**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI**

Batch No. :

**DEPARTMENT OF COMPUTER SCIENCE AND INFORMATION SYSTEMS**

**Artificial Intelligence (BITS F444/ CS F407)**

**I Semester 2017-18**

**Programming Assignment-5**

**Coding Details**

**(November 28, 2017)**

*Instruction: Type the details precisely and neatly*

1. ID : 2015A7PS0078P

Name : Naveen Venkat

1. Mention the names of Submitted files :
   1. <filename.ext>
   2. <filename.ext>
   3. <filename.ext>
   4. <filename.ext>
   5. <filename.ext>
   6. <filename.ext>
   7. <filename.ext>
2. Total number of submitted files: \_\_\_\_\_\_\_\_\_\_\_
3. Name of the folder : **2015A7PS0078P**
4. Have you checked that all the files you are submitting have your name in the top? **yes**
5. Have you checked that all the files you are submitting are in the folder as specified in 4 (and no subfolder exists)? **yes**
6. Modules implemented
   1. Created the Bayesian network? (yes/no) **yes**
   2. Created Markov blanket?( yes/no) **yes**
   3. Created expression from the inputs read ?(yes/no) **yes**
   4. Computed probability ? (yes/no) **yes**
7. Data structures used
   1. To represent the Bayesian network: **python** **dictionary** *(key being the variable node, vale being the corresponding node)*
   2. To represent Markov blanket: **python list** *(eg. ['A', 'C', 'B', 'D', 'G', 'F', 'H', 'L', 'N', 'Y', 'X'] )*
   3. To represent the variables: **python string** *(each character represents a variable, eg. ‘A’, ‘B’ etc.)*
   4. To represent the expression for probabilistic query: **python dictionary** *(with 3 keys, ‘query’ to keep the query variables, ‘cond’ to keep the condition variables if any, ‘string’ to store the expression as a string)*
8. Implementation Details
   1. How did you create the CPT reading the data from the file? **Read the file, and parsed the conditional probability values from the file. Then created a list with the values. This models the CPT. The index of a value in this list will be the number corresponding to the conditional probability with parents ordered as a bit string.**
   2. How did you access the BN to obtain the Markov blanket? **Since a node in the Bayesian network contains the list of parents, and children, I obtained them directly. Also, to get children’s parents, I traversed through the list of parents for each child and added it to the markov blanket. Finally, I removed duplicates by type casting it to a set and back to a list.**
   3. How did you access the CPTs? **CPTs are stored as a list, where the index describes one of the 2^n models, the parents could be in (where n is the number of parents). Given a model, I converted it to a number using the function getRankFromModel and looked up in the list corresponding to that index. *Further description is in the code.***
   4. How did you expand the expression for the conditional dependence on variables? **I converted the conditional probability expression into ratio of joint probabilities according to bayes rule. ie. P(A,B|C,D) will be evaluated as P(A,B,C,D) / P(C,D). The same function P (defined inside class ProbabilityFinder) takes such expressions and evaluates them. *I have given an elaborate description in the code for finding probability. Kindly have a detailed analysis of the code.***
   5. How did you marginalize the expression? **I expanded along all the parents of the variables given in the joint probability expression. I did not consider the markov blanket, and simply used the parents for this case because they are equivalent for a Bayesian network as mentioned in the textbook. *More description in the code.***
   6. How many terms does a query have? Give example. **For eg. Consider the burglar example. Here the query P(A) will be evaluated as:  
      = P(A,B,E) + P(A,B,~E) + P(A,~B,E) + P(A,~B, ~E)  
      This would produce 4 terms, which will further be evaluated as  
      = P(A|B,E)\*P(B)\*P(E) + P(A|B,~E)\*P(B)\*P(~E) + P(A|~B,E)\*P(~B)\*P(E) + P(A|~B,~E)\*P(~B)\*P(~E)**
9. Graphics: Created the graphics (yes/no) **No**
10. Output
    1. Execute your program to answer the following probabilistic queries. Mention the answer obtained by your program. Also compute the Markov blanket of the variable A.

* P(D, A, L| R, X, P, O) = **0.0997432867636**
* P(A)= **0.227587680582**
* P(F,R|A,P)= **0.128149583593**
* P(D)= **0.472122546785**
* P(D|P)= **0.50652782668**
* P(A|Y, C)= **0.0489561897357**
* P(A,D|O,R,P)= **0.224232103788**
* Markov Blanket of A= **['A', 'C', 'B', 'D', 'G', 'F', 'H', 'L', 'N', 'Y', 'X']**

1. Compilation Details:
   1. Code Compiles (Yes/ No): **yes**
   2. Mention the .py files that do not compile: **N/A**
   3. Any specific function that does not compile: **N/A**
   4. Ensured the compatibility of your code with the specified Python version(yes/no) **yes**
   5. Instructions for compilation of your files mentioning the multi file compilation process used by you:  
        
      run **python driver.py** and follow on screen instructions. The complete code runs in console, providing a neat interface to work on. Enter the file name of the input#.txt in the beginning. The file should be present in the folder in which driver.py resides. **I have provided an elaborate explanation in every step in the code. Kindly have a detailed analysis of the documentation in the code to understand the approach taken in solving the problem.**
2. Driver Details: Does it take care of the options specified earlier(yes/no): **yes**
3. Execution status (describe in maximum 2 lines)**: The code compiles and runs in the console. It has been kept neat and intuitive for easy use. The config.py file contains global constants and flags for execution purposes. Details are mentioned in the code.**
4. Declaration: I, **Naveen Venkat** declare that I have put my genuine efforts in creating the python code for the given programming assignment and have submitted only the code developed by me. I have not copied any piece of code from any source. If the code is found plagiarized in any form or degree, I understand that a disciplinary action as per the institute rules will be taken against me and I will accept the penalty as decided by the department of Computer Science and Information Systems, BITS, Pilani.

ID: **2015A7PS0078P** Name: **Naveen Venkat**

Date: 28/11/2017

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* Should not exceed three pages