This question paper consists of 16 printed pages, each of which is identified by the Code Number COMP5830M.

A non-programmable calculator may be used.
Answer All Questions.
Open Book.
Course notes are permitted.

School of Computing

January 2016

COMP5830M

KR+ML: KNOWLEDGE REPRESENTATION AND MACHINE LEARNING (MSc)

Time allowed: 2 hours and 15 minutes

PLEASE DO NOT REMOVE THIS PAPER FROM THE EXAM ROOM

Answer ALL FOUR questions

The marks available for each part of each question are clearly indicated.

TURN OVER FOR QUESTIONS

Question 1

(a) Translate the following sentences into *First-Order Predicate Logic* (using equality where necessary):

(i) All my friends know each other. **[2 marks]**

Answer: $\forall x \forall y[(\mathsf{Friend}(x, \mathsf{me}) \wedge \mathsf{Friend}(y, \mathsf{me}) \rightarrow \mathsf{Knows}(x, y)]$.

(ii) Unicycles have just one wheel. **[3 marks]**

Answer: $\forall x[\mathsf{Unicycle}(x) \rightarrow \exists y[\mathsf{Wheel}(y) \wedge \mathsf{HasPart}(x, y) \wedge \forall z[(\mathsf{Wheel}(z) \wedge \mathsf{HasPart}(x, z)) \rightarrow (y = z)]\ ]\ ]$

(b) Using the *Sequent Calculus* (as specified in the module notes), determine whether the following sequent is valid:

$$A, \neg(A \wedge \neg B),\ B \rightarrow C \ \vdash \ C$$

[6 marks]

Answer: The sequent is valid, as shown by the following proof:

```
    Axiom               Axiom
-------------       --------------
A, -B |- C, A       A, -B |- C, -B                                Axiom
------------------------------------ |- &           -------------------
           A, -B |- C, (A & -B)                      A, C |- C, (A & -B)
           --------------------------------------------------------------  v |-
                         A,  -B v C |- C, (A & -B)
                         -------------------------  -> rw
                         A,  B -> C |- C, (A & -B)
                         -------------------------  - |-
                         A, -(A & -B), B -> C |- C
```

1 mark for each correct rule application. For full marks complete proof is required with Axioms at the top. Some credit may be given for almost correct rule applications.

(c) Give a representation of the following statements in *Propositional Tense Logic*:

- I shall never eat haggis, unless I am starving. **[2 marks]**
  Answer: $\mathbf{G}(\mathsf{IeatHaggis} \rightarrow \mathsf{IamStarving})$
- I knew Tom before he first met Helen. **[2 marks]**
  Answer: $\mathbf{P}(\mathsf{IKT} \wedge \neg\mathsf{TMH} \wedge \neg\mathbf{P}\mathsf{TMH})$

[Question 1 total: 15 marks]

Question 2

(a) The Situation Calculus is used to define a theory describing the actions of a robot that is used to move objects around a factory.

Some of the objects that must be moved are 'heavy'. The robot can pick up any item. However, if it is carrying a heavy item it cannot move to an adjacent room. To move a heavy item, the robot must pick it up and load it onto a trolley. It can then hook itself up to the trolley. While the robot is hooked up, whenever it moves from one location to another, the trolley and any object on the trolley will also move with the robot. The robot can also move while carrying any non-heavy item even if it is also hooked up to the trolley.

Together with the usual Holds and Poss predicates, the theory is stated in terms of the following fluents:

$$\textit{located}(\textit{item}/\textit{robot}/\textit{trolley},\ \textit{room}), \quad \textit{carrying}(\textit{item}), \quad \textit{on\_trolley}(\textit{item}), \quad \textit{hooked\_up}$$

and the following static predicates:

$$\textit{heavy}(\textit{item}), \quad \textit{adjacent}(\textit{room1}, \textit{room2}).$$

The constants include robot and trolley as well as further constants denoting rooms and various items such as hammer, anvil, etc.. Items that are carried or on the trolley are also considered as located in the room occupied by the robot or the trolley.

The theory incorporates the following actions:

pickup(*item*)	**load**(*item*)	**hook-up**	**move**($room1, room2$)
drop(*item*)	**un-load**(*item*)	**un-hook**	

(i) Using the specified vocabulary, write a suitable *precondition axiom* specifying the conditions under which an action **move**($room1, room2$) is possible. **[2 marks]**

Answer: $Poss(move(r1, r2), s) \leftarrow Holds(located(robot, r), s) \wedge adjacent(r1, r2) \wedge$
$\neg\exists x[heavy(x) \wedge Holds(carrying(x), s)]$

(ii) Using the specified vocabulary, write down appropriate *effect* axioms specifying the consequences of the **move**($room1, room2$) action.
Your axioms should describe the effect on the robot, the trolley and any items involved, and should take into account the different effects that may occur depending on whether or not the robot is hooked-up to the trolley. **[6 marks]**

Answer: $Holds(located(\mathsf{robot}, r2), result(move(r1, r2), s)$ [1]
$Holds(carrying(x, s) \rightarrow Holds(located(x, r2), result(move(r1, r2), s)$ [2]
$Holds(hooked\text{-}up, s) \rightarrow Holds(located(trolley, r2), result(move(r1, r2), s)$ [2]
$Holds(hooked\text{-}up, s) \wedge Holds(on\text{-}trolley(x), s) \rightarrow$
$Holds(located(x, r2), result(move(r1, r2), s)$ [1]

(b) For each of the following *Prolog* queries, give the value of the variable `X` after the query has been executed:

(i) `?- X = 12/3.` **[1 mark]**

(ii) `?-  L = [a,b,c,d], [_ | Z] = L, [X | _] = Z.` **[1 mark]**

(iii) `?- L = [a,b], setof( [A,B], (member(A,L), member(B,L)), X).` **[2 marks]**

Answer: `i) 12/3`
`ii) b`
`iii) [[a,a],[a,b],[b,a],[b,b]]`

(c) Consider the following formula involving topological relations of the *Region Connection Calculus* (RCC) and the *convex hull* function, conv. Draw a configuration of the regions a, b and c that satisfies the following formula: **[3 marks]**

$$\mathsf{NTPP}(b, a) \wedge \mathsf{NTPP}(c, conv(b)) \wedge \mathsf{DC}(c, a)$$

Label your diagram to indicate which region is which.

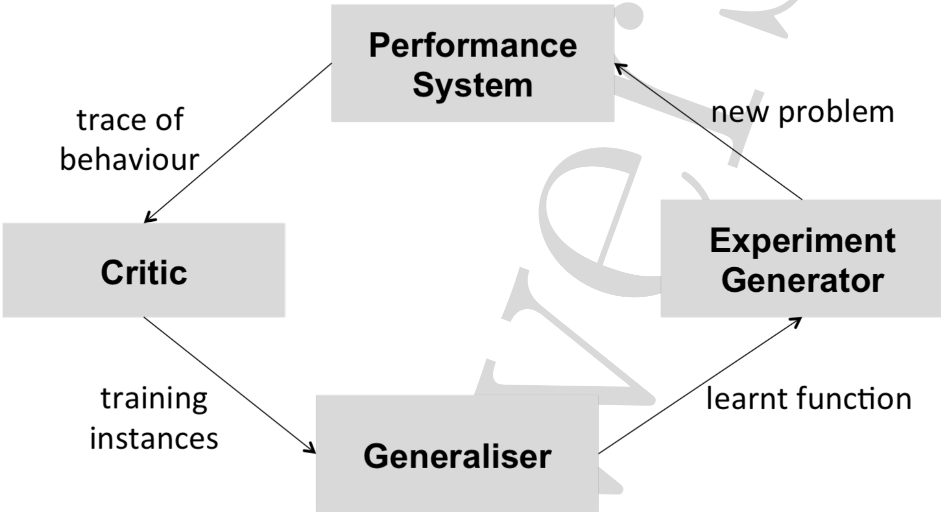
Answer: 1 mark for each of the conjuncts that are true in the given diagram. A possible diagram is the following:

```
----------------------------
|                        a |
|    -------------------   |
|    |         b       |   |
|    |    --------------   |
|    |    |                |
|    |    |   --------------
|    |    |   |
|    |    |   |  -----
|    |    |   |  | c |
|    |    |   |  |   |
|    |    |   |  -----
|    |    |   |
|    |    |   --------------
|    |    |                |
|    |    ---------------  |
|    |                  |  |
|    --------------------  |
|                          |
---------------------------|
```

[Question 2 total: 15 marks]

Question 3

(a) In the following block diagram showing the general structure of a machine learner, explain in detail and using your own words and examples the meaning of any **two** of the components. **[3 marks]**



(b) Consider the following results from a machine learning algorithm:

Test case	Actual class	Predicted class
1	a	c
2	b	b
3	b	b
4	c	c
5	a	b
6	b	a
7	b	c
8	a	a

The *Predicted class* column shows the result from the machine learning algorithm, whilst the *Actual class* column shows the true class for the test case.

Draw a confusion matrix for the above data, making clear what the rows and columns denote. Which classes, if any, are confused? **[2 marks]**

(c) Consider the following dataset, where $\mathbf{X_1}$, $\mathbf{X_2}$ and $\mathbf{X_3}$ are input binary random variables, and $\mathbf{Y}$ is a binary output whose value we want to predict:

$\mathbf{D}$	$\mathbf{Y}$	$\mathbf{X_1}$	$\mathbf{X_2}$	$\mathbf{X_3}$
D_1	1	1	0	1
D_2	1	0	1	0
D_3	1	1	1	1
D_4	0	0	0	0
D_5	0	1	0	1

Given the input $[\mathbf{X_1}, \mathbf{X_2}, \mathbf{X_3}] = [0, 0, 1]$, what would a Naïve Bayes classifier predict for $\mathbf{Y}$? You must show how you calculate your answer. (4 marks are for showing the calculations and explanations. 1 mark for the correct answer.) **[5 marks]**

(d) Louis van Gaal, the manager of Manchester United has a tough decision to make regarding his strikers. In order to help decide which ones to play given their current form, he has collected the team wins over the last 6 matches and whether the striker played or not.

Match	Rooney	Martial	Depay	Win
1	no	no	no	no
2	no	no	yes	no
3	yes	yes	yes	no
4	yes	no	yes	yes
5	yes	no	no	no
6	yes	yes	no	yes

Can you help Louis decide which striker is a key factor by constructing an *ID3* decision tree based on the last 5-matches data?

- Compute the information gain for Depay. **[4 marks]**
- Given that the information gain for Rooney is 0.251 and for Martial is 0.044, which player would you use for the initial split at the root of the tree? **[1 mark]**

Notes:

- To answer this question, you do not need to draw the whole tree, but only to compute which attribute to split on at the root of the tree.
- If your calculator does not have a function for $\log_2$ then you may use the approximation $\log_2(n) = 1.4427 \times \ln(n)$ or $\log_2(n) = 3.322 \times \log_{10}(n)$.

[Question 3 total: 15 marks]

Answer: Answers for Q3

(a) – Performance system: The problem solver, takes as input an instance of a new problem and produces an output solution, a trace of the system behaviour given the inputs, based the learnt function.

– Critic: Takes as input the history or a trace of it and outputs the output of the target function, hence creating training examples for the

– Generaliser: Takes the training examples and improves the learning algorithm to better estimate the target function

– Experiment generator: Takes as input the current learnt function and outputs a new problem that help the system learn better and faster (might for example create new problems from only winning games, e.g. candidate-elimination algorithm).

[2]

(b)

	a	b	c
a	1	1	1
b	1	2	1
c	0	0	1

[1]

The rows are the actual classes and the columns the predicted classes. The classes 'a' and 'b' are confused with each other whilst 'c' is perfectly classified. [1]

(c)

$$
\begin{aligned}
p(Y=0) &= 2/5 && (1)\\
p(Y=1) &= 3/5 && (2)\\
p(X1=0|Y=0) &= 1/2 && (3)\\
p(X1=0|Y=1) &= 1/3 && (4)\\
p(X2=0|Y=0) &= 2/2 && (5)\\
p(X2=0|Y=1) &= 1/3 && (6)\\
p(X3=1|Y=0) &= 1/2 && (7)\\
p(X3=1|Y=1) &= 2/3 && (8)
\end{aligned}
$$

[2]

Predicted Y maximizes:

$$p(X1=0|Y)p(X2=0|Y)p(X3=1|Y)p(Y) \tag{9}$$

[1]

For Y=0:

$$
\begin{aligned}
p(X1=0|Y=0)p(X2=0|Y=0)p(X3=1|Y=0)p(Y=0)\\
= 1/2 * 2/2 * 1/2 * 2/5\\
= 0.1
\end{aligned} \tag{10}
$$

[1]

For Y=1:

$$\begin{aligned} p(X1=0|Y=1)p(X2=0|Y=1)p(X3=1|Y=1)p(Y=1) \\ = 1/3*1/3*2/3*3/5 \\ = 0.044 \end{aligned} \tag{11}$$

[1]

Hence the predicted Y is 0.

(d) We are going to decide on which variable to use as the root node based on the information gain that each variable provides.

Gain is computed as $G(I) = H_0 - \frac{k_y}{N}H(V_y) - \frac{k_n}{N}H(V_n)$

Can be demonstrated in the computation/solution below [2]

• Initial entropy H, there are 2+ and 4−:

$$H_0 = -\frac{2}{6}\log_2\frac{2}{6} - \frac{4}{6}\log_2\frac{4}{6} = 0.918 \tag{12}$$

Can be demonstrated in the computation/solution below [2]

• Rooney:
○ V_y, $k_y = 2$: 2+, 0-

$$H_{V_y} = -\frac{2}{2}\log_2\frac{2}{2} - \frac{0}{2}\log_2\frac{0}{2} = 0.0 \tag{13}$$

○ V_n, $k_n = 4$: 2+, 2-

$$H_{V_N} = -\frac{2}{4}\log_2\frac{2}{4} - -\frac{2}{4}\log_2\frac{2}{4} = 1.0 \tag{14}$$

○ Gain:

$$G(I) = 0.918 - \frac{2}{6}0.0 - \frac{4}{6}1.0 = 0.251 \tag{15}$$

• Martial:
○ V_y, $k_y = 2$: 1+, 1-

$$H_{V_y} = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1.0 \tag{16}$$

○ V_n, $k_n = 4$: 1+, 3-

$$H_{V_N} = -\frac{1}{4}\log_2\frac{1}{4} - -\frac{3}{4}\log_2\frac{3}{4} = 0.811 \tag{17}$$

◦ Gain

$$G(I) = 0.918 - \frac{2}{6}1.0 - \frac{4}{6}0.811 = 0.044 \tag{18}$$

• Depay:

◦ V_y, $k_y = 2$: 1+, 1-

$$H_{V_y} = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1.0 \tag{19}$$

◦ V_n, $k_n = 4$: 2+, 2-

$$H_{V_N} = -\frac{2}{4}\log_2\frac{2}{4} - -\frac{2}{4}\log_2\frac{2}{4} = 1.0 \tag{20}$$

◦ Gain

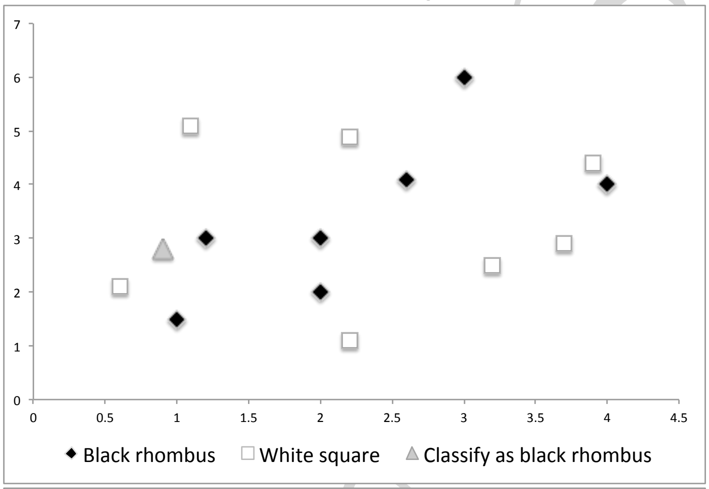
$$G(I) = .918 - \frac{2}{6}1.0 - \frac{4}{6}1.0 = -0.082 \tag{21}$$

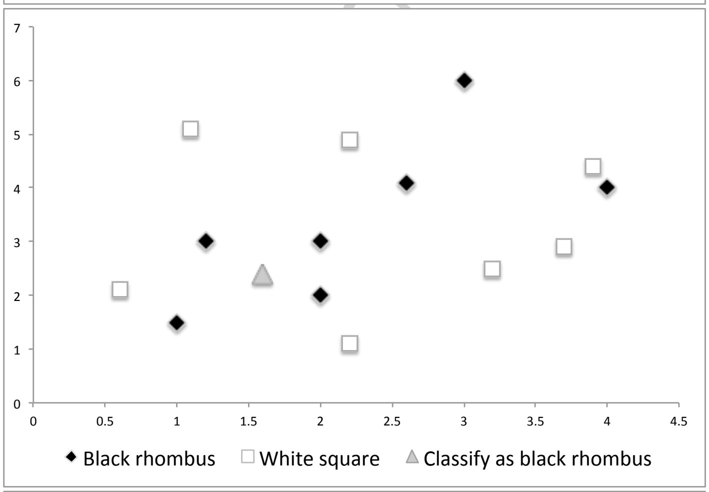
So first node is Rooney, since he has the highest information gain.

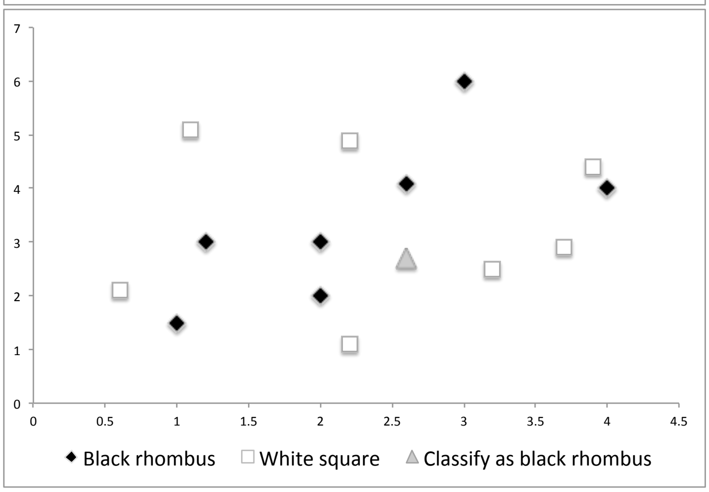
For solution and computation numbers [2]

Question 4

(a) Consider the figures below: find a value of k in a k-nearest neighbourhood classifier, which would classify each of the grey triangles as black rhombuses in all diagrams? Justify your answer. How can k be chosen in general? **[2 marks]**







(b) At start the Candidate Elimination Algorithm has the following version space:

```
G-set:  [ [?, ?, ?]  ]
S-set:  [ [∅, ∅, ∅] ]
```

Assume that the only values possible for each attribute are `a`, `b`, `c` (for each of the three attributes).

Explain and show the new version space which would result from the above version space in the case that each of the following examples is the next example: **[2 marks]**

(i) positive new example `[a, a, a]`.

(ii) positive new example `[b, a, a]`.

(iii) negative new example `[c, c, c]`.

(iv) negative new example `[a, c, c]`.

(c) A large car manufacturing company wants to predict the likelihood of a specific car component being unusable and needing replacement after a period of time (R). Taking the car for regular scheduled services (S) results in an overall good working condition of the car (W). Evidence show that using "cheap" fuel (F) can lead to bad working condition of a car, which in turn is more likely to lead to this specific component becoming unusable and needing replacement.

(i) Draw a Bayes Net corresponding to the information above. **[1 mark]**

(ii) Assume the following:

- 60% of people take their cars for regular scheduled services
- 25% of people put cheap fuel in their car
- There is an 80% chance of the component needing replacement when the car is in bad working condition, but only a 5% chance otherwise.
- If someone takes the car for scheduled services but uses cheap fuel, then there is a 20% chance the car will be in bad working condition, but only a 10% chance if they use premium fuel.
- If someone ignores the regular scheduled services and uses cheap fuel, then there is a 95% chance their car will be in bad working condition, but only a 50% chance if they use premium fuel.

The conditional probability tables for some of this information are given below.

	True	False
P_S	0.60	0.40
P_F	0.25	0.75

P_R	True	False
W=T	0.05	0.95
W=F	0.80	0.20

Give the remaining conditional probability table for the variable W. **[1 mark]**

(iii) Calculate the probability that someone who takes their car for its regular scheduled services and does not use cheap fuel will have this car component failing and needing replacement. **[4 marks]**

(d) Consider the following results from a binary machine learning algorithm:

Test case	Actual class	Predicted class
1	1	1
2	0	0
3	1	1
4	1	0
5	0	1
6	1	1
7	1	1
8	1	1

Show symbolically and compute the following measures:

(i) the numbers of: true positives (TP), false positives (FP), true negatives (TN), false negatives (FN), **[2 marks]**

(ii) the accuracy of the algorithm, **[1 mark]**

(iii) the F1 score of the algorithm. **[2 marks]**

[Question 4 total: 15 marks]

Answer:

Answers to Question 4

(a) k=5. In each case, of the 5 nearest neighbours of the grey triangles, at least 3 are black rhombus. [1]

k can be estimated heuristically, e.g. using cross validation. Choosing odd values is also good if just counting rather than weighting. [1]

(b) (i) +ve `[a, a, a]`:

$$S = [[a, a, a]] \tag{22}$$
$$G = [[?, ?, ?]] \tag{23}$$

(ii) +ve `[b, a, a]`:

$$S = [[?, a, a]] \tag{24}$$
$$G = [[?, ?, ?]] \tag{25}$$

(iii) -ve `[c, c, c]`:

$$S = [[?, a, a]] \tag{26}$$
$$G = [[]] \tag{27}$$

(iv) -ve `[a, c, c]`:

$$S = [\varnothing] \tag{28}$$
$$G = [\varnothing] \tag{29}$$

No solution.

[2]

(c) (i)

```
S   F
 \ /
  W
  |
  R
```

[1]

(ii)

W: Car in good working condition.
S: Keep scheduled services.
F: Cheap fuel.

P_W	S	F	True	False	
	T	T	0.80	0.20	
	T	F	0.90	0.10	[1]
	F	T	0.05	0.95	
	F	F	0.50	0.50	

(iii)

The question asks:

$$P(R = \texttt{t}|S = \texttt{t}, F = \texttt{f}) \tag{30}$$

By the product rule and conditional independence:

$$\begin{aligned} P(R, S, F) &= P(R|S, F)P(S, F) \\ &= P(R|S, F)P(S)P(F) \end{aligned} \tag{31}$$

Using the sum rule to marginalize over W:

$$P(R, S, F) = \sum_{x \in W} P(R, W = x, S, F) \tag{32}$$

From the graph:

$$P(R, W, S, F) = P(R|W)p(W|S, F)P(S)P(F) \tag{33}$$

Using Eq.(2), (3) and (4):

$$P(R|S, F) = \frac{P(R, S, F)}{P(S)P(F)} \tag{34}$$

$$= \frac{\sum_{x \in W} P(R, W = x, S, F)}{P(S)P(F)} \tag{35}$$

$$= \frac{\sum_{x \in W} P(R|W = x)P(W = x|S, F)P(S)P(F)}{P(S)P(F)} \tag{36}$$

$$= \sum_{x \in W} P(R|W = x)P(W = x|S, F) \tag{37}$$

Hence, using Eq.(8) and the table information, Eq.(1) becomes:

$$\begin{aligned} P(R = \texttt{t}|S = \texttt{r}, F = \texttt{f}) &= P(R = \texttt{t}|W = \texttt{t})P(W = \texttt{t}|S = \texttt{t}, F = \texttt{f}) \\ &\quad + P(R = \texttt{t}|W = \texttt{f})P(W = \texttt{f}|S = \texttt{t}, F = \texttt{f}) \\ &= 0.05 \times 0.90 + 0.95 \times 0.10 \\ &= 0.045 + 0.095 \\ &= 0.14 \end{aligned}$$

So the chance of having to repair the component when taking the car for its scheduled services and using premium fuel is 14%.

[4]

(d) (i) TP = 5, FP = 1, TN = 1, FN = 1

(ii) accuracy = (TP+TN)/(P+N) = 0.75

(iii) F1 = (2 x Precision x Recall) / (Precision + Recall) = 0.83

2 marks for first part, 1 mark for Accuracy, 2 marks for F1.

[grand total: 60 marks]