Inference queries

1. *Marginal probability queries over a small subset of variables:*
   - Find Pr(Income='High & Degree='PhD')
     x_1 x_2
   - Find Pr(pixel $y_9 = 1$)

MRF

$$\Pr(x_1) = \sum_{x_2 \ldots x_n} \Pr(x_1 \ldots x_n) = \frac{1}{Z} \prod_C \psi_C(x_C)$$

$$\prod_i P(x_i \mid Pa(x_i))$$

$$= \sum_{x_2=1}^{m} \ldots \sum_{x_n=1}^{m} \Pr(x_1 \ldots x_n)$$

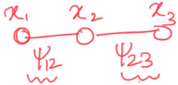
Brute-force requires $O(m^{n-1})$ time.

2. *Most likely labels of remaining variables:* (MAP queries)
   - Find most likely entity labels of all words in a sentence
   - Find likely temperature at sensors in a room

$$\mathbf{x}^* = \operatorname{argmax}_{x_1 \ldots x_n} \Pr(x_1 \ldots x_n)$$

Example of exact inference

$$P(x_1,x_2,x_3) = \frac{1}{Z}\psi_{12}(x_1,x_2)\,\psi_{23}(x_2,x_3)$$



$x_i \in \{0,1\}$ $\quad m=2$

x^*

$$\cancel{P(x_1)} = \frac{1}{Z}\overset{\max}{\cancel{\sum_{x_1\in\{0,1\}}}}\ \overset{\max}{\sum_{x_2}}\ \overset{\max}{\cancel{\sum_{x_3\in\{0,1\}}}}\big(\psi_{12}(x_1,x_2)\,\psi_{23}(x_2,x_3)\big)$$

$$\psi_{123}(x_1,x_2,x_3)$$

$$\overset{\max}{\cancel{\sum_{x_1}}}\ \frac{1}{Z}\ \overset{\max}{\cancel{\sum_{x_2}}}\ \overset{\max}{\cancel{\sum_{x_3}}}$$

$$\frac{1}{Z}\sum_{x_2}\underbrace{\psi^*_{12}(x_1,x_2)}_{\psi^{\max}_{12}(x_1,x_2)}$$

$$\frac{1}{Z}\underbrace{\psi^*_1(x_1)}_{\psi^{\max}_1(x_1)} = \begin{array}{c|c} 0 & 76 \\ 1 & 44 \end{array}$$

$Z = 76+44$

$\psi_{12}(x_1,x_2)$

$x_1 \backslash x_2$	0	1
0	$\psi(0,0)$	$\psi(0,1)$
1	$\psi(1,0)$	$\psi(1,1)$

$x_1 \backslash x_2$	0	1
0	5	2
1	1	4

$x_2 \backslash x_3$	0	1
0	2	10
1	5	3

$x_3 = 0$

$x_1 \backslash x_2$	0	1
0	5×2	2×5
1	1×2	4×5

$x_3 = 1$

$x_1 \backslash x_2$	0	1
0	5×10	6
1	1×10	4×3

$\psi_{12}(0,0)\,\psi_{23}(0,0)$

$\psi_{12}(0,0)\,\psi_{23}(0,1)$

$x_1 \backslash x_2$	0	1
0	60	16
1	12	32

$\overleftarrow{x_2}$

50 $x_3=1$	10 $x_3=0$
10 $x_3=1$	20 $x_3=0$

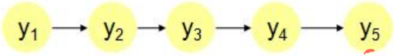
$P(x_1=0) = \frac{76}{Z}$; $P(x_1=1) = \frac{44}{Z}$

$$\begin{array}{c} 0 \\ 1 \end{array}\begin{bmatrix} 50 & x_2=0,\ x_3=1 \\ 20 & x_2=1,\ x_3=0 \end{bmatrix}$$

Example of exact inference

Example of exact inference

Exact inference on chains

- Given,
  - Graph 

  - Potentials: $\psi_i(y_i, y_{i+1})$
  - $Pr(y_1, \ldots y_n) = \prod_i \psi_i(y_i, y_{i+1}), Pr(y_1)$
- Find, $\Pr(y_i)$ for any i, say $\Pr(y_5 = 1)$
  - Exact method: $\Pr(y_5 = 1) = \sum_{y_1,\ldots y_4} Pr(y_1, \ldots y_4, 1)$ requires exponential number of summations.
  - A more efficient alternative...

Exact inference on chains

Graph: $\overset{=5}{n}$ variables, m values for each:
Time required by this algorithm?

$$\Pr(y_5 = 1) = \sum_{y_1,\ldots y_4} Pr(y_1, \ldots y_4, 1)$$

DOUBT

$$= \sum_{y_1}\sum_{y_2}\sum_{y_3}\sum_{y_4} \psi_1(y_1,y_2)\psi_2(y_2,y_3)\psi_3(y_3,y_4)\psi_4(y_4,1) \quad \boxed{O(nm^2)}$$

$$<<$$

$$= \sum_{y_1}\sum_{y_2}\psi_1(y_1,y_2)\sum_{y_3}\underbrace{\psi_2(y_2,y_3)}_{m\ \ m}\sum_{y_4}\psi_3(y_3,y_4)\psi_4(y_4,1) \quad O(m^n)$$

$$= \sum_{y_1}\sum_{y_2}\psi_1(y_1,y_2)\sum_{y_3}\psi_2(y_2,y_3)\boxed{B_3(y_3)}$$

$$= \sum_{y_1}\sum_{y_2}\psi_1(y_1,y_2)B_2(y_2)$$

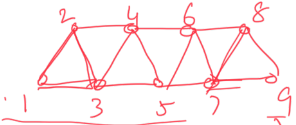
$$= \sum_{y_1} B_1(y_1)$$

An alternative view: flow of beliefs $B_i(.)$ from node $i+1$ to node i

$$y_1 \to y_2 \to y_3 \to y_4 \to y_5$$

More examples of efficient inference

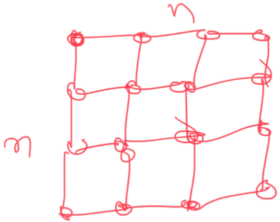
$P(x_1)$



$$\sum_{x_2} \cdots \sum_{x_8} \sum_{x_9} \psi_{123}(x_1, x_2, x_3)\, \psi_{234}(x_2, x_3, x_4) \cdots \psi_{789}(x_7, x_8, x_9)$$

$$O(nm^3) << O(m^n)$$

$P(x_1)$?



Hardness of Inference

Given a graphical model $P(x_1, \ldots x_n)$ which is factorized efficiently in terms of potentials (e.g., polynomial number of potentials, with each potential containing a constant number of variables), can we always find $P(x_i)$ or Z in polynomial time?

Proof.

No. Reduce 3-SAT to inference in Bayesian networks. (Theorem 9.1 of KF textbook) □

The grid graph is an example of such a graph.

Proof of Hardness

Define: 3-SAT problem:

Given n Boolean variables: $x_1, x_2 \ldots x_n$ $x_i \in \{T, F\}$.

Literal: x_i, $\neg x_i$ $\bar{x}_i$

A set of K clauses: $C_1, C_2 \ldots C_k$

Each clause $C_j = l_{j_1} \vee l_{j_2} \vee l_{j_3}$

3-SAT problem is to decide if $\exists$ an assignment of values to the n-variables so that

$C_1 \wedge C_2 \wedge \ldots C_k = \text{True}$.

Example: $n=4$, $k=3$

x_1, x_2, x_3, x_4

$C_1 \wedge C_2 \wedge C_3$

$C_1 = x_1 \vee \bar{x}_2 \vee \bar{x}_3$ T

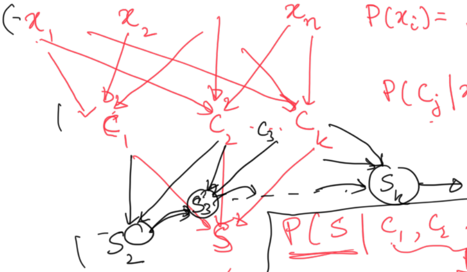
$C_2 = x_2 \vee x_3 \vee \bar{x}_4$ T

$C_3 = x_4 \vee \bar{x}_1 \vee \bar{x}_2$ T

x
(T, T, T, T)

Proof of Hardness

Represent 3-SAT as a BN.



$P(x_i) =$

0	1
1/2	1/2

$C_1 = x_1 \vee \bar{x}_2 \vee \bar{x}_3$

$P(C_j \mid x_{j_1}, x_{j_2}, x_{j_3}) =$

x_1	x_2	x_3	$P(C_j = 1 \mid x_{j_1} x_{j_2} x_{j_3})$
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	0
⋮			

8

$P(S \mid C_1, C_2 \ldots C_k)$

$C_1 \ldots C_k$	$P(S=1 \mid C_1 \cdots C_k)$
0 0 0-0	0
1	0
⋮	⋮
1 1 1 1 1 1	1

$2^k \Rightarrow$ not polynomial

$P(S_j \mid S_{j-1}, C_j, C_{j-1})$

$S_j = S_{j-1} \wedge C_j \wedge C_{j-1}$

$P(S_j = 1) > 0$? Yes → 3-SAT assignment exists else it does not.