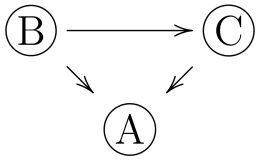
AI 2002 — Artificial Intelligence
Assignment #1

Due Date: Sunday, 25th Feb 2024 8:00 PM on Google Classroom.

Instructions: Assignments are to be done individually. No late assignments will be accepted. You must complete this assignment by yourself. You cannot work with anyone else in the class or with someone outside of the class. The code you write must be your own.
You must **submit a single zip file** containing your notebook and testfiles on Google Classroom named <your_student_id>.zip where <your_student_id> is something like i21-XXXX. This means that you must submit only one file named i21-XXXX.zip containing only your iPython notebook and any test files you created. Each file that you submit must contain your name, student-id, and assignment# on top of the file in comments.
Follow the instructions. Assignments not following the instructions will be awarded zero points.

Search and Problem Solving

Finding optimal graph structures is an important optimization problem that has applications in several areas including machine learning itself, such as Bayesian Network Learning. The problem of Bayes-Net Learning can be simplified as the vertex ordering problem. Given a set of vertices the goal is to determine a vertex ordering with the minimum cost. Consider the network below, which can be represented by the ordering (B, C, A):



Consider the following optimization problem. We are given vertices $V_1, \cdots, V_n$ and possible parent sets for each vertex. Each parent set has an associated cost. Let O be an ordering (a permutation) of the vertices. We say that a parent set of a vertex V_i is consistent with an ordering O if all of the parents come before the vertex in the ordering. Thus, (C, B, A) would be an inconsistent ordering for the above network. Let $mcc(V_i,O)$ be the minimum cost of the parent sets of vertex V_i that are consistent with ordering O. The task is to find an ordering of vertices O that minimizes the total cost of the network: $mcc(V_1, O) + \cdots + mcc(V_n, O)$:

As an example, consider that you are not given the network above. Instead you are given the vertices A,B, and C and their possible parent sets along with their cost as follows and given the task of finding the minimium cost network.

A

parent set	cost
{}	7.5
{B}	2.9
{C}	1.7
{B,C}	1.3

B

parent set	cost
{A,C}	9.8
{}	2.4

C

parent set	cost
{}	12.7
{A}	9.8
{B}	6.3

The cost of ordering (B,C,A) is cost(C→{B}) + cost(A→{B,C})) = 2.4 + 6.3 + 1.3 = 10 whereas the cost of ordering (A,B,C) is cost(C→{B}) + cost(B→{})) = 7.5 + 6.3 + 2.4 = 16.2

Your task is to implement and experiment with the search algorithms(BFS, DFS, UCS) discussed in class for finding a good ordering of the variables. Three datasets are made available for you to experiment on (in addition to the two simple examples described in this document). You will implement the search algorithms in Python and will submit a Python notebook that prints the minimum cost ordering as well as its cost.

Consider the following short data file.

```
1,{},153.466
1,{3},96.093
1,{4},97.913
1,{5},99.835
2,{},141.023
2,{3},122.446
2,{4},121.576
2,{5},123.398
3,{},169.482
3,{1},112.109
3,{2},150.906
3,{1,2},107.516
3,{4},51.681
3,{5},41.775
4,{},169.482
4,{1},113.929
4,{2},150.036
4,{1,2},108.982
4,{3},51.681
4,{5},36.188
5,{},169.802
5,{1},116.171
5,{2},152.178
5,{1,2},111.473
5,{3},42.096
5,{4},36.508
```

which represents the following example:

1

parent set	cost
{}	153.466
{3}	96.093
{4}	97.913
{5}	99.835

2

parent set	cost
{}	141.022
{3}	122.466
{4}	121.576
{5}	123.398

3

parent set	cost
{}	169.482
{1}	112.108
{2}	150.905
{1,2}	107.516
{4}	51.680
{5}	41.775

4

parent set	cost
{}	169.482
{1}	113.929
{2}	150.035
{1,2}	108.981
{3}	51.680
{5}	36.188

Consider the ordering (5, 3, 1, 4, 2). With respect to vertex 1, the parent set {4} is not consistent with the ordering. The parent sets {}, {3}, and {5} are consistent with the ordering

The ordering (5, 3, 1, 4, 2) has a total cost of 96.093 + 121.576 + 41.775 + 36.188 + 169.802 = 465.435. whereas the ordering (1, 2, 3, 4, 5) has a total cost of 153.466 + 141.022 + 107.516 + 51.680 + 36.508 = 490.192

You should create different functions for the each of the search strategies. You will be graded besed on your problem representation, problem decomposition and program hygiene as well as correctness. Did you provide a good structure to the program using functions? Did you minimize the scope of variables to the smallest necessary? Did you use meaningful identifiers? Did you provide comments for your functions? etc.

Honor Policy

This assignment is a individual learning opportunity that will be evaluated based on your ability to think independently, work through a problem in a logical manner solve the problems on your own. You may however discuss verbally or via email the general nature of the conceptual problem to be solved with your classmates or the course instructor, but you are to complete the actual assignment without resorting to help from any other person or other resources that are not authorized as part of this course. If in doubt, ask the course instructor. You may not use the Internet to search for solutions to the problem.