Assignment 7
Theoretical Neuroscience

Peter Dayan (dayan@gatsby.ucl.ac.uk)
Teaching Assistants: Laurence Aitchinson, Vincent Adam

Due 10 January, 2013

All figure and equation numbers refer to *Theoretical Neuroscience*.

1. **Perceptron**

   (a) Construct a perceptron (equation 8.46) that classifies 10 binary inputs according to whether their sum $\sum u_a$ is positive or negative. Use a random set of binary inputs during training and compare the performance (both the learning rate and the final accuracy) of the Hebbian (equation 8.47), delta, and perceptron learning rules.

   (b) Repeat this training protocol, but this time attempt to make the output of the perceptron classify according to the parity of the inputs, which is the sign of their product $\prod u_a$.

   (c) Why is this example so much harder than the first case?

2. **Indirect and direct actor**

   (a) Implement a stochastic three-armed bandit using the indirect actor (equation 9.14) and the action choice softmax rule (equation 9.12). Let arm a produce a reward of 1 with probabilty p_a, with $p_1 = 1/4, p_2 = 1/2, p_3 = 3/4$, and use a learning rate of $\epsilon = 0.01, 0.1, 0.5$ and $\beta = 1, 10, 100$. Consider what happens if after every 250 trials, the arms swap their reward probabilities at random.

   (b) Averaging over a long run, explore to see which values of ϵ and β lead to the greatest cumulative reward. Can you account for this behavior?

   (c) For the case of just two of the bandits, implement the direct actor from the lecture notes (this is actually a different direct actor from the one in the textbook) and compare performance with the indirect actor.

3. **Actor critic and Q learning**

   (a) Implement actor critic learning (equations 9.24 and 9.25) in the maze of figure 9.7, with learning rate $\epsilon = 0.5$ for both actor and critic, and $\beta = 1$ for the critic. Starting from zero weights for both the actor and critic, plot learning curves as in figures 9.8 and 9.9.

   (b) Start instead from a policy in which the agent is biased to go left at both B and C, with initial probability 0.99. How does this affect learning at A?

   (c) Repeat a) and b) using Q learning instead. Do you see any difference?

4. **Kalman filter**

   (a) Implement a Kalman filter model for classical conditioning, and apply it to the cases of forwards and backwards blocking. On what parameter does the relative strength of these depend?

   (b) What happens if you use an assumed density filter in which only the diagonal elements of the covariance matrix of uncertainty about the weights are retained?