Artificial Neural Networks (Gerstner). Exercises for week 12

Reinforcement Learning and Biology

Exercise 1. Biological learning rules (in class)

In this exercise you will show that the softmax output for action selection in combination with a linear read-out function leads to a biologically plausible learning rule.

Consider a network with three output neurons corresponding to actions a_1, a_2 and a_3 with 1-hot coding. If $a_k = 1$, action a_k is taken.

The probability of taking action a_k is given by the softmax function

$$\pi(a_i|x) = \frac{\exp[\sum_k w_{ik} y_k]}{\sum_j \exp[\sum_k w_{jk} y_k]} \tag{1}$$

where $y_k = f(x - x_k)$.

a. Show that

$$\frac{d}{dw_{35}} ln[\pi(a_i|x)] = [a_3 - \pi(a_3|x)] y_5 \,. \tag{2}$$

Hint: simply insert the softmax and then take the derivative.

b. Interpret your result in terms of a 'presynaptic factor' and a 'postsynaptic factor'. Can the rule be implemented in biology?

Hint: Consider the two cases: action a_3 is (or is not) chosen at time t.