Problem 1

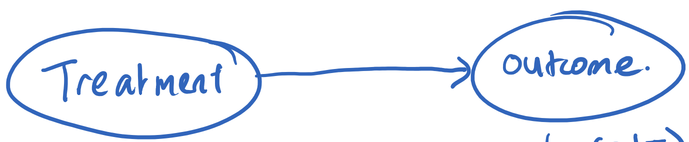
Thursday, 9 February 2023   4:11 PM

Problem 1. (Simpson's Paradox)
A dangerous new virus is sweeping the world. Currently, there are two potential drug treatments (A and B) for patients. Dr. Homer Simpson wants to compare the un-cured rate of patients after receiving either treatment A or B, in order to determine the better drug.

The data indicates that there are 240 patients that are not cured among the 1500 patients who received treatment A. There are 105 patients that are not cured among the 550 who received treatment B. *Note: this is a fictitious scenario and we made up these numbers.*

Problem 1.a. Can you help Homer construct a probabilistic graphical model for the above scenario.

$p(O,T) = p(T)p(O|T)$



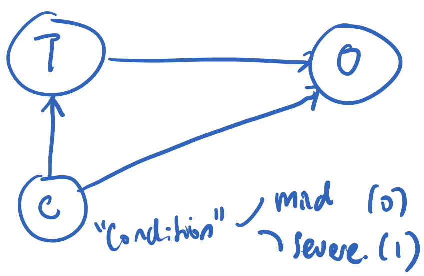
O = Cured
1 = not cured.

T	$p(T)$
A	$1500/(1500+550) = 0.73$
B	0.27

O	T	$p(O\|T)$
0	A	0.84
0	B	0.81
1	A	0.16
1	B	0.19

Problem 1.b. The data seems to indicate that treatment A is more effective. Can Homer confirm (just from the data) that one of the treatments results in more cures? *Hint:* Consider what happens when there are unobserved variables that could affect the treatment and the outcome.

			not cured
A	1500	1400 mild	210
		100 severe	30
B	550	50 mild	5
		500 severe.	100



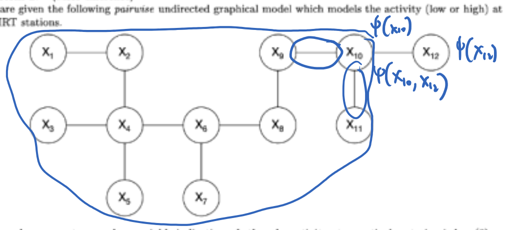
Outcome
0 = cured
1 = not cured.

O	T	C	$p(O\|T,C)$
1	A	0	0.15
1	A	1	0.3
1	B	0	0.1
1	B	1	0.2.

Problem 2a

Thursday, 9 February 2023 4:12 PM

Problem 2. (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). Assume the following factorization:

$$\boxed{p(x_1, x_2, \dots, x_{12}) = \frac{1}{Z}\prod_{i\in V}\psi(x_i)\prod_{(i,j)\in E}\psi(x_i, x_j)} \qquad (11)$$

where V is the set of nodes, E is the set of edges, and that the unary and pairwise factors are given by:

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

Figure 3: Unary Factors

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

Figure 4: Pairwise Factors

Note that the factors are the same across the nodes. Your task is to compute the following conditional probabilities.

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

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

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

$$= \frac{\sum_{x_c} p(x_1, \dots, x_{12})}{\sum_{x_{12}}\sum_{x_c} p(x_1, \dots, x_{12})}$$

$$p(x_{12}, x_{10}) = \sum_{x_c} \frac{1}{Z}\prod_i \psi(x_i)\prod_{(i,j)\in E}\psi(x_i, x_j).$$

$$= \frac{1}{Z}\psi(x_{10})\psi(x_{12})\psi(x_{10}, x_{12})\underbrace{\left[\sum \cdots \psi(x_{10}, x_{11})\psi(x_9, x_{10})\cdots\right]}_{m(x_{10})}$$

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

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

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

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

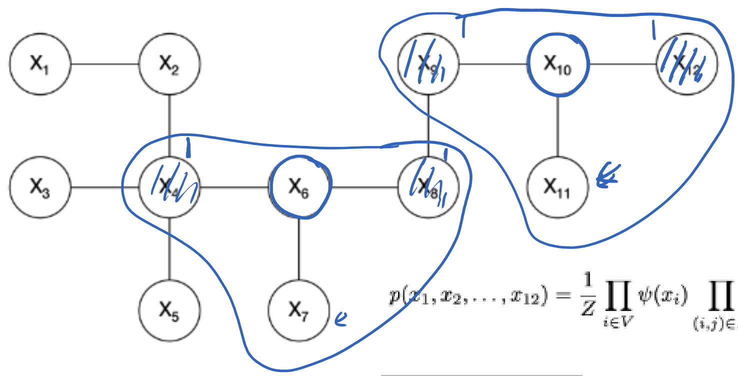
$$= \frac{\cancel{\frac{1}{Z}}\cancel{m(x_{10})}\cancel{\psi(x_{10})}\psi(x_{12})\psi(x_{10}, x_{12})}{\cancel{\frac{1}{Z}}\cancel{m(x_{10})}\cancel{\psi(x_{10})}\sum_{x_{12}}\psi(x_{12})\psi(x_{10}, x_{12})}$$

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

$$p(x_{12} = 1 | x_{10} = 0) = \frac{5 \times 2}{20 \times 10 + 5 \times 2} = \frac{1}{21} = 0.0476.$$

Problem 2c and 2d

Thursday, 9 February 2023 4:13 PM



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

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

$$Z = \sum_{x_{11}} m(x_9)\, \psi(x_9)\, \psi(x_9, x_{10}) \cdots$$

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

Figure 3: Unary Factors

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

Figure 4: Pairwise Factors

Problem 2.c. Compute $p(x_{10} = 1 \mid x_9 = 1, x_{12} = 1, x_2 = 0)$. $= 0.5783$

Problem 2.d. Compute $p(x_6 = 0 \mid x_4 = 1, x_8 = 1, x_{10} = 0)$. $= 1 - 0.5783 = 0.4217.$

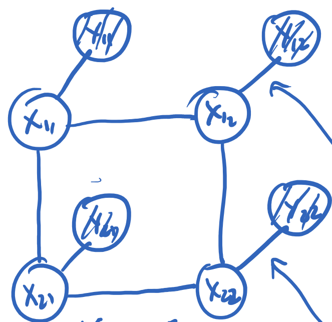
Image Denoising

Thursday, 9 February 2023 4:16 PM

Problem 3. (Image Denoising)

For this problem, you will be working on Image Denoising, taking a noisy image and making it a clean one. Please refer to the provided Image-Denoising-Pre.ipynb notebook. You can download the notebook and relevant images in a zipfile from NUS Canvas (in the MRF module under Home).

To use the notebook, you have to install jupyter (`https://jupyter.org/install`) and a python distro; we use anaconda (`https://www.anaconda.com/products/individual`) but you can use whichever distribution you like.



$\psi(x_{ij}, x^*_{ij})$

~~pEx~~

$$p(X,Y) = \frac{1}{Z(\theta)} \exp\left(-E(X,Y)\right)$$

$$\psi(x_{ij}, x^*_{ij}) = -\beta\, x_{ij} x^*_{ij}$$

neighbor of x_{ij}

$$\psi(x_{ij}, y_{ij}) = -\alpha\, x_{ij} y_{ij}$$

3

x_{ij}	y_{ij}	$x_{ij} y_{ij}$	$-\alpha x_i x_j$
-1	-1	+1	-3
-1	+1	-1	+3
+1	-1	-1	+3
+1	+1	+1	-3