Khadija Khaldi, Romita Banerjee and Christoph F. Eick

COSC 4368: Fundamentals of Artificial Intelligence Spring 2019
Problem Set3 (Individual Tasks) Version 4

Deadline: Su. April 28, **11p** (3% bonus); Wednesday, May 1, **11a** (the latest)
Available Points: 57

12. Logical Reasoning (4 points) Khadija

Show using Resolution (and **not** by using other methods!):

(1) ∀x ∀y∃z (P(x,y,z) → R(x,y))
(2) ∀r ∀s (P(s,s,t) → Q(s,t))
(3) ∀a∃b (Q(a,a) → R(b,a))
(4) ∀x ∀y (R(x,y) → R(y,y))
(5) P(4,4,4)
(6) ~Q(4,5)
|-
(X) R(4,4)

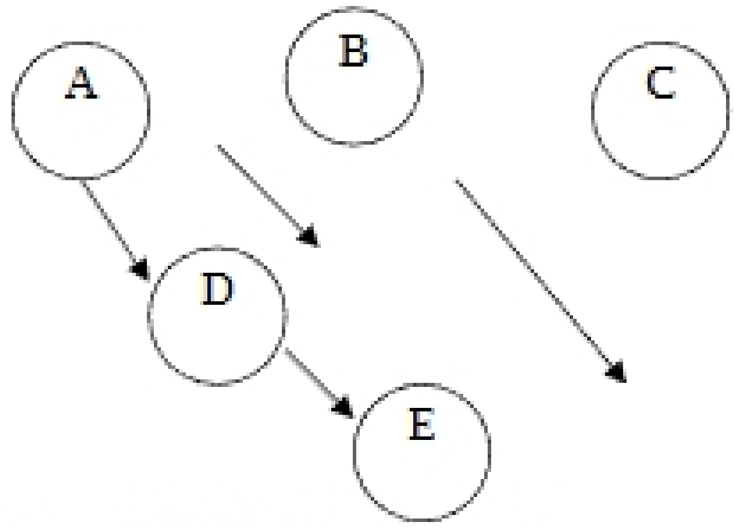
First transform the FOPL formulas into clauses, and then the hunt for the empty clause can begin!

13. Bayes' Theorem and Belief Networks (9 points) Romita

Fig. 1: Thomas Bayes ≈1740—279 years ago!

a) Assume we have 3 symptoms S1, S2, S3, a disease D and the following probabilities: P(D)=0.01 P(S1)=P(S2)=P(S3)=0.02; P(S1|D)=0.1; P(S2|D)=0.02; P(S3|D)=0.002. How would a naïve Bayesian system compute the following probability? **P(D|S1,S2,S3)=…**
b) Now assume the following additional knowledge has become available: P(S1,S2)=0.0002; P(S3|S1,S2)=0.08; P(S1,S2,S3|D)=0.000032; how would you use this information to obtain a "better" estimation of P(D|S1,S2,S3)?
c) How can the discrepancy with respect to the obtained probabilities between cases a) and b) be explained? Why are the numbers you obtain different? What does this discrepancy tell you about naïve Bayesian systems in general?

d) Assume that the following belief network is given that consists of nodes A, B, C, D, and E that can take values of true and false.



- Is A|D,B d-separable from C|D,B? Give reasons for your answer!
- Is A,E|∅ d-separable from C|∅? Give reasons for your answer! ∅:="no evidence"

14. Weight Learning Computations in Neural Networks (7 points) Khadija

Assume we a 3-layer NN that is depicted in Fig. 2 is given that has 3 inputs X1, X2, and X3.

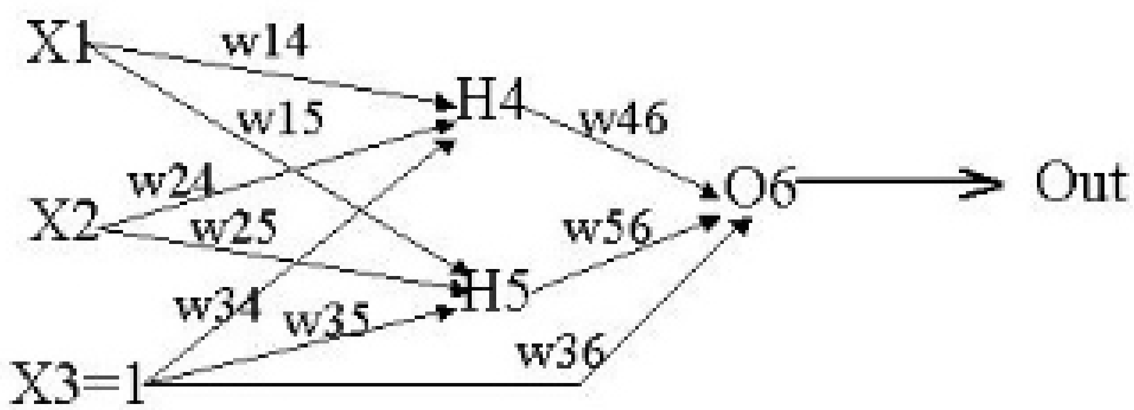


Figure 2

A training set D is given that consists of the following examples: (1,1,1,1) and (0, 0, 1, 0), indicating values for X1, X2, X3 and the "desired" activation for O6. Assume that the initial/current weights are *w*14 = 0.2, *w*15 = 0.2, *w*24 = 0.2, *w*25 = 0.2, *w*34 = 0.2, *w*35 = 0.2, *w*36 = 0.5, *w*46 = 0.5, and *w*56 = 0.5. Use □=0.5 as your learning rate. Show what new weights will be obtained after processing the two training instances of data set D. Use g(x) = 1/(1+e**(-x)) as the activation function; that is: g'(x)=(e**(-x))/(1+e**(-x))**2). List all important computations that lead to the reported weight updates.

S#	X1	X2	X3	O6	Des-O6	Error	w14	w15	w24	w25	w34	w35	w36	w46	w56
0							0.2	0.2	0.2	0.2	0.2	0.2	0.5	0.5	0.5
1	1	1	1		1										
2	0	0	1		0										

15. Using a Belief Network Tool (20 points) Khadija

Fig. 3: Multiple Astronomers Looking at the Sky

Assume we have 3 astronomers in different parts of the world who make measurements M1, M2, and M3 of the number[1] of stars N in some region of the sky. Normally, there is a probability of 0.05 that the astronomer counts a single star twice (overcounts by one star; you can assume that the three astronomers never undercount; moreover, if there is no star visible (N=0) the astronomer never overcounts). Moreover, there is a 10% probability ($P(F_i=1)=0.1$ for i=1,2,3) that a telescope is out of focus (represented using random variables F1, F2, and F3), in which the astronomer undercounts by 2 or more stars (e.g. if N is 4 and her telescope is out of focus, the astronomer will count 2, 1 or 0 stars; you can assume if information is missing that each case has the same probability). Design a belief network, and compute the probability of the other variables assuming the following pieces of evidence are given (feel free to use *Netica* (http://www.norsys.com/download.html) or any another belief network tool to compute your answer[2]!):

1. M1=3 M2=3 M3=1
2. M1=3 M2=3 M3=0
3. N=4, M2=1, M3=0
4. M1=0 M2=3 M3=2
5. N=3 F1=0 F2=0 F3=1
6. M1=4 M2=4 F3=1

Submit the complete Belief Network you created—including all its probability tables—, and the findings you obtained for the six cases listed above!

1 You can assume that N is limited to 4—but the astronomer do not know that: M1, M2 and M3 are therefore limited to values 0 through 5.

2 Including the answer 'inconsistent' in the case that the evidence is inconsistent, e.g, the evidence N=1 M1=3 is inconsistent—as it is 'impossible' because astronomer1 never overcounts by more than 1 star!