NTNU – Trondheim
Norwegian University of
Science and Technology

Department of Computer and Information Science

Examination paper for TDT4171 Methods in AI

Academic contact during examination: Gleb Sizov
Phone: 735 91802

Examination date: June 5th, 2013
Examination time (from-to): 0900 - 1300
Permitted examination support material: *D*
No written and handwritten examination support materials are permitted.
A specified, simple calculator is permitted.

Other information:
Deadline for grading: June 26th, 2013.
All questions (including all sub-questions) shall be answered. Each question is weighted as indicated in the text.

Language: English
Number of pages: 4 (including this cover page)
Number of pages enclosed: 0

Checked by:

25/05 – 2013 ______________________________
Date Signature

<u>Question 1</u> – 35%

a) Describe the syntax and the semantics of Bayesian Networks.

b) Model the following domain using a Bayesian network. Make your model as simple and easy to understand as possible:

Milk from a cow may be infected with probability 0.01. There is a test to detect whether the milk is infected, which may give either a positive result (indicating milk is infected) or a negative result (indicating the milk is clean). The test is not perfect: It gives a positive result on clean milk with probability 0.05 and a negative result on infected milk with probability 0.02.

You are asked to give the graphical structure, ***as well as*** the conditional probability distributions here.

c) A test is done to examine the milk, and comes back with the positive result, i.e., indicating that the milk is infected. Calculate the probability that the milk is indeed infected, given the test result. Use the probabilities in the model built in Question (b).

d) Extend your model from Question (b) as follows:

After getting the test result, the farmer must decide to either store the milk or throw it out. If he stores the milk, and it is not infected, he will earn $100 when selling it later. If he stores it, and it is infected, he will contaminate all the milk he has stored already, thereby loosing $1000 in total. If he throws out the milk, he will earn $0, both if the milk is clean and if it is infected.

You are asked to give the graphical structure, ***as well as*** the quantitative (numeric) parts of the model.

e) What do we call the type of model you have created in Question (d), and how can the farmer use it to make his best decision? Do the calculations, and propose an action for the farmer (remember that we know that the test result was positive). Which principle do you use to come up with the recommended action?

Note: If you were unable to calculate *P(milk is infected | test result is positive)* in Question (c) you can just assume a value for this probability now and make your calculations from there on. State clearly which number you have used.

f) Extend your model from Question (b) to include the following information:

It is known that the more milk is infected when the weather is humid. Further, the test is less reliable when the air-temperature is high.

You are only asked to give the graphical structure, ***not*** the conditional probability distributions in this sub-question.

g) Do you think that Bayesian Networks constitute a *natural* modeling framework for this problem domain? How can one characterize problem domains where Bayesian networks can be used with success? Can you give an example of a problem domain where Bayesian networks do *not* fit well?

Question 2 – 15%

a) *Describe* the main ideas of case-based reasoning (CBR). You can for instance proceed by comparing CBR to a decision tree learner to highlight your points.

b) Briefly describe what happens during each of the four steps of the CBR-cycle, and discuss if and how general domain knowledge can be utilized inside each of these steps.

Question 3 – 20%

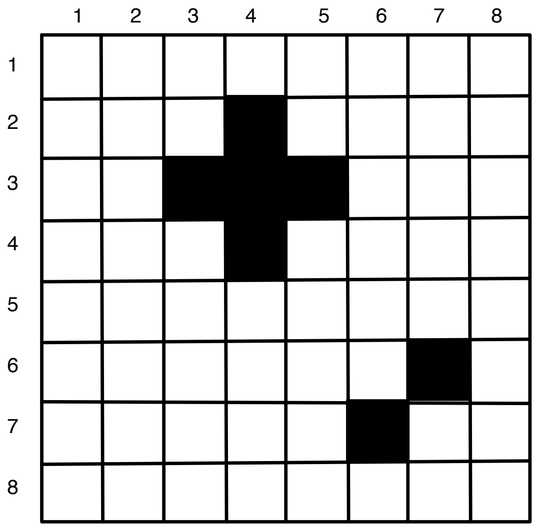
a) One of the assumptions underlying the *Hidden Markov Model* is known as the Markov assumption. Explain why this assumption is made. Show the structure of a Hidden Markov Model, and indicate what the structure would look like if the assumption was *not* made.

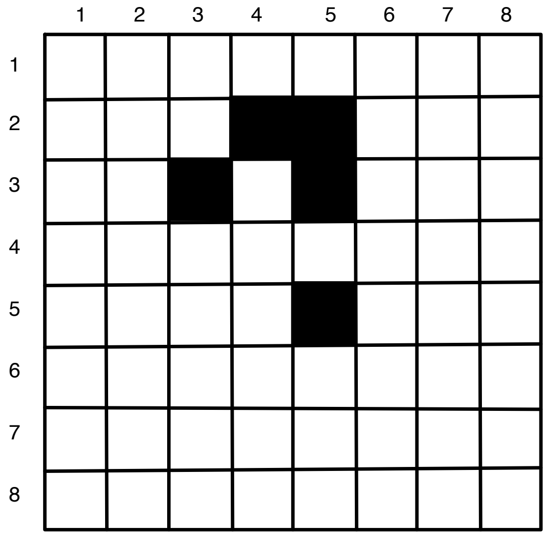
b) In which situations can you use a Markov Decision Process to solve your decision problem? Give a short general description, and give a concrete example. What is the main idea of the *value iteration* algorithm, which is used to solve Markov Decision Processes.

c) Discuss the statement "*The success of AI might mean the end of the human race*". What would be the reasons for this statement, and how can one prevent this from happening? Do *you* think that the statement is true? Why (not)?

<u>Question 4</u> – 30%

You are asked to analyze a large number of binary images. Each image consists of 8 times 8 binary pixels in a grid, and you can read out the value of the pixel at each location, where white pixels are encoded as one and black pixels are coded as zero. Two examples are shown below.





For the image on the left-hand side, you will for instance get the information that the value at (1,1) =1 because the upper left corner is white and (2,4)=0, because the element at position (2,4) in the picture is black. Your task is to create a system that automatically finds the images that contain black pixels that build up a cross. This is true for the picture on the left-hand side, since the pixels (2,4), (3,3), (3,4), (3,5) and (4,4) all are black. The image on the right hand side does not contain a cross.

a) Instead of building a model by hand, you should prefer to use a learning technique to solve this problem. What characterizes situations where it is useful to learn models from data (instead of defining them by hand)?

b) Would you prefer to use artificial neural networks or decision trees for this task? Give reasons for your selection.

c) Describe how you would proceed with the learning phase:
   - How would you structure the input data?
   - Which learning algorithm will you use, and what does this algorithm do? Give a *brief* description.

d) You are asked to adapt your solution to not only detect if a cross is present in each image, but also to determine *how many crosses* there are in each picture. How would you adapt your previous solution to this new problem?