COMP9414: Artificial Intelligence
Tutorial 9: Neural Networks/Reinforcement Learning

1. (i) Construct by hand a perceptron which correctly classifies the following data; use your knowledge of plane geometry to choose values for the weights w_0, w_1 and w_2.

Training Example	x_1	x_2	Class
a	0	1	−
b	2	0	−
c	1	1	+

   (ii) Simulate the perceptron learning algorithm on the above data, using a learning rate of 1.0 and initial weight values of $w_0 = -0.5$, $w_1 = 0$ and $w_2 = 1$. In your answer, clearly indicate the new weight values at the end of each training step.

2. Explain how each of the following could be constructed:

   (i) Perceptron to compute the OR function of m inputs

   (ii) Perceptron to compute the AND function of n inputs

   (iii) 2-Layer neural network to compute any (given) logical expression written in CNF

3. Consider a world with two states $S = \{S_1, S_2\}$ and two actions $A = \{a_1, a_2\}$, where the transitions δ and reward r for each state and action are as follows:

$$\begin{array}{ll} \delta(S_1, a_1) = S_1 & r(S_1, a_1) = 0 \\ \delta(S_1, a_2) = S_2 & r(S_1, a_2) = -1 \\ \delta(S_2, a_1) = S_2 & r(S_2, a_1) = +1 \\ \delta(S_2, a_2) = S_1 & r(S_2, a_2) = +5 \end{array}$$

   (i) Draw a picture of this world, using circles for the states and arrows for the transitions.

   (ii) Assuming a discount factor of $\gamma = 0.9$, determine:

      (a) the optimal policy $\pi^* : S \to A$

      (b) the optimal value function $V^* : S \to R$

      (c) the Q function $Q : S \times A \to R$ for the optimal policy

   (iii) Write the Q values in a table.

   (iv) Trace through the first few steps of the Q-learning algorithm on some randomly chosen input, with all Q values initially set to zero. Explain why it is necessary for the agent to explore the environment through probabilistic choice of actions in order to ensure convergence to the true Q values.