Problem Set #2: Quantitative Models of Behavior

Hsieh Yu-Guan

March 18, 2017

Introduction

In this report we will work on several classic behaviors models. The topics that we will cover include conditional learning (classical conditioning and operant conditioning), decision making and reinforcement learning.

1 Rescola-Wagner model

1.1 Model description

Let's start with the Rescola-Wagner model. It's a model of classical conditioning in which learning signifies building associations between conditioned (CS) and unconditioned (UCS) stimuli. UCS are often represented in form of rewards like food (for an animal) or money (for a person) while CS are some kinds of neutral stimuli that may or may not allow us to predict the occurrence of this reward.

In a single trial, the presence or absence of the reward is respectively denoted by $r = 0$ or $r = 1$. More than one CS can be taken into account, then the presence of the i^{th} stimulus is denoted by $u^{(i)} = 1$ and its absence by $u^{(i)} = 0$. The animal's prediction v is given by the formula

$$v = \sum_{i=1}^{m} w^{(i)} u^{(i)}$$

where m is the number of CS, and for each i, $w^{(i)}$ is the prediction parameter associated with $u^{(i)}$. If we note $u = (u_1, ..., u_m)$ and $w = (w_1, ..., w_m)$, it can also be written in the form

$$v = w \cdot u$$

where $\cdot$ denotes the scalar product. We can now calculate the predition error $\delta = r - v$ which, with the learning rate ϵ, allows us to write down the update rule for every $w^{(i)}$ after the trial at time t

$$w_{t+1}^{(i)} = w_t^{(i)} + \epsilon \delta_t u_t^{(i)}.$$

1.2 A simple test

To examine the reliability of the model, we'll first look at a very simple experiment. Only one stimulus is considered, so we identify for the moment u with $u^{(1)}$ and w with $w^{(1)}$. Let's assume that in the first 25 trials, both stimulus and reward are present, and during the next 25 trials, only the stimulus is present. The plot of r and u is shown in the figure below.

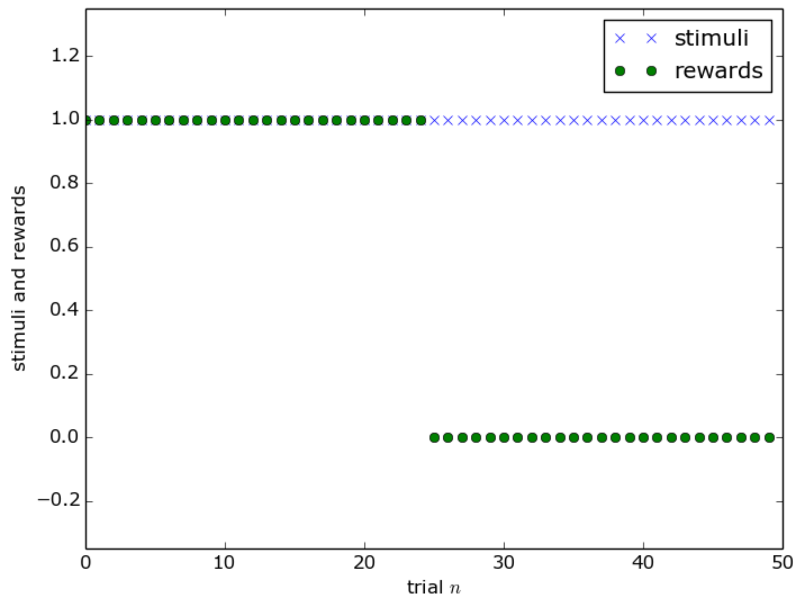

FIGURE 1: A very simple experiment

Now if we do the simulation with the learning rate $\epsilon = 1$, the value of w evolves as shown (since u equals always 1 here, we have also all the time $v = w$).

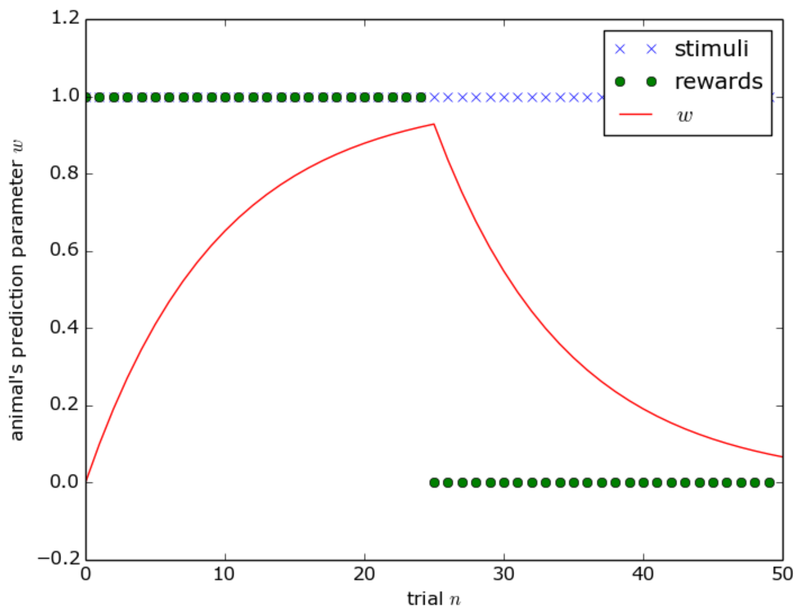

FIGURE 2: Evolution of w in this simple experiment

We see an exponential rise of w to value 1 during the first 25 trials and its value decays exponentially to 0 for the rest of experiment. It's rather reasonable intuitively, and from a mathematical point of view, we write simply $w_{t+1} - w_t = \epsilon\delta_t u_t$. We know that $\delta_t = r_t - w_t$ (remember that $v_t = w_t$) and $u_t = 1$ for all t. For the first 25 trials, we get $w_{t+1} - w_t = \epsilon(1 - w_t)$, so if we put it in a continuous form, it becomes

$$\frac{dw}{dt} = \epsilon(1 - w).$$

We now just need to solve this differential equation to see that $w_t = 1 - e^{-\epsilon t}$. In the same way, for the trials 26 to 50, we can get $w_t = Ae^{-\epsilon(t-25)}$ where A is a constant that can be decided given a particular value of w_t.

The next thing to do is surely to study the impact of the learning rate, so we vary the value of ϵ.

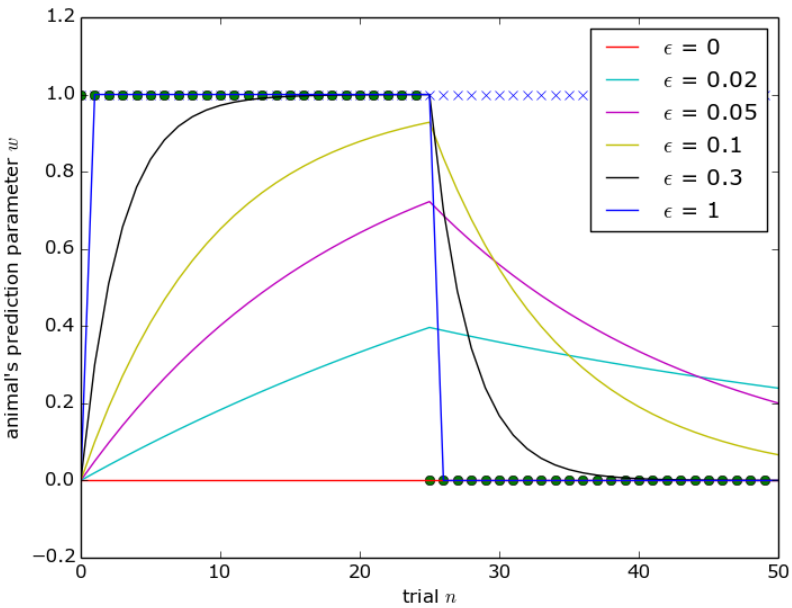

FIGURE 3: Evolution of w in the same experiment but with different ϵ

As what our formula implies, greater the learning rate, faster the animal learns and unlearns the association. It's also interesting to notice that in most of the cases, it's more difficult to unlearn than to learn because the initial δ is smaller. However, even though one learns faster with a greater learning rate, it doesn't necessarily means that it's better. In fact, there are always noises in what is observed and a smaller learning rate indicates that one is somehow doing an average with past experiences.

Two extreme cases are shown in the figure, when $\epsilon = 0$, the animal can never learn, and when $\epsilon = 1$, the animal exploit only the information from the current trial to compute w and is hence very sensitive to noises. In cosequence, the best value of ϵ depends on the environment where one is.