University of St Andrews

DECEMBER 2017 EXAMINATION DIET

SCHOOL OF COMPUTER SCIENCE

MODULE CODE: **CS5010**

MODULE TITLE: Artificial Intelligence Principles

EXAM DURATION: 2 hours

EXAM INSTRUCTIONS:
(a) Answer three questions.
(b) Each question carries 20 marks.
(c) Answer questions in the script book.

PERMITTED MATERIALS: Non-programmable calculator

YOU MUST HAND IN THIS EXAM PAPER AT THE END OF THE EXAM

PLEASE DO NOT TURN OVER THIS EXAM PAPER UNTIL YOU ARE INSTRUCTED TO DO SO.

1. Let S be the training data given in Table 1, in which weather attributes influence decisions on whether or not to play tennis.

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	No
D13	Overcast	Hot	Normal	Weak	Yes
D14	Sunny	Mild	High	Weak	Yes
D15	Rain	Cool	Normal	Weak	No
D16	Rain	Mild	High	Strong	No

Table 1

(a) Given the new observation $[D17, Sunny, Hot, Normal, Weak]$, use Naïve Bayesian classification to decide whether or not tennis was played on Day 17. **[6]**

(b) Suppose that the observations and decisions for days 1 and 2 were removed. Describe the additional step that would be needed to apply Naïve Bayesian classification to new data $[D18, Sunny, Hot, High, Weak]$. **[2]**

(c) Describe the back propagation method for learning optimal neural network weights. **[2]**

(d) Given that the rectifier function is $f(x) = max(0, x)$ (so that, for example, $f(2.5) = 2.5$ and $f(-2.5) = 0$), create a diagram that is labelled with the components for the neural net defined by the following formula

$$\hat{y_1} = \hat{\beta}_0 + \hat{\beta}_1 z_1 + \hat{\beta}_2 z_2 + \hat{\beta}_3 z_3 + \hat{\beta}_4 z_4$$
$$\hat{y_2} = \hat{\beta}_5 + \hat{\beta}_6 z_1 + \hat{\beta}_7 z_2 + \hat{\beta}_8 z_3 + \hat{\beta}_9 z_4$$

where

$$
\begin{aligned}
z_1 &= f(\hat{\alpha}_1 x_1 + \hat{\alpha}_2 x_2 + \hat{\alpha}_3 x_3) \\
z_2 &= f(\hat{\alpha}_4 x_1 + \hat{\alpha}_5 x_2 + \hat{\alpha}_6 x_3) \\
z_3 &= f(\hat{\alpha}_7 x_1 + \hat{\alpha}_8 x_2 + \hat{\alpha}_9 x_3) \\
z_4 &= f(\hat{\alpha}_{10} x_1 + \hat{\alpha}_{11} x_2 + \hat{\alpha}_{12} x_3)
\end{aligned}
$$

[5]

(e) For the neural net defined in part (d), suppose that $\hat{\beta}_2 = -2$, $\hat{\beta}_8 = 3$, $\hat{\alpha}_4 = -1.5$, $\hat{\alpha}_{10} = 4$, and all other $\hat{\alpha}$ and $\hat{\beta}$ values are equal to one. Calculate the response given data $x_1 = -5.2$, $x_2 = 3$ and $x_3 = -3.8$. **[3]**

(f) For the neural net defined in part(d)

(i) show how the output could be restricted to the range $-1, 1$ **[1]**

(ii) show how the output could be restricted to the range $0, 1$ **[1]**

2. (a) Describe the Iterative Deepening search algorithm. Use pseudocode or some other method of giving a succinct but essentially complete description of the algorithm. **[5]**

(b) Consider the Lecturers and Students puzzle. There are two Lecturers and two Students. They are at the Jack Cole Building. We have to find a plan which ends with all four people at Leuchars railway station. The only transport available is a two-seat car. All people involved can drive. If the plan ever leaves more Lecturers than Students at either end of the journey, the Lecturers will bore to death the Student at that location, so this must be avoided. Lecturers do not bore each other to death so it is acceptable to have some Lecturers and no Students at a location. We are to use AI search to find an acceptable plan for this problem.

To represent a search state we only need to know how many Lecturers and Students are at each location, and where the car is. For example, we will write `2/2 JCB 0/0` to represent the start state with 2 Lecturers and Students and the car at Jack Cole Building. The state `1/0 LRS 1/2` would represent a state with 1 lecturer at Jack Cole, with 1 Lecturer and 2 Students and the car at Leuchars Railway Station. Moves involve swapping the position of the car and either one or two people, making sure that the rule about numbers of Lecturers is obeyed.

(i) Draw the *complete* search space, containing a node for each state reachable from the start state up to and including the goal state. Edges between nodes should be labelled with the occupants of the relevant car journey, i.e. L, S, LL, SS, or LS. Since all moves are reversible it is not necessary for edges to be directed. At any node where some possible moves are disallowed because they result in too many lecturers at one place, include a note of these illegal moves. **[7]**

(ii) Using your answer to the previous part (or otherwise), give each of:

- A sequence of legal moves which moves all people to Leuchars, i.e. obtains the state `0/0 LRS 2/2`;
- A sequence of legal moves from the start state which returns to the start state without any move immediately reversing the previous move;
- A state where not everybody is at Leuchars, but which can *only* be reached legally by first transporting everybody to Leuchars.

[3]

(iii) Compare and contrast how the following search methods would perform on this search problem: depth-first search; breadth-first search; iterative deepening search. **[3]**

(iv) This puzzle is a variant of a puzzle which goes back more than 1200 years. Suggest why this puzzle is often hard for people to solve, even though it has a small search space. **[2]**

3. (a) For a Knowledge Base KB and sentence α, what does "$KB \models \alpha$" mean? For a given inference procedure, what does "$KB \vdash \alpha$" mean? What does it mean for the inference procedure to be *sound*, and what does it mean for it to be *complete*? **[4]**

(b) Describe the "resolution" proof rule for propositional logic. Explain why it is *sound*. **[4]**

(c) For each of the following examples, determine if the entailment $KB \models \alpha$ is valid or not in propositional logic. If it is valid, give a proof of this fact using the resolution proof rule. If it is not valid, give an assignment to the propositional variables which makes KB true but α false.

(i) $P \vee Q, P \vee \neg Q \models P$ **[3]**

(ii) $Q \rightarrow \neg R, (P \rightarrow P) \rightarrow R \models R$ **[4]**

(iii) $P \rightarrow (Q \vee R), Q \rightarrow \neg P \models \neg P \vee Q \vee \neg R$ **[5]**

4. Let D be the joint distribution specified by Table 2, in which credit applications to a bank are accepted (or not) for applicants having a salaried job (or not) and are educated to degree level (or not).

	accept		**¬accept**	
	job	¬job	job	¬job
degree	0.027	0.008	0.108	0.012
¬degree	0.144	0.576	0.016	0.064

Table 2

(a) The value 0.027 in the first row and column of D is incorrect. Calculate the correct value (and use this value for subsequent calculations). **[2]**

(b) (i) Calculate $P(accept)$ **[1]**

(ii) Calculate $P(degree \vee \neg accept)$ **[2]**

(iii) Calculate $P(\neg degree | \neg accept)$ **[3]**

(c) State the time and space complexities of the enumeration algorithm for exact inference queries of a Bayesian network. **[2]**

(d) For exact inference queries of a Bayesian network, explain how variable elimination can often be more efficient in time than the enumeration algorithm, but less efficient in space. **[5]**

(e) For approximate inference queries of a Bayesian network, explain how Markov chain Monte Carlo (MCMC) methods can be used to estimate $P(X|\mathbf{e})$ (i.e. the conditional probability of outcome X given evidence $\mathbf{e}$). **[3]**

(f) Are MCMC approximations of inference queries of a Bayesian network aimed at reducing epistemic uncertainty or aleatoric uncertainty, neither, or both? Justify your answer. **[2]**

END OF PAPER
