Examination in Machine Intelligence

Thomas Dyhre Nielsen

January 8, 2018

On the next six pages you will find six questions covering different aspects of the course. The questions differ in their level of difficulty, and for each correctly answered question you will get a certain amount of points as indicated by each question. When solving the questions you are allowed to use all available material such as books, pocket calculator, etc., however, laptops/tablets and other networking devices are *not* allowed.

Before you answer a question make sure that you have read the question carefully. Moreover, make sure that you argue for your answers (e.g. include intermediate results) so that it is possible to follow your line of thought. Finally, it is important that your solutions are presented in a readable form. The answers to the questions should be written in English.

In addition to the six pages with questions, you are also provided with 10 pages that you can use when writing your answers to the questions.

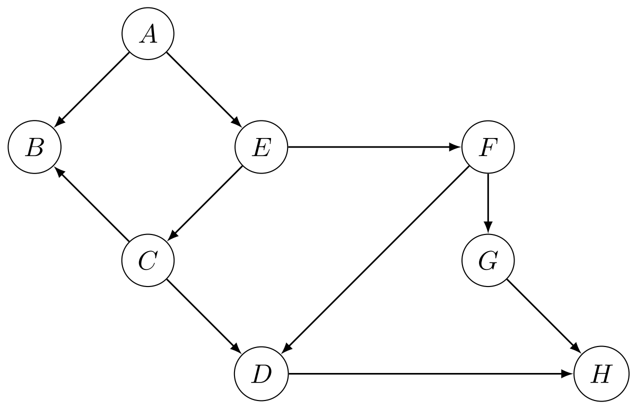
- Two of the pages contain the grid world found in Question 2. You can use these grids when answering this question.
- For each sheet of paper containing your response to the questions, please include your name, study number, current page number, and the total number of pages.
- If you need more paper, simply raise your hand to contact one of the guards in the examination room.

Good luck with the questions

Thomas Dyhre Nielsen

Question 1 - 10 points

Consider the graph below:



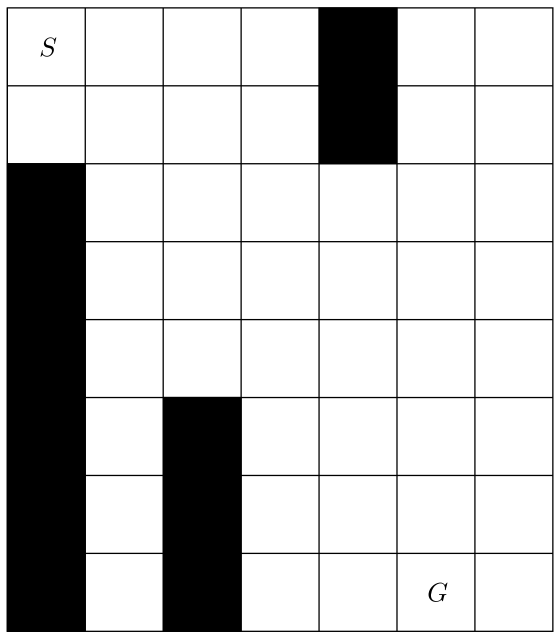
1. List all probability distributions (on the form $P(X|Y_1,\ldots,Y_n)$) that should be specified in order to obtain a Bayesian network from the graph?
2. Assume that the variable D has three states labeled $\{a,b,c\}$ and that the remaining variables (i.e., $\{A,B,C,E,F,G,H\}$) all have two states labeled t and f. Give a table representation of a valid conditional probability distribution for variable D.
3. Which variables are d-separated from H given hard evidence on D and E?
4. Which variables are d-separated from A given hard evidence on B and E?
5. Which variables are d-separated from C given hard evidence on E and H?

Solution:

1. $P(A)$, $P(B|A,C)$, $P(C|E)$, $P(D|C,F)$, $P(E|A)$, $P(F|E)$, , $P(G|F)$, $P(H|G,D)$
2. ...
3. $\{A\}$
4. $\{G,F\}$
5. $\{A\}$

Question 2 - 15 points

Consider a robot that can move in the grid shown below. The robot can only move *right* or *down* and only one step at a time; no step can be made into the shaded areas or outside the grid.



1. Use dynamic programming to calculate the number of paths leading to the goal cell G from each of the cells in the grid.
2. Specify the path the robot should take from S to G so that at each step the robot moves to the cell that has the maximum number of paths leading to G.

Solution:

S 384	236	88	21		1	0
148	148	67	21		1	0
	81	46	21	6	1	0
	35	25	15	5	1	0
	10	10	10	4	1	0
	0		6	3	1	0
	0		3	2	1	0
	0		1	1	G	0

Question 3 - 20 points

The software company *Macrosoft* needs to upgrade four of its key software systems. The software systems (labeled S_1, S_2, S_3, and S_4) should be upgraded over the course of a single day and, to keep things organized, seven possible starting times (labeled 1, 2, 3, 4, 5, 6, and 7) for the software updates have been identified. The software systems are, however, inter-connected, and these connections induce constraints on when the upgrades can start:

1. S_1 should start after S_2.
2. S_3 should start before S_2 and S_4.
3. S_4 should start before time 7.
4. S_4 must start at least two time slots after S_1.

You should:

1. Model the problem as a constraint satisfaction problem, i.e., identify the variables, their domains, and the constraints.
2. Represent the problem as a constraint network.
3. Make the network arc-consistent.
4. Find a satisfying solution to the problem (if one exists) using variable elimination with the elimination ordering S_1, S_2, S_3, S_4.

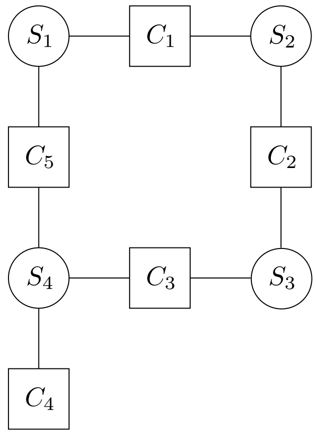
Solution:

Sub-problem 1 and 2

There is one variable for each system, and the state space of each of these variables corresponds to the time slots $\{1, 2, 3, 4, 5, 6, 7\}$. The constraints for the problem can be specified as:

1. C_1: $S_1 > S_2$
2. C_2: $S_3 < S_2$
3. C_3: $S_3 < S_4$
4. C_4: $S_4 < 7$
5. C_5: $S_1 + 2 \leq S_4$

The constraint network for this problem is shown in the figure below.



Sub-problem 3

After making the network arc-consistent we end up with the following domains:

- $dom(S_1) = \{3, 4\}$
- $dom(S_2) = \{2, 3\}$
- $dom(S_3) = \{1, 2\}$
- $dom(S_4) = \{5, 6\}$

Sub-problem 4

Eliminating S_1

$$C_6 = (C_1 \bowtie C_5)^{\downarrow S_2, S_4} = \left(\begin{array}{c|c} S_1 & S_2 \\ \hline 3 & 2 \\ 4 & 2 \\ 4 & 3 \end{array} \bowtie \begin{array}{c|c} S_1 & S_4 \\ \hline 3 & 5 \\ 3 & 6 \\ 4 & 6 \end{array} \right)^{\downarrow S_2, S_4} = \begin{array}{c|c} S_2 & S_4 \\ \hline 2 & 5 \\ 2 & 6 \\ 3 & 6 \end{array}$$

Eliminating S_2

$$C_7 = (C_2 \bowtie C_6)^{\downarrow S_3, S_4} = \left(\begin{array}{c|c} S_2 & S_3 \\ \hline 2 & 1 \\ 3 & 1 \\ 3 & 2 \end{array} \bowtie \begin{array}{c|c} S_2 & S_4 \\ \hline 2 & 5 \\ 2 & 6 \\ 3 & 6 \end{array} \right)^{\downarrow S_3, S_4} = \begin{array}{c|c} S_3 & S_4 \\ \hline 1 & 5 \\ 1 & 6 \\ 2 & 6 \end{array}$$

Eliminating S_3

$$C_8 = (C_3 \bowtie C_7)^{\downarrow S_4} = \left(\begin{array}{c|c} S_3 & S_4 \\ \hline 1 & 5 \\ 1 & 6 \\ 2 & 5 \\ 2 & 6 \end{array} \bowtie \begin{array}{c|c} S_3 & S_4 \\ \hline 1 & 5 \\ 1 & 6 \\ 2 & 6 \end{array} \right)^{\downarrow S_4} = \begin{array}{c} S_4 \\ \hline 5 \\ 6 \end{array}$$

By back tracking we find multiple solutions:

S_1	S_2	S_3	S_4
3	2	1	5
3	2	1	6
4	2	1	6
4	3	1	6
4	3	2	6

Question 4 - 15 points

Consider the variables A, B, and C with state spaces $sp(A) = \{a_1, a_2\}$, $sp(B) = \{b_1, b_2, b_3\}$, and $sp(C) = \{c_1, c_2\}$, respectively. The joint probability distribution over the three variables is defined by the conditional probability distributions assigned to the variables:

$P(A) =$

A	
a_1	a_2
0.6	0.4

$P(B) =$

B		
b_1	b_2	b_3
0.4	0.5	0.1

$P(C|A, B) =$

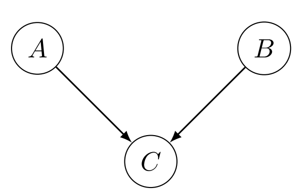
		B		
		b_1	b_2	b_3
A	a_1	$(0.9, 0.1)$	$(0.2, 0.8)$	$(0.4, 0.6)$
	a_2	$(0.7, 0.3)$	$(0.6, 0.4)$	$(0.1, 0.9)$

You should:

1. Show the structure of the Bayesian network representation for the variables A, B, and C.
2. Calculate the probability table $P(A, B, C = c_1)$.
3. Calculate the probability table $P(A, C = c_1)$.
4. Calculate the conditional probability distribution $P(A \,|\, C = c_1)$.
5. Calculate the conditional probability distribution $P(A \,|\, B = b_2, C = c_1)$.
6. Based on the calculated probability distributions, argue whether A is independent of B given $C = c_1$.

Solution:

Sub-problem 1



Sub-problem 2

$$P(A, B, C = c_1) = P(C = c_1 \mid A, B)P(A)P(B)$$

=

		B		
		b_1	b_2	b_3
A	a_1	0.216	0.06	0.024
	a_2	0.112	0.12	0.004

Sub-problem 3

$$P(A, C = c_1) = \sum_B P(A, B, C = c_1)$$

$= \sum_B$ (

		B		
		b_1	b_2	b_3
A	a_1	0.216	0.06	0.024
	a_2	0.112	0.12	0.004

)

=

A	
a_1	a_2
0.3	0.236

Sub-problem 4

$$P(A \mid C = c_1) = \frac{P(A, C = c_1}{\sum_A P(A, C = c_1)}$$

$\approx$

A	
a_1	a_2
0.5597	4403

Sub-problem 5

$$P(A \mid B = b_2, C = c_1) = \frac{P(A, B = b_2, C = c_1)}{\sum_A P(A, B = b_2, C = c_1)} = (1/3, 2/3)$$

Sub-problem 6

A is not independent of B given $C = c_1$ since $P(A \mid B = b_2, C = c_1) \neq P(A \mid C = c_1)$.

Question 5 - 20 points

The book store *Smart books* wants to make book recommendations for its online customers. The store has recorded customer feedback (*Likes* with the states *yes* and *no*) for prior book purchases. For these purchases, the book store has recorded the following information about the books: The genre of the book (*Genre* with states *action*, *biography*, and *romance*), the format of the book (*Format* with states *paperback* and *hardcover*), the length of the book (*Length* with states *long* and *short*), and the price (*Price*) of the book (in DKK). The data that has been collected is shown in the table below.

	Attributes				Target
	Genre	*Format*	*Length*	*Price*	*Likes*
1	*bio*	*paperback*	*long*	110	*no*
2	*bio*	*paperback*	*short*	100	*no*
3	*action*	*hardcover*	*long*	70	*yes*
4	*romance*	*paperback*	*long*	80	*yes*
5	*romance*	*hardcover*	*short*	90	*yes*
6	*romance*	*hardcover*	*long*	105	*no*

In order to compare previous book purchases, the book store has defined distances for the states of the features characterizing a book:

- For the *Price* feature, the book store defines the distance as the absolute difference in prices.
- For the three discrete features, the book store uses the following distances for the states of the features:

Genre	*bio*	*action*	*romance*
bio	0	2	1
action	2	0	1
romance	1	1	0

Format	*paperback*	*hardcover*
paperback	0	1
hardcover	1	0

Length	*Short*	*Long*
Short	0	1
Long	1	0

The total distance between two instances representing a book is the sum of the distances of the four features with the price distance weighted with 0.1.

1. A new book is characterized by the feature values $\langle bio, hardcover, short, 90\rangle$. Calculate its distance to the six recorded instances in the table.
2. Classify the new book according to the 1-nearest-neighbor-rule.
3. Classify the new book according to the 3-nearest-neighbor-rule.

Solution:

Sub-problem 1

The distances are:

Instance	Distance
1	4
2	2
3	5
4	4
5	1
6	3.5

Sub-problem 2

The closets neighbor is instance 5. The new book will therefore be classified as *yes* according to the 1-NN.

Sub-problem 3

The three closets neighbors are instances 2, 5, and 6. The new book will therefore be classified as *no* according to the 3-NN.

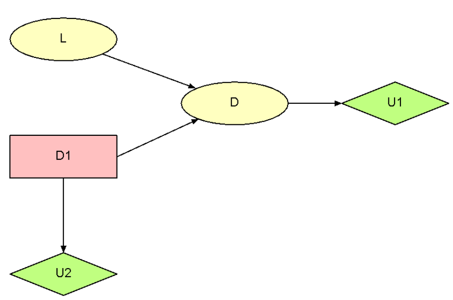
Question 6 - 20 points

On a vacation near the sea, you consider going on a guided tour in the hope of seeing dolphins. You have been told that there is a probability of 0.7 that there are dolphins at the location where the tour is heading, with probability 0.1 there are many dolphins, and with probability 0.2 there are no dolphins. However, even if there are dolphins at the location, there is no guarantee that you will actually see any dolphins: if there are some dolphins in the area you estimate that the probability of seeing a dolphin is 0.6, but if there are many dolphins then the probability increases to 0.9. The tour costs 200 DKK and you estimate that seeing a dolphin will be worth 500 DKK to you.

1. You should decide (D) whether you should *go* on the tour or *stay* at home.
   (a) Construct an influence diagram for your decision problem based on the description above. This include specifying the graphical structure and the probability and utility tables.
   (b) Calculate the expected utility of *go* and *stay*. Which decision should you take in order to maximize the expected utility.
2. Before deciding on whether to go, an experienced local sailor offers you accurate information about the dolphin population in the area (i.e, whether there *some*, *many*, or *no* dolphins). The cost of the information is 30 DKK. Determine whether it is worth to pay for the information by treating this information gathering decision as a value of information problem.

Solution:

Sub-problem 1



The probability and utility tables are given by: $P(L) = (0.2, 0.7, 0.1)$, $C(D) = (-200_{yes}, 0_{no})$, $U(F) = (500_{yes}, 0_{no})$, and

$$P(DF|D,L) = \begin{array}{r|ccc} & L = no & L = some & L = many \\ \hline D = go & (0,1) & (0.6, 0.4) & (0.9, 0.1) \\ D = stay & (0,1) & (0,1) & (0,1) \end{array}$$

First we marginalize out L:

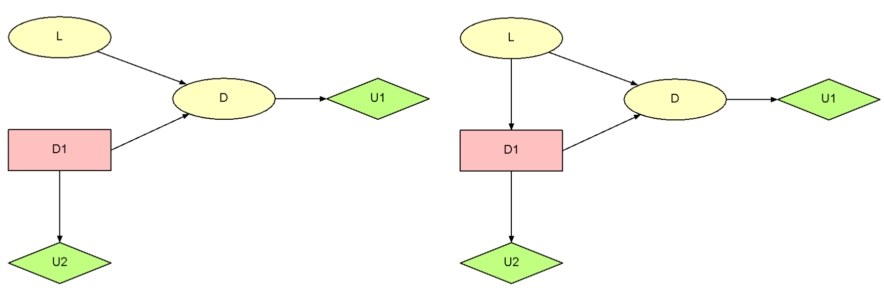
$$\begin{aligned} P(DF|D) &= \sum_L P(DF, L \mid D) \\ &= \sum_L P(DF \mid L, D) P(L) \\ &= \begin{array}{r|cc} & DF = yes & DF = no \\ \hline D = go & 0.51 & 0.49 \\ D = stay & 1 & 0 \end{array} \end{aligned}$$

This yields the expected utilities:

$$\begin{aligned} EU(go) &= 0.51 \cdot 500 + 0.49 \cdot 0 - 200 \\ &= 55 \\ EU(stay) &= 0 \end{aligned}$$

Sub-problem 2

The value of information problem basically boils down to calculating the differences in maximum expected utility between the two models:



The expected utility of the left-most model was calculated above.

The expected utility of the right model is given by:

$$EU = \sum_L P(L) \max_D (\sum_{DF} P(DF \mid D, L) U(DF) + U(C))$$

For the first part ($\sum_{DF} P(DF \mid D, L)U(DF) + U(C)$) of the calculations we get the expected utility $EU(L, D)$:

$$EU(L,D) = \begin{array}{r|ccc} & L = no & L = some & L = many \\ \hline D = go & -200 & 100 & 250 \\ D = stay & 0 & 0 & 0 \end{array}$$

By marginalizing out L and D we find:

$$EU = 95$$

and by also taking the cost (30DKK) of the information into account, we have a value of information given by $95 - 55 - 30 = 10$. Since the value is positive you should pay for the information.