x_1)\cdot\sum_{x_2}\left[\varphi(x_2)\,\psi(x_1,x_2)\left[\left[\sum_{x_3}\varphi(x_3)\,\psi(x_2,x_3)\right]\times\left[\sum_{x_4}\varphi(x_4)\,\psi(x_2,x_4)\right]\right]\right]$$

<u>equiv</u>:

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

$$m_{21}(x_1) = \sum_{x_2} \varphi(x_2)\,.\,\psi(x_1,x_2)\,.\prod_{j\in N(2)-1} m_{j2}(x_j)$$

$$m_{32}(x_2) = \sum_{x_3} \varphi(x_3)\,.\,\psi(x_2,x_3)$$

etc.

this gives a <u>general</u> process for finding marginals on trees.

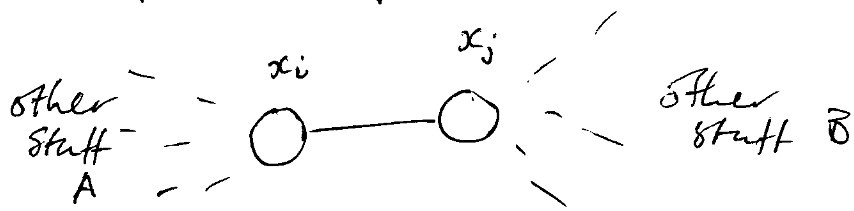
$$p(x_i) \propto \phi_i(x_i)\,.\prod_{j\in N(i)} m_{ji}(x_i)\,.$$

$$m_{ji}(x_i) \leftarrow \sum_{x_j} \phi_j(x_j)\,.\,\psi(x_i,x_j)\left[\prod_{k\in N(j)-i} m_{kj}(x_j)\right]$$

Now

- choose a root.
- orient all edges ~~root~~ ~~away~~ ~~from~~ toward root
- compute rootward messages, leaf → root.
- now leafward messages, root → leaf.
- Done - you have all msgs, so can compute any marginal on one var that amuses.

Modifying to compute pairwise marginals



$$p(x_i, x_j)$$

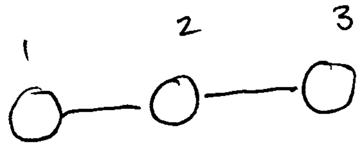
$$= \sum_{OSA} \sum_{OSB} p(\overbrace{\cdots}^{A \text{ vars}}, x_i, x_j, \overbrace{\cdots}^{B \text{ vars}})$$

$$= \frac{1}{Z} \left[\prod_{u \in N(i) - j} m_{ui}(x_i) \right] \phi_i(x_i) \, \psi_{ij}(x_i, x_j) \, \phi_j(x_j) \times$$

$$\left[\prod_{u \in N(j) - i} m_{uj}(x_j) \right]$$

(and Z is easy cause $p(x_i, x_j)$ is a PDF)

It is straightforward to do MAP inference like this, by modifying messages.



$$\max_{x_1, x_2, x_3} p(x_1, x_2, x_3)$$

$$= \max \frac{1}{Z} \phi_1(x_1) \psi(x_1, x_2) \phi(x_2) \psi(x_2, x_3) \varphi(x_3)$$

$$= \max_{x_1, x_2} \left[\phi_1(x_1) \psi(x_1, x_2) \phi(x_2) \underbrace{\left\{ \max_{x_3} \psi(x_2, x_3) \varphi(x_3) \right\}}_{\text{cost to go function equiv, message}} \right]$$

cost to go function equiv, <u>message</u>

<u>Not</u> same as previous message (max vs sum).

So we get.

$$x_i^{MAP} = \arg\max \left[\phi_i(x_i) . \prod_{u \in N(i)} M_{ui}^{MAP}(x_i) \right]$$

$$M_{ji}^{MAP}(x_i) \leftarrow \max_{x_j} \left[\phi_j(x_j) . \Psi(x_i, x_j) \left[\prod_{k \in N(j) - i} M_{kj}^{MAP}(x_j) \right] \right]$$

again, tree works great

- pass from leaves to root to compute M^{map}'s
- " " root to leaves to get x's.

Idea:

- for non-trees, compute messages anyhow, iterate until it all settles down

  ||

- loopy belief propagation

- often rather successful, but what does it mean?