University of St Andrews

DECEMBER 2016 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.

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. Symbolic Search

(a) Explain the relation of A* to Greedy best-first and Uniform cost symbolic search and how A* works. When and in what way is A* superior to these other two types of search? Describe an exception to this superiority.

[4 marks]

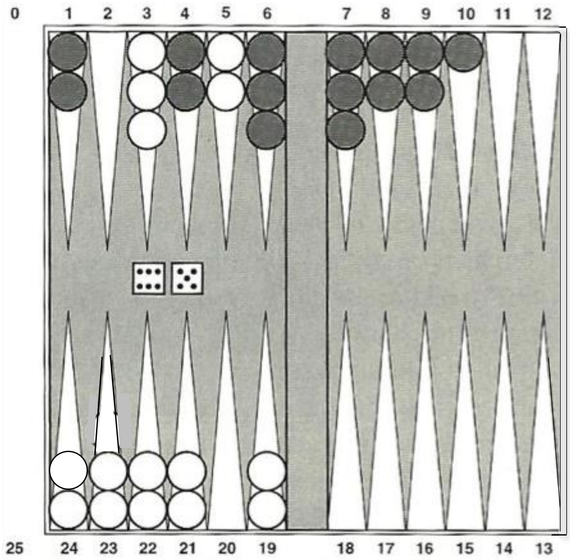
(b) Explain what is meant by the minimax and expectiminimax search procedures as used in AI game playing. Include a description of how values are processed. [4 marks]

(c) Figure 1 shows a possible backgammon position. Draw a part of the expectiminimax tree that represents the dice roll of [5,6] for MAX (White) followed by two alternative dice rolls for MIN (Black) of [6,6] and [1,5] and give the expected leaf utilities associated with these alternative rolls. For MAX, take into account *only* the two possible moves for MAX, MAX1 and MAX2, that hit a black counter. For MIN's roll of [1,5], take into account *only* the two moves MIN1 and MIN2 that: expose a black counter at point 5 (MIN1), or move from point 7 onto point 6 (MIN2).

You should suppose the utility components that are summed to give a board state utility are (from MAX's perspective): Hit an opponent counter => +10; can't get back on the board => -5; each new exposed counter => -3; maximal advance => + 1 per position advanced by a counter advancing maximally (e.g. 2 counters advancing 5 and 6 => +6). You should also assume that utilities are added in sequence from MAX's perspective to yield an accumulated utility for the leaf. For example: MAX1: hits Black and moves 1 counter 11 positions from point 5, 2 new counters exposed, so MAX1utility = 10 + 11 – (2x3) = 15. Then MIN3 throws double 6, can't get back on the board, so MIN3utility = MAX1utility + 5 = 20. [9 marks]

(d) Which move does MAX select using expectiminimax and why? In what ways is expectiminimax in particular limited for Backgammon ? [3 marks]

Black Home



White Home

Figure 1.

2. Learning

(Remember to show your working for parts requiring numerical answers.)

(a) Suppose a hidden layer neuron H in a layered architecture has a weight W1 of 0.45 along a link from input IN1 and an input value from IN1 at time T1 of 2.0. Compute the neuron's next analogue output at time T1+ΔT, assuming the neuron has a logistic sigmoidal activation function and the next output takes ΔT to compute. Give a conventional binary classification of the analogue output. [2 marks]

(b) Add a recurrent connection at time T1+ ΔT from the neuron H to itself with a weight W2 of -2.45. Compute the next output at time T1 + 2 ΔT in analogue and binary terms given that the input is again IN1 and the recurrent connection takes ΔT to transmit its output. If the input remains fixed at IN1, predict the trend of behaviour of the neural output in both analogue and binary terms. [4 marks]

(c) Why is the above behaviour limited from a Finite State Machine perspective and how may the limitation be removed from recurrent networks with a layered architecture? [4 marks]

(d) A point robot positioned at R and with a goal location of G in Figure 2 is currently heading towards H. There is an L-shaped obstacle with a nearest corner at C. Suppose the motion computation involves the following coordinates in (x,y): R = (0, 0); P1 = (0.96, 1.04); P2 = (1.04, 0.96); C = (1.5, 2.5); G = (2.4, 3.1).

Consider the following two potential functions:

An obstacle potential fall-off function f

$$f(d) = \frac{1}{d}.e^{-\frac{1}{a}}$$

where d is the Euclidean distance of a point from C with obstacle potential generated by C alone. a is given by $a = (s - d)$, where s is the fall-off range and set to a value of 2.0, and a goal potential function g

$$g(h) = h^2$$

where h is the Euclidean distance of a point to the goal G.

Use the classical static potential field approach, and the coordinates and potential functions given above, to compute whether the robot will turn left towards P1 or right towards P2 in the next move. Make a numerically based prediction as to what the robot will do after the next move. Give your reasons for your prediction.

(The Figure is not drawn to scale.)

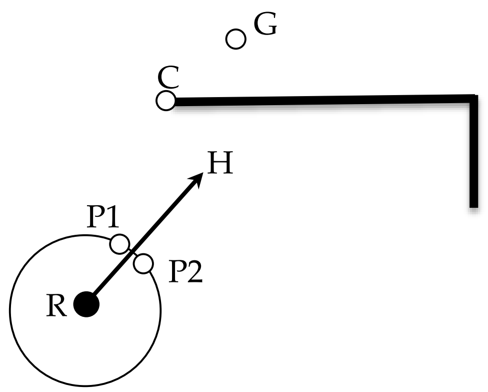


Figure 2.

[10 marks]

3. Logic

(a) What are circuit-based and inference-based propositional agents according to Russell and Norvig? [2 marks]

(b) Suppose a Russell and Norvig inference-based propositional agent has to keep track of its location in (x,y), $L_{x,y}$, and has 4 types of propositions to use in order to do this: Facing Left/Right/Up/Down, always in conjunction with a Moving Forward proposition.

(i) How could moving from $L_{1,1}$ to $L_{1,2}$ in one move while FacingUp be represented by combined propositions?

(ii) How does the use of $L_{x,y}^{t}$ instead of $L_{x,y}$ help prevent system failure where t refers to time at which the agent is at location (x,y)? Correct your initial above representation accordingly.

(iii) Why would a circuit-based representation be a further improvement?

[5 marks]

(c) Consider the following sentence:

$$[(Snow \Rightarrow Cold) \vee (Ice \Rightarrow Cold)] \Rightarrow [(Snow \wedge Ice) \Rightarrow Cold]$$

(i) Convert the left-hand and right-hand sides of the main implication into CNF, showing each step. Give a reason that decides whether the form of the results shows that the sentence is valid, satisfiable (but not valid), or unsatisfiable.

(ii) Prove your answer to (i) using resolution.

(iii) What is the aim of DPLL and what are its theoretical and practical search implications for AI?

(iv) How does DPLL cope with circuit-based inference?

[10 marks]

(d) What is the main additional step to unit propagation in DPLL and how might the sentence investigated above in (c) be extended in a simple way so this extra step is needed? [3 marks]

4. **Uncertainty**

(a) Suppose you wish to compute the likelihood of a hypothesis H given two events $S1$ and $S2$, $P(H|S1, S2)$, and you have no conditional independence information. Which of the three following sets of numbers are sufficient?

Classify each set as either sufficient or insufficient and give your reasons.

(i) $\{P(S1|S2), P(H), P(S1|H), P(S2|H)\}$

(ii) $\{P(S1|S2), P(H), P(S1, S2|H)\}$

(iii) $\{P(H), P(S1|H), P(S2|H)\}$ [3 marks]

(b) What is normalisation in the context of conditional probability? Give an example where normalisation can fruitfully be applied in the above to reduce the computation of $P(H|S1, S2)$. [3 marks]

(c) Suppose new knowledge is made available in the form

$P(S1|H, S2) = P(S1|H)$

for all values of $H, S1, S2$. Is your classification the same or different in each case? Give your reasoning. [4 marks]

(d) Show your reasoning in (a) being borne out numerically using the full joint distribution in Table 1 below and any alternative table entries useful for explanation. [8 marks]

(e) What is the general merit of conditional independence information for AI? [2 marks]

	H	H	NOT H	NOT H
	S1	NOT S1	S1	NOT S1
S2	0.11	0.01	0.07	0.01
NOT S2	0.02	0.06	0.14	0.58

Table 1.

END OF PAPER