Recap

- Part 1. Probabilistic Graphical Models, Markov Properties ✓
- Part 2. Inference Problems ┬ efficient algorithms.
- part 3. └ Variational inference & Sampling.
- part 4. Learning Graphical models.
- part 5. Extra topics.

Today: Efficient algorithms for approximately solving inference problems

Inference problems:

given $P(x)$ with an undirected graphical model G

$$P(x) = \frac{1}{Z} \prod_{c \in \mathcal{C}} f_c(x_c)$$

set of maximal cliques

① compute marginal $P(x_1)$

② compute MAP: $\arg\max_{x \in \mathcal{X}^n} P(x)$ ← $P(x) \Leftarrow \mathbb{P}(X|Y)$ G.M

③ compute partition function Z

④ sample from $P(x)$

example> Y ~ observation, X ~ cause/state, $P(x_1)$ ~ denoise estimation

noisy observation	pixels	denoising $\mathbb{E}[X_1 \mid Y]$
MRI measurements	human brain	estimation
noisy received bits	bits to send	decoding BER (bit error rate)

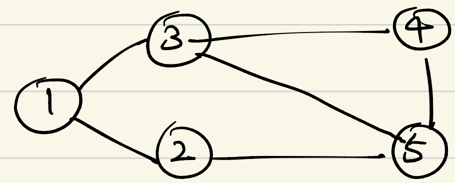
MARG ┬ decoding BER (bit error rate)
$P(X_i|Y)$
└ decoding Block error rate

MAP $\arg\max P(x|y)$

* Elimination Algorithm for Marginalization

(Exact but can take $O(|\mathcal{X}|^{n})$ operations) — n "=" $|V|$; $\mathcal{X}$: alphabet of X_i's

$$P(x) = \frac{1}{Z} f_{12}(x_1,x_2)\, f_{13}(x_1,x_3)\, f_{25}(x_2,x_5)\, f_{345}(x_3,x_4,x_5)$$



Goal: $P(x_1)$

brute force: $\underset{\text{summation}}{|\mathcal{X}|^4} \cdot \underset{\text{enumerate } x_1}{|\mathcal{X}|}$

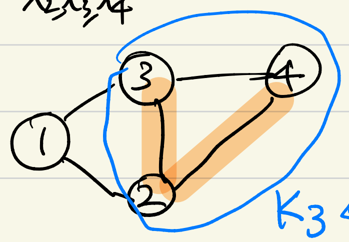
$|\mathcal{X}|^n$

consider ordering $\overrightarrow{(5,4,3,2,1)}$

$$P(x_1) = \frac{1}{Z} \sum_{x_2,x_3,x_4} f_{12}(x_1,x_2)\, f_{13}(x_1,x_3) \underbrace{\boxed{\sum_{x_5} f_{25}(x_2,x_5)\, f_{345}(x_3,x_4,x_5)}}_{m_5(x_2,x_3,x_4)}$$

$\underset{\text{enum}}{|\mathcal{X}|^3} \cdot \underset{\text{sum}}{|\mathcal{X}|}$

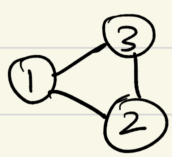
G_1:



$$= \frac{1}{Z} \sum_{x_2,x_3} f_{12}(x_1,x_2)\, f_{13}(x_1,x_3) \underbrace{\sum_{x_4} m_5(x_2,x_3,x_4)}_{m_4(x_2,x_3)}$$

$\underset{\text{enum}}{|\mathcal{X}|^2} \cdot \underset{\text{sum}}{|\mathcal{X}|}$

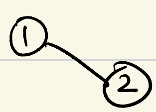
G_2:



$$= \frac{1}{Z} \sum_{x_2} f_{12}(x_1,x_2) \underbrace{\sum_{x_3} f_{13}(x_1,x_3)\, m_4(x_2,x_3)}_{m_3(x_1,x_2)}$$

$|\mathcal{X}|^2 \cdot |\mathcal{X}|$

G_3:



$$= \frac{1}{Z} \underbrace{\sum_{x_2} f_{12}(x_1,x_2)\, m_3(x_1,x_2)}_{m_2(x_1)}$$

$|\mathcal{X}| \cdot |\mathcal{X}|$

$\left.\begin{bmatrix} \\ \\ \end{bmatrix}\right\} |\mathcal{X}| \rightarrow P(x_1) = \dfrac{m_2(x_1)}{\sum_{x_1'} m_2(x_1')}$

$O(n \cdot |\mathcal{X}|^{TW})$ = Total: $|\mathcal{X}|^4 + 2|\mathcal{X}|^3 + \ldots$

tighter: $O(\sum |\mathcal{X}|^{|C|})$

vs. brute $|\mathcal{X}|^5$

Def. Treewidth is $\min_{\text{all orderings}}$ { maximum size of a clique in reconstituted graph }

size of largest node/vertex in a tree decomposition

(5,4,3,2,1) → 4

Def. Reconstituted graph of $G=(V,E)$

$\tilde{G} = (V, \tilde{E})$ w.r.t elimination ordering $\rho = (\rho_1, \cdots, \rho_n)$

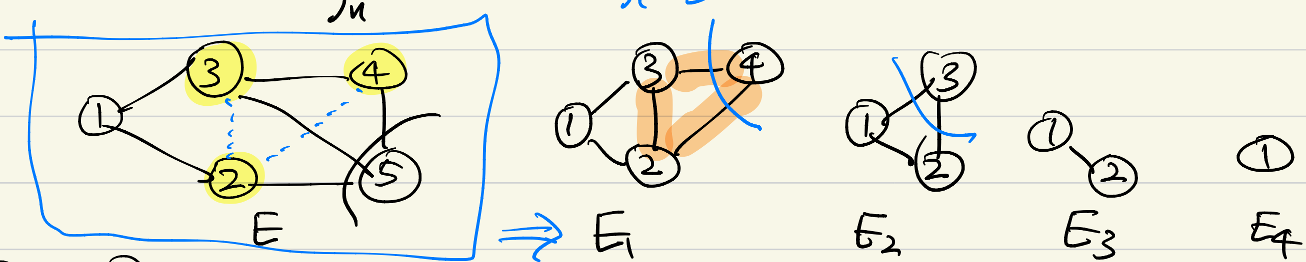
Remove ρ_1 and $E_1 = \{(i,j) \mid i,j$ are in neighborhood of ρ_1 in the remaining graph $\}$

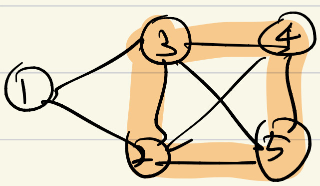
⋮

ρ_n

bottle neck

$\rho_1 = 5$

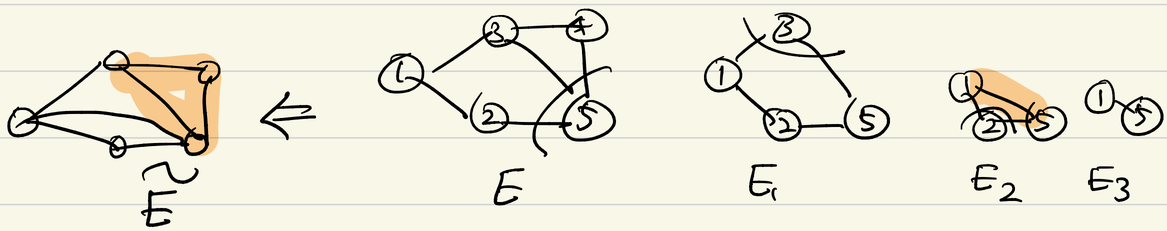




$\tilde{E} = E \cup E_1 \cup E_2 \cdots \cup E_n$

max clique size = 4 → $|\mathcal{X}|^4$

(4, 3, 2, 5, 1)

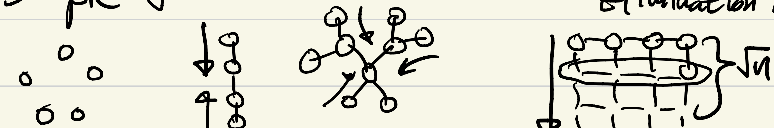


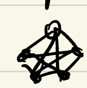
max clique size = 3 → $|\mathcal{X}|^3$

{1, ⋯, n}

Small ← Treewidth $G = ([n], E)$ → large

Simple ← Elimination Algo Complexity $P(x)$ → Complex





	chain	Tree	2D grid	3D grid	Erdős-Rényi $p = \frac{c \log n}{n}$	K_n
$E = \emptyset$						
TW = 1	2	2	$\sqrt{n}$	$n^{2/3}$	n	n
$O(n\|\mathcal{X}\|)$		$O(n\|\mathcal{X}\|^2)$				$O(\|\mathcal{X}\|^n)$
max-degree 0	2	2	4	6	$2 \cdot C \cdot \log n$ ≪	n

* Elimination Algorithm for MAP inference.

$$x^* \in \arg\max_{x \in \mathcal{X}^n} P(x)$$

Q. Can we use $P(x_1)$ to find x_1^* ?

$$\arg\max_{x_1} P(x_1) \overset{?}{=} x_1^*$$

$$\arg\max_{x_1} \sum_{x_2 \cdots x_n} P(x) \neq \arg\max_{x_1} \max_{x_2 \cdots x_n} P(x)$$

Elimination ordering $(5,4,3,2,1)$.

$$\arg\max_{x_1 \cdots x_5} P(x) = \frac{1}{Z} \arg\max_{x_1 \cdots x_4} f_{12}(x_1,x_2) f_{13}(x_1,x_3) \underbrace{\max_{x_5} f_{25}(x_2,x_5) f_{345}(x_3,x_4,x_5)}_{\substack{m_5^*(x_2,x_3,x_4) \\ x_5^*(x_2,x_3,x_4)}}$$

← maximize

$\Rightarrow (x_1^*, \ldots, x_5^*)$

$$= \frac{1}{Z} \arg\max f_{12} f_{13} \underbrace{\max_{x_4} m_5^*(x_2,x_3,x_4)}_{\substack{m_4^*(x_2,x_3) \\ x_4^*(x_2,x_3)}}$$

$$= \frac{1}{Z} \arg\max_{x_1,x_2} f_{12} \underbrace{\max_{x_3} f_{13}(x_1,x_3) m_4^*(x_2,x_3)}_{\substack{m_3^*(x_1,x_2) \\ x_3^*(x_1,x_2)}} \rightarrow (x_1^*, x_2^*, x_3^*, x_4^*)$$

$$= \frac{1}{Z} \arg\max_{x_1} \underbrace{\max_{x_2} f_{12}(x_1,x_2) m_3^*(x_1,x_2)}_{\substack{m_2^*(x_1) \\ x_2^*(x_1)}} \rightarrow (x_1^*, x_2^*, x_3^*)$$

$x_2^*(x_1) \rightarrow (x_1^*, x_2^*)$

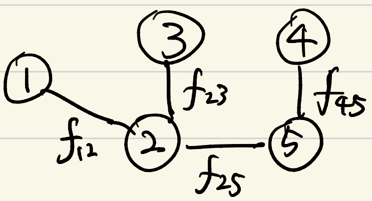
$$= \frac{1}{Z} \underbrace{\max_{x_1} m_2^*(x_1)}_{\substack{m_1^* \\ x_1^*}}$$

(Exact on Tree) (Approximate on G)

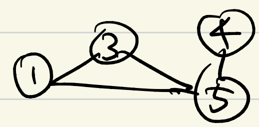
* Elimination Algorithm on a Tree $\Longrightarrow$ sum-product algorithm
  (loopy) belief propagation
  $\cong$ message-passing algorithm

Def. <u>Tree</u> is a graph
- no cycle
- connected
- $|E| = n-1$

2 redundant



wrong elimination order: $(2, \dots)$



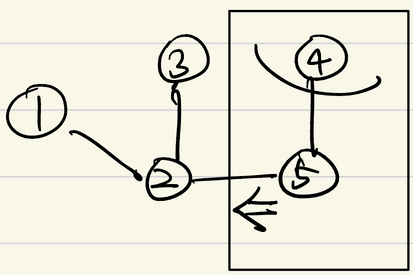
$|\mathcal{X}|^3$

Good order: recursively eliminate <u>leaves</u>

$\triangleq$ node with degree 1.

$g = (4, 5, 3, 1, 2)$ $\quad (1\ 3\ 4\ 5\ 2), (1\ 3\ 2\ 5\ 4)$

$$P(x_2) = \frac{1}{Z} \sum_{x_1, x_3, x_5} f_{12} \cdot f_{23} \cdot f_{25} \cdot \underbrace{\sum_{x_4} f_{45}(x_4, x_5)}$$

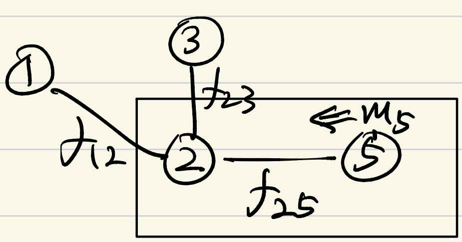


$$= m_5(x_5) = [\ \ ] \}|\mathcal{X}|$$

$$= \frac{1}{Z} \sum_{x_1, x_3} f_{12} \cdot f_{23} \cdot \sum_{x_5} f_{25}(x_2, x_5) \cdot m_{5 \to 2}(x_5)$$

parent leave to be eli.

pairwise compatibility func.



$$= m_2(x_2)$$