CS726: Programming Assignment 1

February 8, 2025

1 General Instructions

1. Code plagiarism will be strictly penalized including but not limited to reporting to DADAC and zero in assignments.
2. If you use tools such as ChatGPT, Copilot, you must explicitly acknowledge their usage in your report. Code borrowed from other sources must be cited in the comments.
3. Submit a report explaining your approach, implementation details, and results. Clearly mention the contributions of each team member in the report.
4. Edit the template.py file given and use Python's default libraries such as json, math, itertools, collections, functools, random, heapq.
5. If you use external sources (e.g., tutorials, papers, or open-source code), you must cite them properly in your report and comments in the code. Submit your code and report as a compressed <TeamName>_<student1rollno>_ <student2rollno>_ <student3rollno>.zip file Fill a student as NOPE if less than 3 members.
6. Start well ahead of the deadline. Submissions up to one day late will be capped at 80% of the total marks, and no marks will be awarded beyond that.

2 Problem Statement

Given an undirected graph with cliques and their associated potentials, the goal is to compute marginal probabilities and the Maximum A Posteriori (MAP) assignment using message passing algorithms. The input includes predefined cliques that are complete subgraphs within each other, ensuring structured factorization. Follow the steps below to solve the problem using message passing algorithms:

1. Given a set of variables and their potentials as an undirected graph.
2. The graph needs to be triangulated to convert it into a chordal form. Any triangulation method may be used.
3. A junction tree is constructed from the triangulated graph while maintaining the running intersection property. Each clique in this tree retains its assigned potential values.
4. The sum-product algorithm is used to compute marginal probabilities, while the max-product algorithm is used to determine the MAP assignment. Given an input k, the top k most probable assignments should be determined based on computed probabilities.

3 Details for report

The final report must include detailed steps, explanations, and structured pseudocode, formatted in LaTeX for clarity and precision.

1. **Triangulation**: Explain the process of triangulating the graph. Include diagrams if necessary.
2. **Junction Tree Construction**: Describe how to construct a junction tree from the triangulated graph and how to assign potentials to each clique.

3. **Marginal Probability**: Show how to calculate the marginal probability of each variable using the junction tree.

4. **MAP Assignment**: Define the MAP assignment and explain how to find it in the context of message passing algorithms.

5. **Top k Assignments**: Discuss how to find the top k assignments of probability values and their significance.

4 Sample Test Case

The UGM structure is as follows:

- Nodes: x_1, x_2, x_3, x_4
- Edges: $(x_1, x_2), (x_2, x_3), (x_1, x_4)$

The potentials are defined as:

- $\phi(x_1, x_2)$
- $\phi(x_2, x_3)$
- $\phi(x_1, x_4)$

4.1 Input Data

- $\phi(x_1, x_2)$

X_1	X_2	$\phi(X_1, X_2)$
0	0	3
0	1	4
1	0	5
1	1	6

- $\phi(x_2, x_3)$

X_2	X_3	$\phi(X_2, X_3)$
0	0	2
0	1	7
1	0	1
1	1	3

- $\phi(x_1, x_4)$

X_1	X_4	$\phi(X_1, X_4)$
0	0	5
0	1	8
1	0	2
1	1	7

4.2 Cliques and Junction Tree:

4.3 Messages:

- Message from C_2 to C_1 with seperator $S_{12} = \{x_2\}$:

$$m_{C_2 \to C_1}(X_2) = \sum_{X_3} \phi(X_2, X_3)$$

$$m_{C_2 \to C_1}(X_2 = 0) = 2 + 7 = 9$$

$$m_{C_2 \to C_1}(X_2 = 1) = 1 + 3 = 4$$

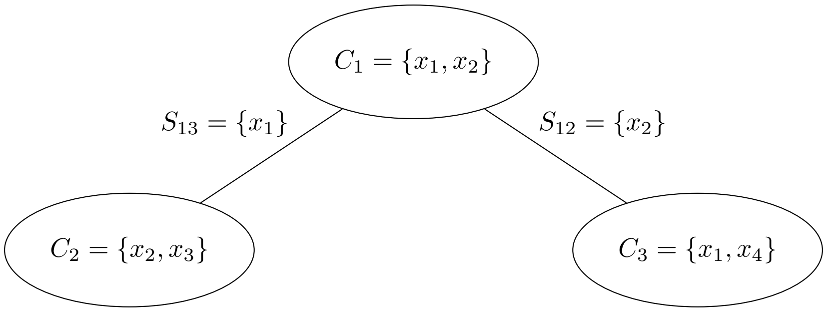

Figure 1: Junction tree with cliques C_1, C_2, C_3 and the separators S_{12}, S_{13}

- Message from C_3 to C_1 with seperator $S_{13} = \{x_1\}$:

$$m_{C_3 \to C_1}(X_1) = \sum_{X_4} \phi(X_1, X_4)$$
$$m_{C_3 \to C_1}(X_1 = 0) = 5 + 8 = 13$$
$$m_{C_3 \to C_1}(X_1 = 1) = 2 + 7 = 9$$

- Message from C_1 to C_2 with seperator $S_{12} = \{x_2\}$:

$$m_{C_1 \to C_2}(X_2) = \sum_{X_1} \phi(X_1, X_2) \cdot m_{C_3 \to C_1}(X_1)$$

$$m_{C_1 \to C_2}(X_2 = 0) = 3 \cdot 13 + 5 \cdot 9 = 39 + 45 = 84$$
$$m_{C_1 \to C_2}(X_2 = 1) = 4 \cdot 13 + 6 \cdot 9 = 52 + 54 = 106$$

- Message from C_1 to C_3 with seperator $S_{13} = \{x_1\}$:

$$m_{C_1 \to C_3}(X_1) = \sum_{X_2} \phi(X_1, X_2) \cdot m_{C_2 \to C_1}(X_2)$$

$$m_{C_1 \to C_3}(X_1 = 0) = 3 \cdot 9 + 4 \cdot 4 = 27 + 16 = 43$$
$$m_{C_1 \to C_3}(X_1 = 1) = 5 \cdot 9 + 6 \cdot 4 = 45 + 24 = 69$$

4.4 Computing Z value:

$$Z = \sum_{X_1, X_2, X_3, X_4} \phi(X_1, X_2)\phi(X_2, X_3)\phi(X_1, X_4)$$
$$Z = \sum_{X_1, X_2} \phi(X_1, X_2) \sum_{X_3} \phi(X_2, X_3) \sum_{X_4} \phi(X_1, X_4)$$

- Let $R(X_1) = \sum_{X_4} \phi(X_1, X_4)$:
  - $R(X_1 = 0) = 5 + 8 = 13$
  - $R(X_1 = 1) = 2 + 7 = 9$
- Let $Q(X_1, X_2) = \sum_{X_3} \phi(X_2, X_3) R(X_1)$:
  - $Q(0, 0) = 117$
  - $Q(0, 1) = 52$
  - $Q(1, 0) = 81$
  - $Q(1, 1) = 36$
- Calculating Z using the potentials: $Z = \sum_{X_1 X_2} \phi(X_1 X_2) Q(X_1, X_2) = 1180$

4.5 Marginal Probability:

Calculate the Marginal Probability of $X_2 = 1$

$$P(X_2 = 1) = \frac{\sum_{X_1} \phi(X_1, X_2 = 1) m_{C_2 \to C_1}(X_2 = 1) m_{C_3 \to C_1}(X_1)}{Z}$$

$$P(X_2 = 1) = \frac{424}{1180}$$

4.6 MAP for top k (k=2):

$$\underset{x_1,x_2,x_3,x_4}{\mathrm{top}_2} = \underset{x_1,x_2}{\mathrm{top}_2}\left(\phi(x_1,x_2)\right) \times \underset{x_3}{\mathrm{top}_2}\left(\phi(x_2,x_3)\right) \times \underset{x_4}{\mathrm{top}_2}\left(\phi(x_1,x_4)\right)$$

- Let $R(X_1) = \underset{x_4}{\mathrm{top}_2}\left(\phi(x_1,x_4)\right)$:
  - $R(X_1 = 0) = ((x_4 = 1)8, (x_4 = 0)5)$
  - $R(X_1 = 1) = ((x_4 = 1)7, (x_4 = 0)2)$
- Let $Q(X_1, X_2) = \underset{x_3}{\mathrm{top}_2}\left(\phi(x_2,x_3)\right) \times R(X_1)$:
  - $Q(0,0) = ((x_3 = 1, x_4 = 1)56, (x_3 = 1, x_4 = 0)35)$
  - $Q(0,1) = ((x_3 = 1, x_4 = 1)24, (x_3 = 1, x_4 = 0)15)$
  - $Q(1,0) = ((x_3 = 1, x_4 = 1)49, (x_3 = 1, x_4 = 0)14)$
  - $Q(1,1) = ((x_3 = 1, x_4 = 1)21, (x_3 = 1, x_4 = 0)6)$

$$\underset{x_1,x_2,x_3,x_4}{\mathrm{top}_2} = \underset{x_1,x_2}{\mathrm{top}_2}\left(\phi(x_1,x_2)\right) \times Q(X_1, X_2)$$

$$\underset{x_1,x_2,x_3,x_4}{\mathrm{top}_2} = ((x_1 = 1, x_2 = 0, x_3 = 1, x_4 = 1)245, (x_1 = 0, x_2 = 0, x_3 = 1, x_4 = 1)168)$$

The top 2 assignments that gives higher probabilites are:

- $x_1 = 1, x_2 = 0, x_3 = 1, x_4 = 1$ with probability $\frac{245}{1180}$
- $x_1 = 0, x_2 = 0, x_3 = 1, x_4 = 1$ with probability $\frac{168}{1180}$