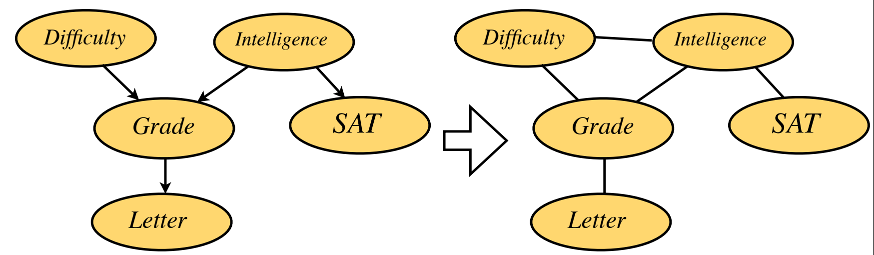
From directed to undirected

- Complete algorithm
  - Make the graph moral.
  - Initialize maximal clique potential to one.
  - Multiply each factor in p() into an appropriate clique potential.
  - Note that Z=1

Example of converting directed to undirected



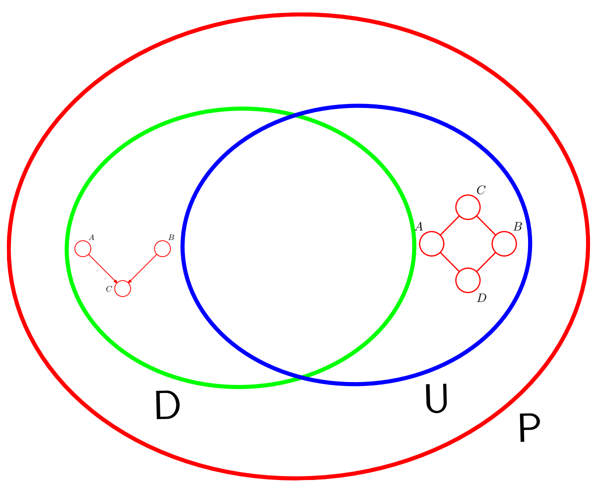
$$P(I,D,G,L,S) = P(I)P(D)P(G|I,D)P(L|G)P(S|I)$$

$$P = \psi(D,G,I)\psi(S,I)\psi(L,G)$$

$$\psi(D,G,I) = P(I)P(D)P(G|I,D) \quad \psi(S,I) = P(S|I) \quad \psi(L,G) = P(L|G)$$

$$\psi(D,G,I) = P(D)P(G|I,D) \quad \psi(S,I) = P(I)P(S|I) \quad \psi(L,G) = P(L|G)$$

Directed and undirected perfect maps



Inference on graphs

- Given a graph and its conditionals or potentials compute

$p(\theta|e)$ (particular θ and e, marginalizing out other variables)

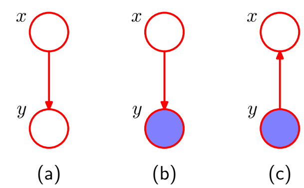
$p(X)$ (particular event, marginalizing out other variables)

$\operatorname{argmax} p(\theta|e)$ (particular θ and e; marginalizing other variables)

$\operatorname{argmax} p(\theta,\theta_N,e_N|e)$ (all variables, will nuiscence / unobserved)

Inference on graphs

- Simplest example (Bayes' rule)
  - (a) model
  - (b) illustrates observed
  - (c) inference **reverses** the arrow



- Computationally

$$p(x|y) = \frac{p(y|x)\,p(x)}{\sum_{x'} p(y|x')\,p(x')}$$

Marginals on a chain

Recall

$$p(\mathrm{x}) = p(x_1)\,p(x_2|x_1)\,p(x_3|x_2)\ \cdots\ p(x_{N-1}|x_{N-2})\,p(x_N|x_{N-1})$$

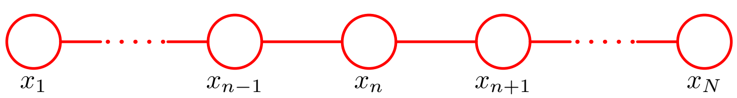
Converted to

$$p(\mathrm{x}) = \psi_{1,2}(x_1,x_2)\,\psi_{2,3}(x_2,x_3)\ \ldots\ \psi_{N-2,N-1}(x_{N-2},x_{N-1})\,\psi_{N-1,N}(x_{N-1},x_N)$$

Assume N discrete variables, with K values each.

Compute the marginal of a node in the middle, $p(x_n)$

Marginals on a chain



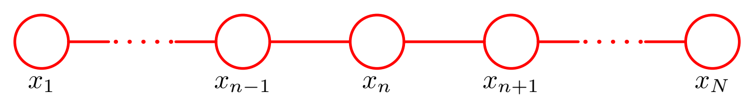
$$p(\mathrm{x}) = \psi_{1,2}(x_1,x_2)\,\psi_{2,3}(x_2,x_3)\ \ldots\ \psi_{N-2,N-1}(x_{N-2},x_{N-1})\,\psi_{N-1,N}(x_{N-1},x_N)$$

Direct calculation of $p(x_n)$

$$p(x_n) = \sum_{x_1}\sum_{x_2}\ \ldots\ \sum_{x_{n-1}}\sum_{x_{n+1}}\ \ldots\ \sum_{x_{N-1}}\sum_{x_N} p(\mathbf{x})$$

Skip x_n

Marginals on a chain

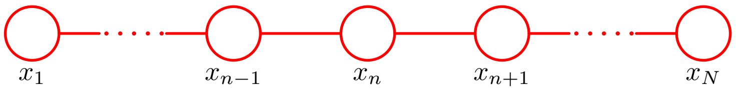


$$p(\mathrm{x}) = \psi_{1,2}(x_1,x_2)\,\psi_{2,3}(x_2,x_3)\ \ldots\ \psi_{N-2,N-1}(x_{N-2},x_{N-1})\,\psi_{N-1,N}(x_{N-1},x_N)$$

Direct calculation of $p(x_n)$

$$p(x_n) = \sum_{x_1}\sum_{x_2}\ \ldots\ \sum_{x_{n-1}}\sum_{x_{n+1}}\ \ldots\ \sum_{x_{N-1}}\sum_{x_N} p(\mathbf{x})$$

Computational complexity is $O(K^N)$. Way too slow!



$$p(\mathrm{x}) = \psi_{1,2}(x_1,x_2)\psi_{2,3}(x_2,x_3) \quad \dots \quad \psi_{N-2,N-1}(x_{N-2},x_{N-1})\psi_{N-1,N}(x_{N-1},x_N)$$

Main idea is to rearrange terms to exploit conditional independence.

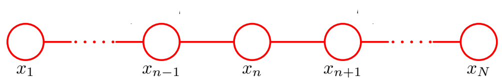
Fancy formulas from algebra

$$\underbrace{\sum_{x_1,\, x_2,\, \dots,\, x_N}}_{\text{all values of each}} f(x_1, x_2, \dots, x_N) = \underbrace{\sum_{x_1}\sum_{x_2}\dots\sum_{x_N}}_{\text{any order you like}} f(x_1, x_2, \dots, x_N)$$

(essentially a definition)

$$\left(\sum a_i\right)\left(\sum b_j\right) = \sum\sum a_i b_j$$

Back to marginals on a chain



$$p(\mathrm{x}) = \psi_{1,2}(x_1,x_2)\psi_{2,3}(x_2,x_3) \quad \dots \quad \psi_{N-2,N-1}(x_{N-2},x_{N-1})\psi_{N-1,N}(x_{N-1},x_N)$$

$$\begin{aligned}
p(x_n) &= \sum_{x_1}\sum_{x_2}\dots\sum_{x_{n-1}}\sum_{x_{n+1}}\dots\sum_{x_{N-1}}\sum_{x_N} f_L(x_1,x_1,\dots,x_n) f_R(x_n,x_{n+1},\dots,x_N) \\
&= \left(\sum_{x_1}\sum_{x_2}\dots\sum_{x_{n-1}} f_L(x_1,x_1,\dots,x_n)\right)\left(\sum_{x_{n+1}}\dots\sum_{x_{N-1}}\sum_{x_N} f_R(x_n,x_{n+1},\dots,x_N)\right) \\
&= \left(\sum_{x_1}\sum_{x_2}\dots\sum_{x_{n-1}}\prod_{i=1}^{n-1}\psi_{i,i+1}(x_i,x_{i+1})\right)\left(\sum_{x_{n+1}}\dots\sum_{x_{N-1}}\sum_{x_N}\prod_{i=n}^{N-1}\psi_{i,i+1}(x_i,x_{i+1})\right) \\
&= \left(\sum_{x_{n-1}}\sum_{x_{n-2}}\dots\sum_{x_1}\prod_{i=1}^{n-1}\psi_{n-i-1,n-i}(x_{n-i},x_{n-i+1})\right)\left(\sum_{x_{n+1}}\dots\sum_{x_{N-1}}\sum_{x_N}\prod_{i=n}^{N-1}\psi_{i,i+1}(x_i,x_{i+1})\right)
\end{aligned}$$

$$p(x_n) = \left(\sum_{x_{n-1}}\sum_{x_{n-2}}\dots\sum_{x_1}\prod_{i=1}^{n-1}\psi_{n-i,\,n-i+1}(x_{n-i},x_{n-i+1})\right)\left(\sum_{x_{n+1}}\dots\sum_{x_{N-1}}\sum_{x_N}\prod_{i=n}^{N-1}\psi_{i,\,i+1}(x_i,x_{i+1})\right)$$

$$\prod_{i=1}^{n-1}\psi_{n-i,\,n-i+1}(x_{n-i},x_{n-i+1}) = \psi_{n-1,\,n}(x_{n-1},x_n)\psi_{n-2,\,n-1}(x_{n-2},x_{n-2}) \quad \dots \quad \psi_{2,\,3}(x_2,x_3)\psi_{1,\,2}(x_1,x_2)$$

Warmup example

$$\sum_{x_2}\sum_{x_1} \underbrace{\psi(x_2,x_3)}_{\substack{\text{No dependency on } x_1 \\ \text{hence we can move} \\ \text{factor outside sum} \\ \text{over } x_1 .}} \psi(x_1,x_2) = \sum_{x_2}\psi(x_2,x_3)\sum_{x_1}\psi(x_1,x_2)$$

(Recall the distibutive law: $ba + ca = a(b+c)$)

$$p(x_n)=\left(\sum_{x_{n-1}}\sum_{x_{n-2}}\cdots\sum_{x_1}\prod_{i=1}^{n-1}\psi_{n-i,\,n-i+1}(x_{n-i},x_{n-i+1})\right)\left(\sum_{x_{n+1}}\cdots\sum_{x_{N-1}}\sum_{x_N}\prod_{i=n}^{N-1}\psi_{i,\,i+1}(x_i,x_{i+1})\right)$$

$$\prod_{i=1}^{n-1}\psi_{n-i,\,n-i+1}(x_{n-i},x_{n-i+1})=\psi_{n-1,\,n}(x_{n-1},x_n)\,\psi_{n-2,\,n-1}(x_{n-2},x_{n-2})\;\cdots\;\psi_{2,\,3}(x_2,x_3)\,\psi_{1,\,2}(x_1,x_2)$$

$$\begin{aligned}\sum_{x_{n-1}}\sum_{x_{n-2}}\cdots\sum_{x_1}\prod_{i=1}^{n-1}\psi_{n-i,\,n-i+1}(x_{n-i},x_{n-i+1})&=\sum_{x_{n-1}}\sum_{x_{n-2}}\cdots\sum_{x_2}\prod_{i=1}^{n-2}\psi_{n-i,\,n-i+1}(x_{n-i},x_{n-i+1})\left\{\sum_{x_1}\psi_{1,2}(x_1,x_2)\right\}\\&=\sum_{x_{n-1}}\sum_{x_{n-2}}\cdots\sum_{x_3}\prod_{i=1}^{n-3}\psi_{n-i,\,n-i+1}(x_{n-i-1},x_{n-i})\left\{\sum_{x_2}\psi_{2,3}(x_2,x_3)\left\{\sum_{x_1}\psi_{1,2}(x_1,x_2)\right\}\right\}\\&\cdots\\&=\left\{\sum_{x_{n-1}}\psi_{n-1,n}(x_{n-1},x_n)\cdots\left\{\sum_{x_3}\psi_{3,4}(x_3,x_4)\left\{\sum_{x_2}\psi_{2,3}(x_2,x_3)\left\{\sum_{x_1}\psi_{1,2}(x_1,x_2)\right\}\right\}\right\}\right\}\end{aligned}$$

$$p(x_n)=\left(\sum_{x_{n-1}}\sum_{x_{n-2}}\cdots\sum_{x_1}\prod_{i=1}^{n-1}\psi_{n-i,\,n-i+1}(x_{n-i},x_{n-i+1})\right)\left(\sum_{x_{n+1}}\cdots\sum_{x_{N-1}}\sum_{x_N}\prod_{i=n}^{N-1}\psi_{i,\,i+1}(x_i,x_{i+1})\right)$$

$$\sum_{x_{n-1}}\sum_{x_{n-2}}\cdots\sum_{x_1}\prod_{i=1}^{n-1}\psi_{n-i,n-i+1}(x_{n-i},x_{n-i+1})=\left\{\sum_{x_{n-1}}\psi_{n-1,n}(x_{n-1},x_n)\cdots\left\{\sum_{x_3}\psi_{3,4}(x_3,x_4)\left\{\sum_{x_2}\psi_{2,3}(x_2,x_3)\left\{\sum_{x_1}\psi_{1,2}(x_1,x_2)\right\}\right\}\right\}\cdots\right\}$$

and

$$\sum_{x_{n+1}}\cdots\sum_{x_{N-1}}\sum_{x_N}\prod_{i=n}^{N-1}\psi_{i,\,i+1}(x_i,x_{i+1})=\left\{\sum_{x_{n+1}}\psi_{n,n+1}(x_n,x_{n+1})\cdots\left\{\sum_{x_{N-1}}\psi_{N-2,N-1}(x_{N-2},x_{N-1})\left\{\sum_{x_N}\psi_{N-1,\,N}(x_{N-1},x_N)\right\}\right\}\cdots\right\}$$

(Deriving the right factor (red) is similar.)

Matrix interpretation (for two variables)

$$\sum_{x_{n-1}}\sum_{x_{n-2}}\cdots\sum_{x_1}\prod_{i=1}^{n-1}\psi_{n-i,n-i+1}(x_{n-i},x_{n-i+1})=\left\{\sum_{x_{n-1}}\psi_{n-1,n}(x_{n-1},x_n)\cdots\left\{\sum_{x_3}\psi_{3,4}(x_3,x_4)\left\{\sum_{x_2}\psi_{2,3}(x_2,x_3)\left\{\sum_{x_1}\psi_{1,2}(x_1,x_2)\right\}\right\}\right\}\cdots\right\}$$

$\psi_{i,\,i+1}(x_i,x_{i+1})\Leftrightarrow Q_{i+1,\,i}$ (note transposition!)

$\sum_i \psi_{i,\,i+1}(x_i,x_{i+1})\Leftrightarrow$ sums colums to get a vector V_{i+1}

$$\sum_{x_{i+1}}\psi_{i+1,\,i+2}(x_{i+1},x_{i+2})\left\{\sum_{x_i}\psi_{i,\,i+1}(x_i,x_{i+1})\right\}\Leftrightarrow Q_{i+2,\,i+1}\bullet V_i$$

Computational Complexity

Suppose each variable has K values

What is the cost of evaluating the first factor?

$$\sum_{x_{n-1}}\sum_{x_{n-2}}\cdots\sum_{x_1}\prod_{i=1}^{n-1}\psi_{n-i,n-i+1}(x_{n-i},x_{n-i+1})=\left\{\sum_{x_{n-1}}\psi_{n-1,n}(x_{n-1},x_n)\cdots\left\{\sum_{x_3}\psi_{3,4}(x_3,x_4)\underbrace{\left\{\sum_{x_2}\underbrace{\psi_{2,3}(x_2,x_3)\left\{\underbrace{\sum_{x_1}\psi_{1,2}(x_1,x_2)}_{\text{K sums of K values}}\right\}}_{\text{K evaluations of K products}}\right\}}_{\text{K sums of K values}}\right\}\cdots\right\}$$

The other factor is similar.

We see that the overall cost is $O(N\cdot K^2)$.

Much better than the naive computation where we had $O(K^N)$.