ECE 285

HW-1 Solutions

January 2022

1 Problem - 1

- T and U are independent - False, since the path from T to U via I is not blocked.
- T and U are conditionally independent given I, E, and H - False, since the path from T to U via E meets head-to-head at E which is in the node list hence is not blocking the path.
- T and U are conditionally independent given I and H- True, since all the paths from T to U are blocked. First path from T to U is via E which meets head-to-head at E and is not included in the node list hence this path is blocked. Second path from T to U is via I which meets tail-to-tail and is included in the node list and hence is blocked. Thus all the paths are blocked.
- E and H are conditionally independent given U - False, since the path from H to E via U-I-T is not blocked. The node U meets head-to-head which is included in the node list hence U is not blocking the path. The node I meets tail-to-tail and is not in the node list thus is not blocking the path. Similarly node T which meets head-to-tail is not in the node list and hence does not block the path.
- E and H are conditionally independent given U, I, and T - True, since the path from E to H via U is blocked at U since it meets head-to-tail and is in the node list. The path from E to H via T-I-U is also blocked by node T and I which meet head-to-tail and tail-to-tail respectively and are in the node list.
- I and H are conditionally independent given E - False, since the path from I to H via U, which meets head-to-head, is not blocked since its descendant E is in the node list.
- I and H are conditionally independent given T - True. The path from I to H via U is blocked since it meets head-to-head and neither U nor its descendant E is not included in the node list. The other path from I to

I = 0	I = 1
0.5	0.5

Figure 1: Probability of I

H = 0	H = 1
0.4	0.6

Figure 2: Probability of H

H via T-E-U is also blocked at node T which meets head-to-tail and is included the node list.

- T and H are independent - True. The first path from T to H via I and U is blocked by U which meets head-to-head and neither U nor its descendant E is not included in the node list. The second path from T to H via E-U is blocked by E which meets head-to-head and is not not in the node list.
- T and H are conditionally independent given E - False, since the path from T to H via E-U is not blocked. Node E is not blocking the path since it meets head-to-head and is included in the node list. Node U is not blocking the path too since it meets head-to-tail and is not included in the node list.
- T and H are conditionally independent given E and U - False, since the path from T to H via I-U is not blocked. The node I meets tail-to-tail and is not included in the node list thus I is not blocking. The node U meets head-to-head and both U and its descendant E are included in the node list thus are not blocking the path.

2 Problem 2

- We assume that intelligence is equally likely, Figure 1.
- We assume that there are more hardworking students in a class, Figure 2.
- If a student is neither intelligent nor hardworking then with a high probability the student might not understand the material well. Similarly we can explain the rest of the table, Figure 3.
- There is a higher chance of an intelligent student being a good test taker however the converse is lesser likely to happen, Figure 4.

	U = 0	U = 1
I = 0, H = 0	0.9	0.1
I = 1, H = 0	0.4	0.6
I = 0, H = 1	0.6	0.4
I = 1, H = 1	0.1	0.9

Figure 3: Conditional Probability of U given I and H

	T = 0	T = 1
I = 0	0.6	0.4
I = 1	0.3	0.7

Figure 4: Conditional Probability of T given I

- If a student is a good test taker or understands the material well then with a good probability the student will score well. Similarly we can explain the rest of the table, Figure 5.

Note: There is no unique answer.

3 Problem 3

$$P(I, H, T, U, E) = P(I)P(H \mid I)P(T \mid I, H)P(U \mid I, H, T)P(E \mid I, H, T, U)$$

The expression of joint probability after simplifying the expression using the conditional independence from the graph structure.

$$P(I, H, T, U, E) = P(I)P(H)P(T \mid I)P(U \mid I, H)P(E \mid T, U)$$

4 Problem 4

$$P(E = 1 \mid H = 1) = \frac{P(E = 1, H = 1)}{P(H = 1)}$$

Observe that we know the denominator from Figure 2. Our goal is to determine P(E = 1, H = 1).

$$P(E = 1, H) = \sum_{I,T,U} P(I, H, T, U, E = 1)$$

	E = 0	E = 1
T = 0, U = 0	0.9	0.1
T = 1, U = 0	0.4	0.6
T = 0, U = 1	0.4	0.6
T = 1, U = 1	0.1	0.9

Figure 5: Conditional Probability of E given T and U

$$
\begin{aligned}
\textstyle\sum_{I,T,U} P(I,H,T,U,E=1) &= \textstyle\sum_{I,T,U} P(I)P(H)P(T \mid I)P(U \mid I,H)P(E=1 \mid T,U) \\
&= P(H) \textstyle\sum_{I,U} P(I)P(U \mid I,H) \sum_{T} P(T \mid I)P(E=1 \mid T,U) \\
&= P(H) \textstyle\sum_{I} P(I) \sum_{U} P(U \mid I,H)\{\sum_{T} P(T \mid I)P(E=1 \mid T,U)\} \\
&= P(H) \textstyle\sum_{I} P(I)\{\sum_{U} P(U \mid I,H)m_T(I,U)\} \\
&= P(H)\{\textstyle\sum_{I} P(I)m_U(I,H)\} \\
&= P(H)m_I(H)
\end{aligned} \tag{1}
$$

where,

$$m_T(I,U) = \sum_{T} P(T \mid I)P(E=1 \mid T,U)$$

$$m_U(I,H) = \sum_{U} P(U \mid I,H)m_T(I,U)$$

$$m_I(H) = \sum_{I} P(I)m_U(I,H)$$

$$P(E=1,H=1) = P(H=1)m_I(H=1)$$

Note, the solution is written in a way that explains the variable elimination process. You don't need to write it in terms of message passing. Please substitute the values according to your conditional probability tables and calculate the summation.

5 Problem 5

The cities connected by an edge are reachable from one another. We assume that the San Diego County has 9 cities. X in our model (Figure 6) is the city graph with node values corresponding to the past week cases. Y is the city graph with the node values corresponding to the current number of cases. Drawing

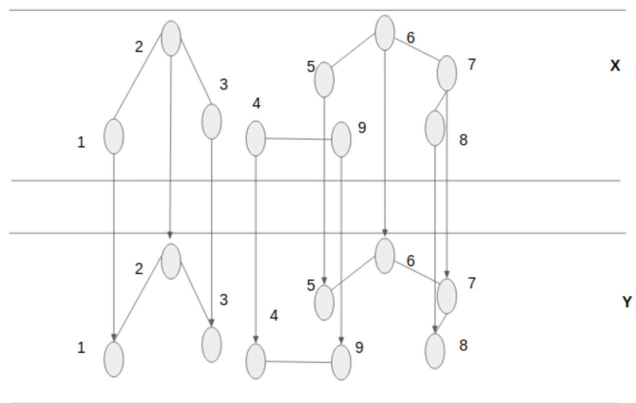

Figure 6: Markov random field model. X is the City graph with node values equal to past week's cases. Y is the City graph with node values being the estimated present week's cases.

inspiration from the conditional random field, we draw a directed edge between X and Y. Let E be the list of edges in the city graph and V be the number of nodes in the city graph.
Define,

$$\phi(Y \mid X) = {W_1}^T \overline{MSE_{YY}} + {W_2}^T \overline{MSE_{XY}}$$

where,

- MSE_{YY} is a vector of length $|E|$ and is defined as follows,

  $$\{MSE_{YY}\}_{ij} = -||Y_i - Y_j||^2$$

  where (i,j) $\in$ E, i.e., there is an edge between Y_i and Y_j. The intuition behind this term is that the cities that are connected must have similar number of cases. Each of $MSE_{YY\,ij}$ has $[W_1]_{(i,j)}$ weight associated.

- MSE_{XY} is a vector of length $|V|$ and is defined as follows,

  $$\{MSE_{XY}\}_{ii} = -||X_i - Y_i||^2$$

  where X_i is the i^{th} vertex in the city graph with past cases, X. Y_i is the i^{th} vertex in the city graph with current (to be predicted) cases, Y. The intuition behind this term is to account for a smooth variation in the number of cases from the past data to the estimated present data. Each of $MSE_{XY\,ii}$ has $[W_2]_{(i)}$ weight associated.

Using this we derive the conditional probability as follows,

$$P(Y \mid X) = \frac{1}{Z} exp\{\phi(Y \mid X)\}$$

We observe that while learning we tend to maximize the probability thereby minimizing $||Y_i - Y_j||^2$ and $||X_i - Y_i||^2$ $\forall (i, j) \in E, i \in V$ because we have a negative sign in the vector definition. Thus satisfying our criterion.
Note: There is no unique answer.