CS5340:
Tutorial 4

Asst. Prof. Harold Soh

TAs: Eugene Lim

Course Schedule (Tentative)

Week	Date	Lecture Topic	Tutorial
1	16 Jan	Introduction to Uncertainty Modeling + Probability Basics	~~Introduction~~
2	23 Jan	Simple Probabilistic Models	Introduction and Probability Basics
3	30 Jan	Bayesian networks (Directed graphical models)	More Basic Probability
4	6 Feb	Markov random Fields (Undirected graphical models)	DGM modelling and d-separation
5	13 Feb	Variable elimination and belief propagation	MRF ~~+ Sum/Max Product~~
6	20 Feb	Factor graphs	**Quiz 1**
-	-	**RECESS WEEK**	
7	5 Mar	Mixture Models and Expectation Maximization (EM)	Linear Gaussian Models
8	12 Mar	Hidden Markov Models (HMM)	Probabilistic PCA
9	19 Mar	Monte-Carlo Inference (Sampling)	Linear Gaussian Dynamical Systems
10	26 Mar	Variational Inference	MCMC + Langevin Dynamics
11	2 Apr	Inference and Decision-Making	Diffusion Models + Sequential VAEs
12	9 Apr	Gaussian Processes (optional)	**Quiz 2**
13	16 Apr	Project Presentations	Closing Lecture

Administrative Issues

- **Project Groups:**
  - **Abstracts due:** 4 March 2024
- **Quiz 1**
  - **Next week:** 20 Feb 2024
  - Covers everything up to variable elimination and belief propagation
  - Last Year's quiz has been uploaded to Canvas.
- **Survey** online: **https://forms.gle/FKSZ5MbtPXPnC8Mv6**

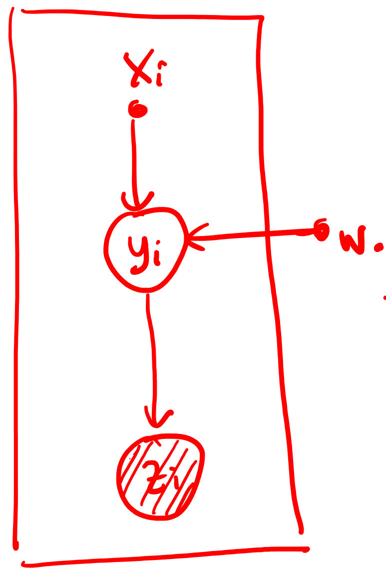
Questions?

https://pollev.com/haroldsohsoo986

SCAN ME

4. Label Errors

4.c. Bayesian Network with Label Errors



$p(y_i \mid x_i, w)$

z	y	$p(z \mid y)$
0	0	0.75
0	1	0.05
1	0	0.25
1	1	0.95.

4. Label Errors

4.d. New MLE that accounts for label errors.

$$\hat{w} = \arg\max_{w} \lg \prod_{i=1}^{N} \underbrace{p(z_i \mid x_i, w)}.$$

$$= \sum_{y_i} p(z_i, y_i \mid x_i, w).$$

$$= \sum_{y_i} p(z_i \mid y_i)\, p(y_i \mid x_i, w).$$

$$\Downarrow$$

$$= \sum_{i=1}^{N} \lg\Big(\sum_{y_i} p(z_i \mid y_i)\, p(y_i \mid x_i, w)\Big).$$

magic?

$$= \sum_{i}^{N} z_i \lg\left[p(z_i = 1 \mid y_i = 1)\, p_i + p(z_i = 1 \mid y_i = 0)\,(1 - p_i)\right] +$$

$$(1 - z_i) \lg\left[p(z_i = 0 \mid y_i = 1)\, p_i + p(z_i = 0 \mid y_i = 0)\,(1 - p_i)\right].$$

$$\sum_{i=1}^{N} \lg \Big(\underbrace{\sum_{y_i} p(z_i | y_i) \underbrace{p(y_i | x_i, w)}} \Big)$$

$$\lg \underbrace{p(z_i | x_i, w)}.$$

$$= \text{Bern}[\pi_i].$$

$$\text{Bern}[\pi_i] = \pi_i^{z_i} (1-\pi_i)^{1-z_i}$$

$$\lg \pi_i^{z_i} (1-\pi_i)^{1-z_i}$$

$$= z_i \lg \pi_i + (1-z_i) \lg (1-\pi_i)$$

$$\lg p(z_i = 1 | x_i, w) = \lg \Big(\sum_{y_i} p(z_i = 1 | y_i) p(y_i | x_i, w) \Big).$$

$$= \lg \big(p(z_i = 1 | y_i = 1) \underbrace{p(y_i = 1 | x_i, w)} + p(z_i = 1 | y_i = 0) p(y_i = 0 | x_i, w) \big)$$

$$= \lg \big(p(z_i = 1 | y_i = 1) \rho_i + p(z_i = 1 | y_i = 0)(1-\rho_i) \big).$$

$$\lg p(z_i = 0 | x_i, w) = \lg \big(p(z_i = 0 | y_i = 1) p(y_i = 1 | x_i, w) + p(z_i = 0 | y_i = 0) p(y_i = 0 | x_i, w) \big).$$

$$= \lg \big(p(z_i = 0 | y_i = 1) \rho_i + p(z_i = 0 | y_i = 0)(1-\rho_i) \big).$$

$$\lg p(z_i | x_i, w) = \begin{cases} \lg p(z_i = 1 | x_i, w) & \text{if } z_i = 1 \\ \lg p(z_i = 0 | x_i, w) & \text{if } z_i = 0 \end{cases}$$

$$= z_i \lg p(z_i = 1 | x_i, w) + (1-z_i) \lg p(z_i = 0 | x_i, w). \quad \|$$

CS5340:

Tutorial 4

Asst. Prof. Harold Soh

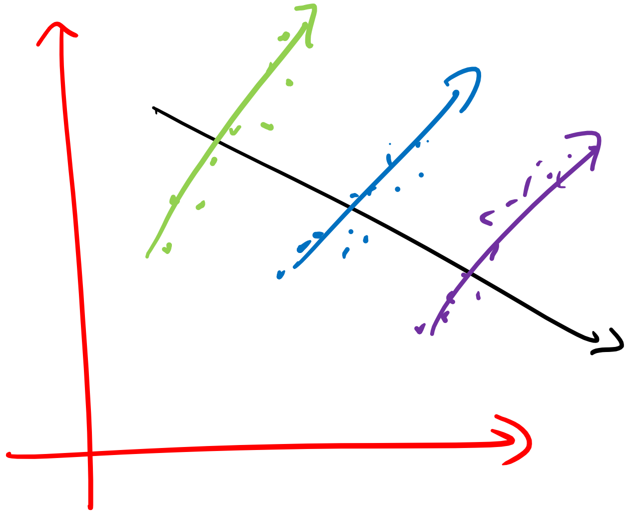
TAs: Eugene Lim

1. Simpson's Paradox

1.a. There are two potential drug treatments (A and B).

Dr. Homer Simpson wants to compare the outcome of patients after receiving either treatment A or B, to determine the better drug.

Draw a PGM for the scenario above.



Seems like Drug A is better

$$T \longrightarrow R$$

Treatment (A, B)

Recovered (0, 1)

Not — Recov.

t	$p(T=t)$
A	0.73
B	0.27

$$\frac{1500}{1500+550}$$

t	r	$p(R=r \mid T=t)$
A	0	0.16
A	1	0.84
B	0	0.19
B	1	0.81

$\boxed{A}$

1500

240 Not cured — 1260 cured

$\boxed{B}$

550

105 Not cured — 445 cured

$$p(R=0 \mid T=A) = \frac{240}{1500} = 0.16$$

$$p(R=0 \mid T=B) = \frac{105}{550} = 0.19$$

Condition on the condition (C), treatment B is better

$$p(R=1 \mid T=A, C=0) = \frac{1190}{1400} = 0.85$$

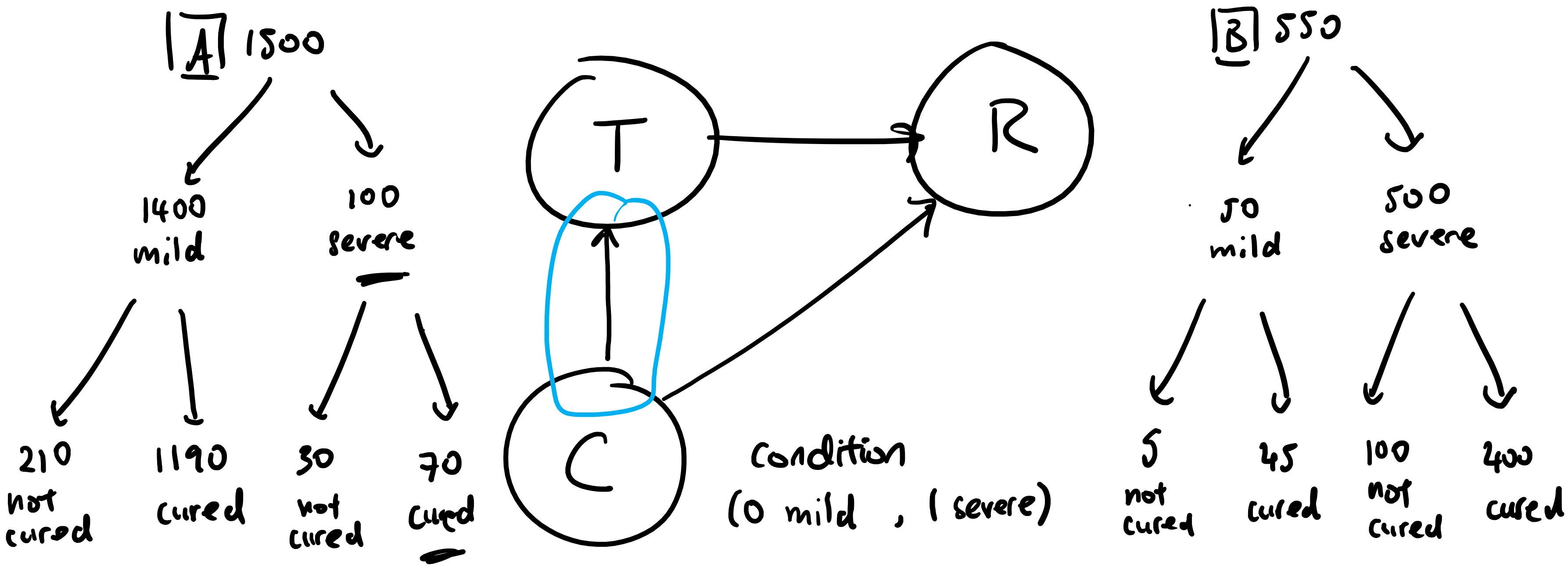
$$p(R=1 \mid T=A, C=1) = \frac{70}{100} = 0.7$$

$$p(R=1 \mid T=B, C=0) = \frac{45}{50} = 0.9$$

$$p(R=1 \mid T=B, C=1) = \frac{400}{500} = 0.8$$

Simpson's Paradox

1.b. The data seems to indicate that treatment A is more effective. Can Homer confirm this just from the data?



Questions?

https://pollev.com/haroldsohsoo986

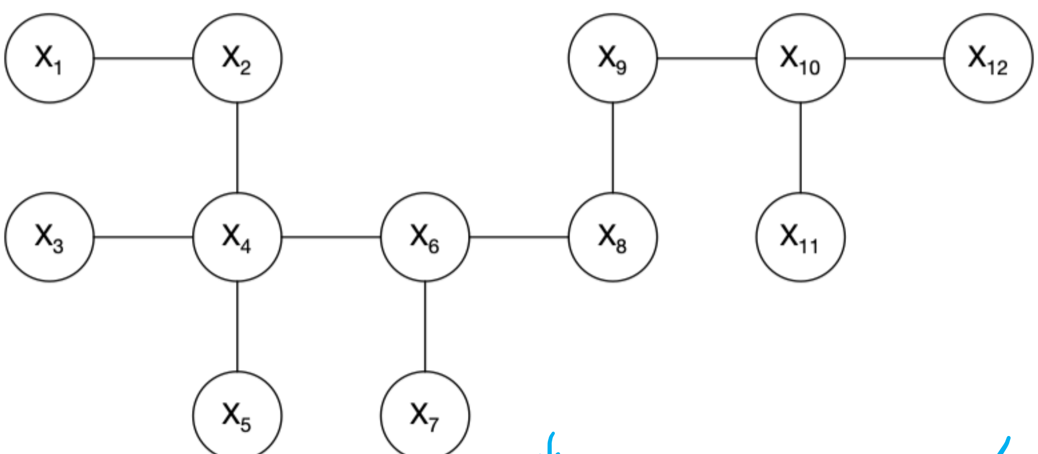
SCAN ME

MRT Inference

You are given the following *pairwise* undirected graphical model which models the activity (low or high) at 12 MRT stations. Each node represents a random variable indicating whether the activity at a particular station is low (0) or high (1).

x_i	$\psi(x_i)$
0	10
1	2

Unary Factors



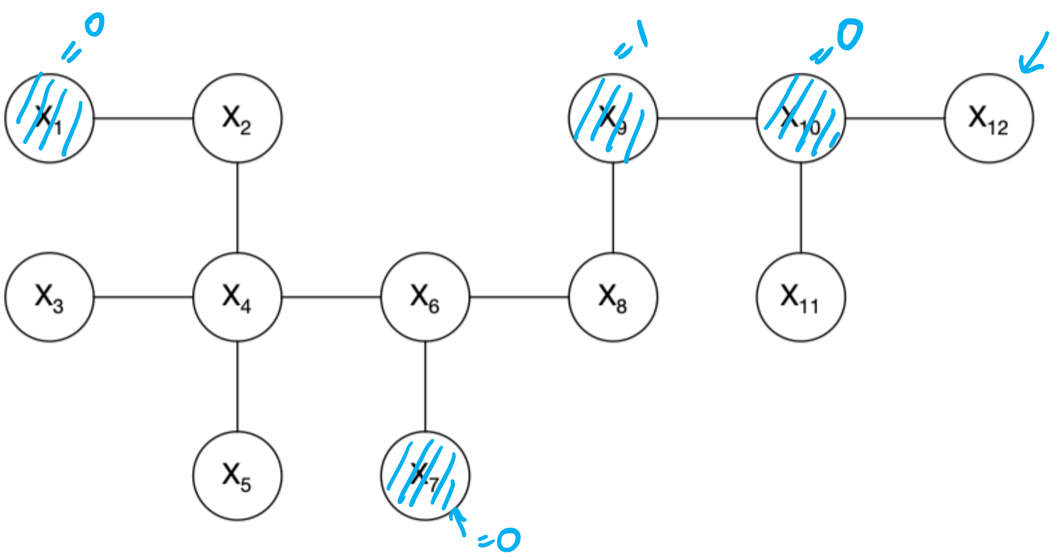
x_i	x_j	$\psi(x_i, x_j)$
0	0	20
0	1	5
1	0	5
1	1	20

Pairwise Factors

$$p(x_1, x_2, \ldots, x_{12}) = \frac{1}{Z} \prod_{i \in V} \psi(x_i) \prod_{(i,j) \in E} \psi(x_i, x_j)$$

MRT Inference

2.a. Compute $p(x_{12} = 1 | x_1 = 0, x_7 = 0, x_9 = 1, x_{10} = 0)$.



x_i	$\psi(x_i)$
0	10
1	2

Unary Factors

x_i	x_j	$\psi(x_i, x_j)$
0	0	20
0	1	5
1	0	5
1	1	20

Pairwise Factors

MRT Inference

2.a. Compute $p(x_{12} = 1 | x_1 = 0, x_7 = 0, x_9 = 1, x_{10} = 0)$

$$p(x_{12} = 1 | x_1 = 0, x_7 = 0, x_9 = 1, x_{10} = 0)$$
$$= p(x_{12} = 1 | \underline{x_{10} = 0})$$

2.a. Compute $p(x_{12} = 1 | x_1 = 0, x_7 = 0, x_9 = 1, x_{10} = 0)$

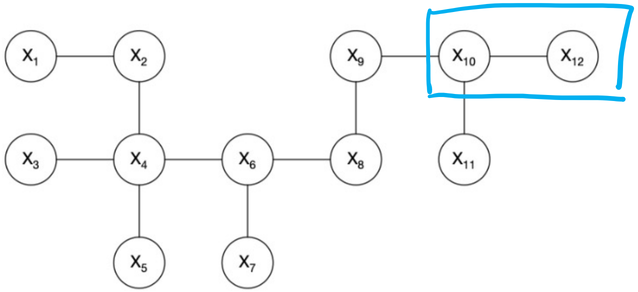
$$p(x_{12} = 1 | x_1 = 0, x_7 = 0, x_9 = 1, x_{10} = 0)$$
$$= p(x_{12} = 1 | x_{10} = 0)$$

x_i	$\psi(x_i)$
0	10
1	2

Unary Factors

x_i	x_j	$\psi(x_i, x_j)$
0	0	20
0	1	5
1	0	5
1	1	20

Pairwise Factors



$$\underline{\underline{p(x_1, x_2 \ldots x_{12})}} = \frac{1}{Z} \prod \psi(x_i) \prod \psi(x_i, x_j).$$

x_i	$\psi(x_i)$
0	10
1	2

Unary Factors

x_i	x_j	$\psi(x_i, x_j)$
0	0	20
0	1	5
1	0	5
1	1	20

Pairwise Factors

$$p(x_{12} \mid x_{10}) = \frac{p(x_{12}, x_{10})}{p(x_{10})}$$

$$\underline{\underline{p(x_{12}, x_{10})}} = \sum_{\underline{\underline{V \setminus \{x_{10}, x_{12}\}}}} \frac{1}{Z} \psi(x_1)\psi(x_2) \ldots \psi(x_{12})\, \psi(x_1, x_2) \ldots \psi(x_{10}, x_{12})$$

$$= \frac{1}{Z} \psi(x_{10})\psi(x_{12})\psi(x_{10}, x_{12}) \underbrace{\sum \psi(x_1) \ldots \psi(x_{11})\, \psi(x_1, x_2) \ldots \psi(x_{11}, \underline{\underline{x_{10}}})}_{m(\underline{\underline{x_{10}}}).}$$

$$p(x_{10}) = \sum_{x_{12}} p(x_{12}, x_{10})$$

$$= \frac{1}{Z} m(x_{10})\psi(x_{10}) \sum_{x_{12}} \psi(x_{12})\psi(x_{10}, x_{12}).$$

$$\frac{p(x_{12}, x_{10})}{p(x_{10})} = \frac{\cancel{\frac{1}{Z}}\, m(\cancel{x_{10}})\, \psi(\cancel{x_{10}})\, \psi(\overset{=1}{x_{12}})\, \psi(\overset{=0}{x_{10}}, \overset{=1}{x_{12}})}{\cancel{\frac{1}{Z}}\, m(\cancel{x_{10}})\, \psi(\cancel{x_{10}}) \sum_{x'_{12}} \psi(x'_{12})\, \psi(\overset{=0}{x_{10}}, x'_{12})} = \frac{2 \times 5}{(20 \times 10) + (5 \times 2)} = \frac{1}{21}$$

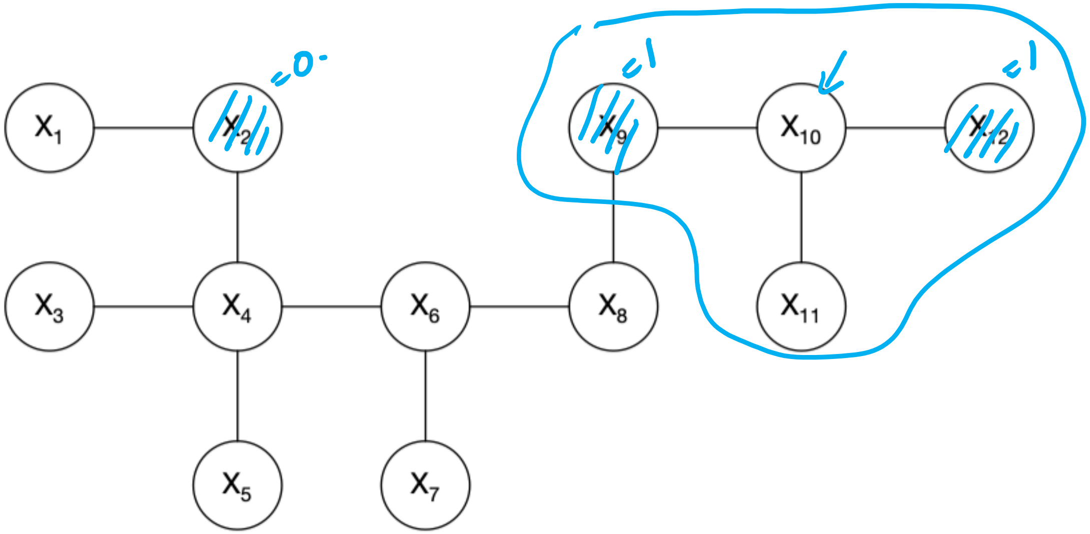
MRT Inference

2.c. Compute $p(x_{10} = 1 \mid x_9 = 1, x_{12} = 1, x_2 = 0)$

$$p(x_{10} \mid x_9, x_{12}, x_2) = p(x_{10} \mid x_9, x_{12})$$

$$= \frac{p(x_{10}, x_9, x_{12})}{p(x_9, x_{12})}$$

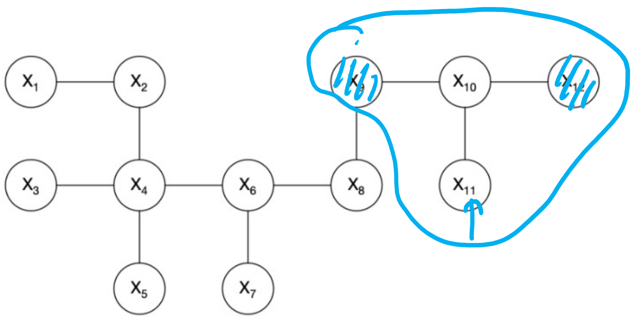
2.c. Compute $p(x_{10} = 1 \mid x_9 = 1, x_{12} = 1, x_2 = 0)$



Unary Factors

x_i	x_j	$\psi(x_i, x_j)$
0	0	20
0	1	5
1	0	5
1	1	20

Pairwise Factors



$$p(x_{10} \mid x_9, x_{12}) = \frac{p(x_{10}, x_9, x_{12})}{p(x_9, x_{12})}$$

x_i	$\psi(x_i)$
0	10
1	2

Unary Factors

x_i	x_j	$\psi(x_i, x_j)$
0	0	20
0	1	5
1	0	5
1	1	20

Pairwise Factors

$$p(x_9, x_{10}, x_{12}) = \sum_{x_{11}} \underbrace{p(x_9, x_{10}, x_{11}, x_{12})}.$$

$$= \sum_{x_{11}} \sum_{V \setminus \{x_9, x_{10}, x_{11}, x_{12}\}} \frac{1}{Z} \psi(x_1)\psi(x_2) \cdots \psi(x_{12})\psi(x_1, x_2) \cdots \psi(x_{10}, x_{12})$$

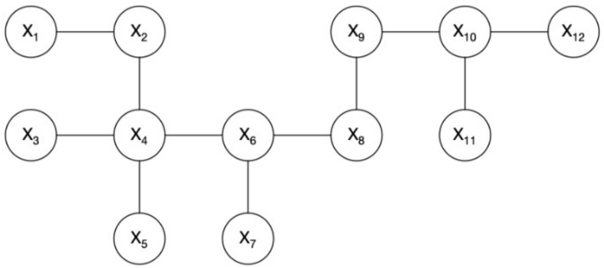
$$= \frac{1}{Z} \sum_{x_{11}} \psi(x_9)\psi(x_{10})\psi(x_{11})\psi(x_{12})\psi(x_9, x_{10})\psi(x_{10}, x_{11})\psi(x_{10}, x_{12})$$

$$\boxed{\sum_{V \setminus \{x_9, 10, 11, 12\}} \psi(x_1)\psi(x_2) \cdots \psi(x_8, x_9)} \; m(x_9)$$

$$= \cancel{\frac{1}{Z}} \cancel{m(x_9)} \cancel{\psi(x_9)} \psi(x_{10}) \cancel{\psi(x_{12})} \psi(x_9, x_{10})\psi(x_{10}, x_{12}) \sum_{x_{11}} \psi(x_{11})\psi(x_{10}, x_{11})$$

$$p(x_9, x_{12}) = \sum_{x_{10}} p(x_9, x_{10}, x_{12})$$

$$= \cancel{\frac{1}{Z}} \cancel{m(x_9)} \cancel{\psi(x_9)} \cancel{\psi(x_{12})} \sum_{x_{10}} \psi(x_{10})\psi(x_9, x_{10})\psi(x_{10}, x_{12}) \sum_{x_{11}} \psi(x_{11})\psi(x_{10}, x_{11})$$



x_i	$\psi(x_i)$
0	10
1	2

Unary Factors

x_i	x_j	$\psi(x_i, x_j)$
0	0	20
0	1	5
1	0	5
1	1	20

Pairwise Factors

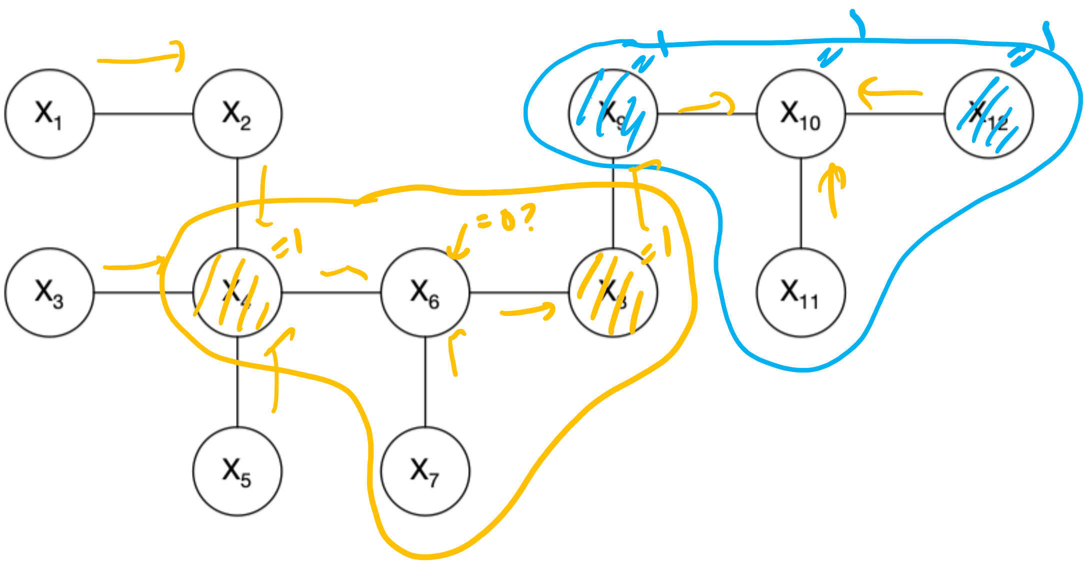
$$\frac{p(x_{10}, x_9, x_{12})}{p(x_9, x_{12})} = \frac{\psi(x_{10})\,\psi(x_9, x_{10})\,\psi(x_{10}, x_{12}) \sum_{x_{11}} \psi(x_{11})\,\psi(x_{10}, x_{11})}{\sum_{x_{10}'} \psi(x_{10}')\,\psi(x_9, x_{10}')\,\psi(x_{10}', x_{12}) \sum_{x_{11}} \psi(x_{11})\,\psi(x_{10}', x_{11})}$$

$$= \frac{20 \times 20 \times 2\,(10 \times 5 + 2 \times 20)}{124{,}500} = \frac{48}{83} \approx 0.5783$$

MRT Inference

2.d. Compute $p(x_6 = 0 \mid x_4 = 1, x_8 = 1, x_{10} = 0)$

$1 - \frac{48}{83} = \frac{35}{83} \approx 0.4217$



x_i	$\psi(x_i)$
0	10
1	2

Unary Factors

2.d. Compute $p(x_6 = 0 \mid x_4 = 1, x_8 = 1, x_{10} = 0)$

Pairwise Factors

Questions?

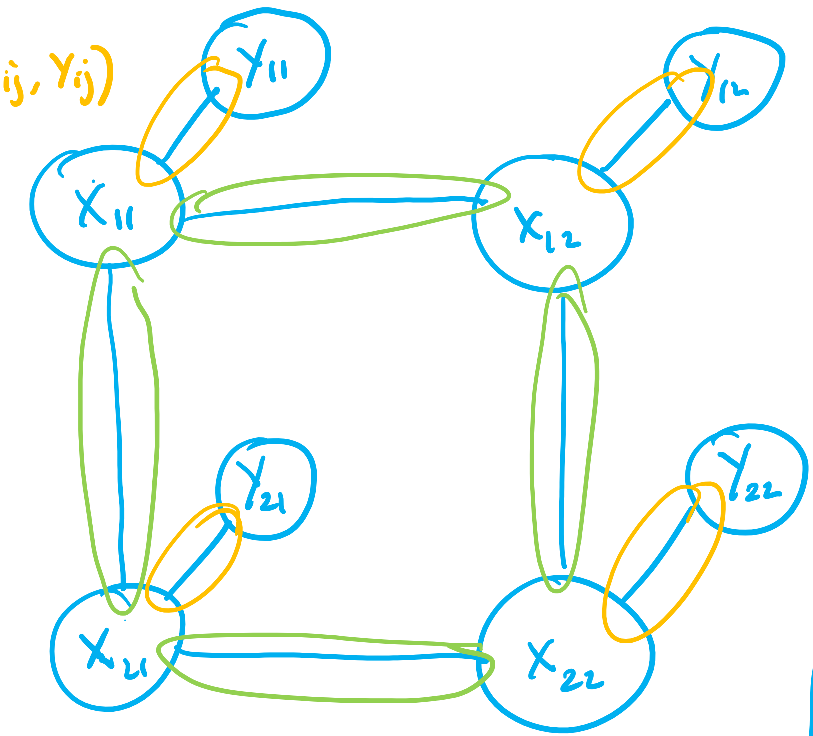
https://pollev.com/haroldsohsoo986

SCAN ME

Denoising

- Let's move to the Jupyter notebook.

"local energy" – $\psi(x_{ij}, y_{ij})$



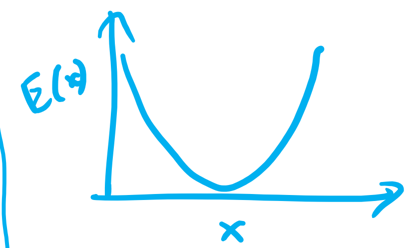
$\psi(x_{ij}, x_{ij}^{n})$

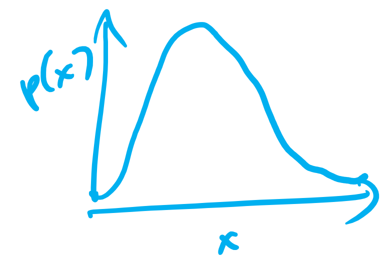
<u>idea</u>

model energy functions.

↑ energy → ↓ probable

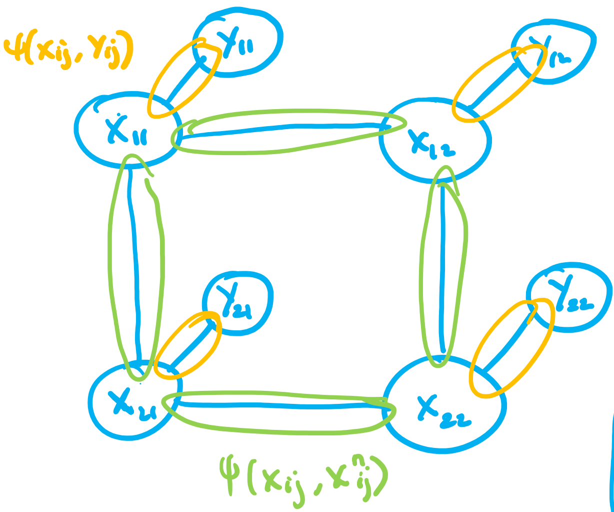
↓ energy → ↑ probable





$p(x) = \exp(-E(x))$

$E(x) = \sum \varphi(x_i, x_j)$

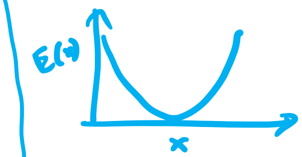


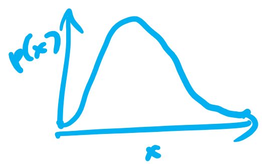
idea

model energy functions

↑ energy → ↓ probable

↓ energy → ↑ probable





$$p(x) = \exp(-E(x))$$

$$E(x) = \sum \psi(x_i, x_j)$$

X_{ij}	Y_{ij}	$X_{ij} \cdot Y_{ij}$	$\psi(x_{ij}, Y_{ij}) = -\alpha X_{ij} Y_{ij}$
-1	-1	$+1$	$-\alpha$
-1	$+1$	-1	$+\alpha$
$+1$	-1	-1	$+\alpha$
$+1$	$+1$	$+1$	$-\alpha$

X_{ij}	X^n_{ij}	$X_{ij} \cdot X^n_{ij}$	$\psi(x_{ij}, x^n_{ij}) = -\beta X_{ij} X^n_{ij}$
-1	-1	$+1$	$-\beta$
-1	$+1$	-1	$+\beta$
$+1$	-1	-1	$+\beta$
$+1$	$+1$	$+1$	$-\beta$.

Questions?

https://pollev.com/haroldsohsoo986

SCAN ME

Homework. ☺

- Spend 5 minutes now filling up this survey:
- **https://forms.gle/FKSZ5MbtPXPnC8Mv6**
- Watch the lecture videos and prep for quiz!