Lecture 10.1 Exercises

1. Consider the Rescorla/Wagner model of Pavlovian conditioning.
   Assume that $r_t = 1$ on trials when reward is provided and $r_t = 0$ on trials when it is omitted, and that A, B, and C refer to three different stimuli, s_1, s_2, and s_3. For this question, you need not consider the learning rule, only the prediction rule $V_{net} = \sum_i V(s_i)$, where the sum is over all stimuli i present on the trial.
   Suppose I train one set of animals, Group 1, with equal numbers of interleaved trials of two types: $AB \rightarrow R$ and $AC \rightarrow\ 0$. This means s_1 and s_2, i.e. A and B, are presented together in the first case, with reward, and s_1 and s_3 are presented together in the second case, without reward.
   Suppose I train a second set of animals, Group 2, with equal numbers of interleaved trials of two types: $AB \rightarrow 50\%\ R$, $AC \rightarrow 50\%\ R$. By this I mean that the same stimulus compounds are presented, but they are each rewarded on half the trials in which they are presented, so there are actually four trial types, $AB \rightarrow R$, $AB \rightarrow 0$, etc. all presented interleaved in equal numbers.
   Suppose in each case I conduct this training to asymptote.

   a. Now I test the animals for their conditioned responding to A alone. (In the model, this is probing the value $V(A)$. How would you expect that responding to A should compare between the two groups?

      Explain, in English, why it would make sense for these groups to make the pattern of predictions you describe. Here, we are only considering the logic of the experiment – imagine a “pigeon detective” trying to figure out which stimuli predict rewards given the training described above. (Hint: Note that in both groups, in training, A was reinforced on half the trials it was presented, but in the context of compounds where the *other* stimuli B and C have different predictive relationships with reward.)

   b. For each group of animals, write down a set of values $V(A)$, $V(B)$, and $V(C)$ such that the animal can correctly predict the probability of reward in each case. (e.g., $V_{net} = V(A) + V(B) = 1$ on AB trials in the first group; $V_{net} = V(A) + V(C) = 0.5$ on AB trials in the second group, and so on for AC. There are many sets of weights that meet these criteria, but you should exhibit one which also complies with your prediction from part **a** about responding to A alone.

2. Now consider the second-order conditioning task. In particular, we first train $A \rightarrow R$ until asymptote, $V(A) = 1$. Now we train a single trial of $B \rightarrow A$, i.e. we precede A by a second stimulus. Since the Rescorla-Wagner model does not consider the order of events in a trial, this is the same as $AB \rightarrow 0$, the compound AB presented with reward $r = 0$.
   Prove, by reference to the Rescorla-Wagner learning rule, that V(B) will be negative for any nonzero learning rate α, inconsistent with the second-order conditioning effect.

3. Suppose an animal is repeatedly exposed to some stimulus A accompanied by reward $r = 1$ with probability p and nonreward ($r = 0$) with probability $1 - p$.
   Given infinite training with these events and any nonzero learning rate α, what is the asymptotic expected value for the reward prediction $V(A)$? Use the Rescorla-Wagner prediction error δ to show that this is the fixed point of the update rule, in the sense that it is the value for which the average prediction error, in expectation over the two trial types, is zero.