CS5340: Uncertainty Modeling in AI

Coding Assignment 1

Deadline: Open for now, to be decided later

Problem 1. (Independence Queries)

Your goal is to implement a variant of the Bayes-Ball algorithm for answering conditional independence queries.

Inputs

Your program needs to accept 2 input files: the graph structure (`graph.txt`) and the conditional independence queries (`queries.txt`).
Input Format:

- The graph structure (`graph.txt`): Given a Graph with nodes numbered 1 to K. The first line contains the integer K, followed by a list of edges (one on each line). Each edge is denoted by a space-separated pair of integers a and b which represents an edge directed from node a to node b.

  Example file for the graph shown in Fig. 1:

```
4
1 2
2 3
2 4
3 4
1 4
```

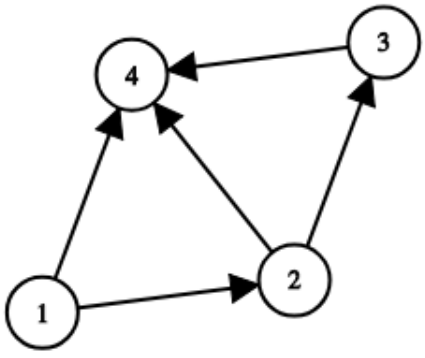

Figure 1: Example Bayesian network

- The conditional independence queries (`queries.txt`): Each line in this file represents a query of the from $(X \perp Y|Z)$. Each line consists of three space-separated sets for X, Y, and Z. Each set is denoted by a comma-separated list of integers inside braces, e.g., $\{a_1, a_2, \ldots, a_k\}$. An empty set is denoted by {}.

  Example file:

  ```
  {1} {2} {}
  {1} {3} {2}
  {1} {3} {2,4}
  ```

  This file represents the following queries:

  - $(\{1\} \perp \{2\}|\emptyset)$
  - $(\{1\} \perp \{3\}|\{2\})$
  - $(\{1\} \perp \{3\}|\{2,4\})$

Output:

Your program should print out a single integer value for each query. Print out a 1 if the graph satisfies the conditional independence query, 0 otherwise.
Example output:

```
0
1
0
```

Code Skeleton

The skeleton code is provided below with helper functions to create the graph and read queries. The code for printing the output in the required format is also provided. You are only required to complete the `is_independent` function which takes in the graph and sets X, Y, and Z, and returns `True` if $X \perp Y|Z$ and `False` otherwise. Please **do not** modify other parts of the code skeleton. Before you begin coding, **do not** forget to enter your names and matric numbers in the file's header.

Submission Format

Submit only the python (`.py`) file renamed to `YourMatricNumber-PartnerMatricNumber.py` on IVLE. If your matric number is A0174067B and your partner's is A0175067A, then the file should be named `A0174067B-A0175067A.py`. If you're doing the assignment as an individual, name it as `YourMatricNumber.py`. Submit **only one** python file per group.

```python
'''
Description: CS5340 - The Bayes-Ball algorithm
Name: Your Name, Your partner's name
Matric No.: Your matric number, Your partner's matric number
'''


def create_graph():
    """Reads graph.txt and returns a dictionary
    with nodes as keys and the value is a list of
    nodes that the given node has a directed edge to.

    Returns:
        dict: the graph as a dictionary
    """
    with open('graph.txt', 'r') as g_file:
        K = int(g_file.readline())
        graph = {i: [] for i in range(1, K + 1)}
        for line in g_file:
            i, j = map(int, line.split())
            graph[i].append(j)
    return graph


def read_queries():
    """Reads queries.txt and returns a list of X, Y, Z
    triplets.

    Returns:
        list: the list of queries
    """
    with open('queries.txt', 'r') as q_file:
        queries = []
        for line in q_file:
            X, Y, Z = [], [], []
            x, y, z = line.split()
            X.extend(map(int, filter(bool, x[1:-1].split(','))))
            Y.extend(map(int, filter(bool, y[1:-1].split(','))))
            Z.extend(map(int, filter(bool, z[1:-1].split(','))))
            queries.append([X, Y, Z])
    return queries


def is_independent(graph, X, Y, Z):
```

```
    """Checks if X is conditionally indepedent
    of Y given Z.

    Args:
        graph (dict): the Bayesian network
        X (list): list of nodes in set X
        Y (list): list of nodes in set Y
        Z (list): list of nodes in set Z

    Returns:
        bool: True if X is conditionally indepedent
    of Y given Z, False otherwise.
    """
    # TODO
    return True


if __name__ == '__main__':
    graph = create_graph()
    Qs = read_queries()
    for X, Y, Z in Qs:
        output = 1 if is_independent(graph, X, Y, Z) else 0
        print(output)
```