CCN Assignment 2: Models of Decision-Making and Computational Psychiatry

Peggy Series, pseries@inf.ed.ac.uk

February 22, 2016

1 Practicalities

Deadline is March 14th 2016 at 4 pm (standard late policies apply). Please submit the pdf of your report to ITO using the command submit. Please name your document using yourname-assign2-ccn16.pdf.

Please also submit the paper copy to ITO by the deadline or just after (it will be the time of the submit command that will matter).

Report your findings. Particularly well-researched answers can receive additional points. Ideally you substantiate your explanations, for instance by additional simulations. Plots should always include axes labels and units. Figures should always have a caption. The presentation and format will count in the final mark. The report should look like a scientific report - no need to include any code.

Copying results is not allowed. It's OK to ask for help from your friends. However, this help must not extend to copying code or written text that your friend has written, or that you and your friend have written together. I assess you on the basis of what you are able to do by yourself. It's OK to help a friend. However, this help must not extend to providing your friend with code or written text. If you are found to have done so, a penalty will be assessed against you as well.

Email me (pseries@inf.ed.ac.uk) the Matlab script that you used. I will not assess the programming style, but I might check them if results are unexpected. I will also run plagiarism detectors on them.
2 Problem 1: Decision as a Diffusion Process (50 points)

2.1 Introduction

Psychologists have often modelled the dynamics of decision processes using a random walk model or a diffusion process (see lectures slides, and [6] for a review). This model starts from the basic principle of accumulation of information. When an individual is asked to make a binary choice on the basis of an available stimulus, the assumption is that evidence from the stimulus is accumulated over time and a decision is made as soon as an upper or lower boundary is reached. Which boundary is reached determines which response is given and the time required to reach it determines the reaction time (RT). In this section, we analyze the properties of such a model.

2.2 Model

We here model a decision process between two hypotheses which we call $h_-$ and $h_+$. A Wiener diffusion process $W_t$ with drift rate $v$ and variance $s^2$ can be simulated using the following equation [3]:

$$W(t + dt) = W(t) + v \cdot dt + s \cdot \eta$$

When $W(t) < 0$, a decision in favor of $h_-$ is made. On the contrary if $W(t) > a$, a decision in favor of $h_+$ is made. $dt$ denotes the time step used to simulate the process. $\eta$ represents a Gaussian noise process with standard deviation $\sqrt{dt}$: $\eta \sim N(0,\sqrt{dt})$.

The parameters are: the mean drift rate $v$, the separation of the boundaries $a$, the starting point $W(0) = z$. The drift rate $v$ models the amount of evidence in favor of each of the hypotheses. If $v > 0$, there is more evidence in favor of $h_+$ (which is then the correct decision). On the contrary, if $v < 0$, there is more evidence in favor of $h_-$. The amplitude of the drift rate is related to the difficulty of the task: easy tasks are modelled using large $v$ while difficult tasks are modelled using relatively small $v$. For example, in the context of the motion discrimination task with random dots, the amplitude of $v$ would be related to the coherence level. Positive $v$ would signal for e.g. available information (or evidence) in favor of 'up' while negative $v$ would signal evidence in favor of 'down'.

1. We assume that $v = 0.2$, $a = 0.1$, $z = a/2$, $s = 0.1$. You can use $dt = 0.001$ (or smaller). Plot some examples of the paths taken by the diffusion process on a few trials.

2. Try different values of $v$ to explore the expected performance and RTs distributions (of correct and incorrect responses) on easier or more difficult tasks.
3. On the same task, it is possible to ask subjects to respond as fast as possible, or as accurately as possible. Speed-accuracy tradeoffs are modeled by changing the boundary positions. When accuracy is emphasized, the boundaries are set far from the starting point. On the contrary, when speed is emphasized, boundaries are set close to the starting point. It has also been proposed that poor decision-making in certain disorders such as schizophrenia could be explained by the distance between the decision boundaries being smaller or more variable than in healthy subjects [2, 1]. Show how the distance between the boundaries affects the performance and RTs distributions.

4. In a different version of the model (known as the Ratcliff diffusion process), variability is introduced in the drift rate $v$ and the starting point $z$. What difference does this make?

5. Suppose that the subject knows that one hypothesis is more likely than the other, for example the prior $p(h_+) = 2p(h_-)$. How can you include such information in the model? Can this be justified theoretically (inspired from what you know of the log-likelihood ratio test, for e.g.)?

6. Now assume the task to be performed is the random-dot motion discrimination task where the subject has to decide whether the noisy motion is going 'up' or 'down'. We are interested in understanding how such a model could be implemented neurally. Assume that the evidence that is available comes from the firing of two MT neurons (or pools of neurons), one preferring 'up', and one preferring 'down'. Can you propose a biologically plausible circuit, that would be able to implement the diffusion process (or equivalent)?

3 Problem 2: TD learning and Psychosis (50 points + 10 bonus points)

This problem is inspired from [5].

Dopamine plays a central role in the pathophysiology of schizophrenia, particularly with respect to the manifestation of psychosis (i.e. delusions, hallucinations). There are multiple sources of evidence for this. First, clinically effective antipsychotic drugs (APDs) such as haloperidol act by blocking dopamine receptors. Second, dopamine enhancing agents such as amphetamines can induce psychotic symptoms in otherwise normal people. There are also a large number of studies using animal models that have served to clarify the impact of dopamine manipulation on behaviour, in particular Conditioned Avoidance, and Latent Inhibition (see below). Conditioned avoidance (CA) is a classic animal model in the study of antipsychotic drugs and their dopamine-blocking properties and one that has been used extensively as a preclinical screen for
antipsychotic efficacy. Meanwhile, latent inhibition is widely used in the study of selective attention in the context of reward learning and is disrupted not only in animals and people following induced-hyperdopaminergic states but also in patients with schizophrenia.

The purpose of the present problem is to link animal models of psychosis to computational models of dopamine function.

3.1 TD learning and Conditioned Avoidance (CA)

We first aim at modeling Conditioned Avoidance in normal conditions.

In a typical CA experiment, a rat is placed in a 2 compartment box and presented with a neural conditioned stimulus (CS) such as a light or tone for 10 seconds, immediately followed by an aversive unconditioned stimulus (US), such as a footshock. The animal may escape the US when it arrives by running from one compartment to the other. However, after several presentation of the CS-US pair, the animal typically runs during the CS and before the onset of the US, thereby avoiding the US altogether – this behaviour is called conditional avoidance.

Following [5], we model the CA experiment by assuming it can be decomposed in a series of states: a state representing the CS, a number of internal states ($I_1, I_2, I_3, I_4$) and a state representing the US. Each state is associated with a value $V(S)$, which represents the agent's internal estimate of the total future reward expected to follow that state. Initially, all $V(S)$ are 0 but will be adapted with experience. Each state is also associated with a reward $r(S)$ $r(S) = 0$ for all states, except for the $r(US) = 1$.

- Using the equations of TD learning (cf [5], equations 1 and 2), simulate the evolution of $V(S)$ and $\delta(s)$ of all the states with training. Illustrate.
  The discount factor is set to $\gamma = 0.93$. The learning rate is $\alpha = 0.5$.

We assume that $V(t)$ can be interpreted as motivation to produce an avoidance response in the current state. The probability of observing an avoidance response at time $t$ is $p(t) = V(t)$. As the first avoidance response of a trial will end that trial, the overall probability of observing an avoidance response at time $t$ is: $\bar{p}(t) = p(t)(1 - \sum_{i=1}^{t-1} \bar{p}(i))$ while the overall probability of producing an avoidance response $P(avoidance)$ before the US is presented is $\bar{p}(CS) + \bar{p}(I_1) + \bar{p}(I_2) + \bar{p}(I_3) + \bar{p}(I_4)$.

- Illustrate the evolution of $P(avoidance)$ with training.

3.2 CA under pharmacological manipulation

It is well established that dopamine antagonists, such as antipsychotic drugs (APDs) disrupt the acquisition of conditioned avoidance. The simplest way
to simulate dopamine receptor manipulation is to add some constant \( \theta \) to all prediction errors: \( \delta_{rec}(t) = \delta(t) + \theta \). Blocking of dopamine receptors is modelled using \( \theta < 0 \). This simulates the fact that dopamine has less of an impact at the receptor site. Low, medium and high doses of APD are simulated by setting \( \theta = -0.2; \theta = -0.3 \) and \( \theta = -0.4 \) respectively.

- Simulate training with the different doses of APD and plot the probability of avoidance response\(^1\). What do you observe?
- Can you reproduce Figure 2 of [5]? Discuss.

### 3.3 Latent Inhibition (LI)

Latent inhibition refers to a subject’s increased difficulty to form a new association between a stimulus and a reward due to prior exposure of that stimulus without consequence. LI is disrupted in animals treated with amphetamines (which enhance dopamine). This deficit can be reversed if the animal is given APDs.

To account for LI, [5] offer a simple modification of the classic TD model. Pre-exposure is modelled by adding an associativity parameter \( \phi \) which controls the rate of conditioning. Equation 2 is replaced by \( V(t) = V(t) + \alpha \delta_{rec}(t) \). The idea is that a non-pre-exposed CS has \( \phi = 1 \) whereas a pre-exposed CS has \( \phi < 1 \) which is constant following pre-exposure.

- Using \( \phi = 0.4 \) for the pre-exposed group, show how \( P(\text{avoidance}) \) evolves with training in the pre-exposed group compared to the exposed group.
- Repeat this analysis when the rats have been given Amphetamines (\( \theta = 0.3 \)). Can you reproduce Figure 4 of [5]? Comment.
- Repeat this analysis when the rats have been given the APD Haloperidol (\( \theta = -0.2 \)). Can you reproduce Figure 5 of [5]? Comment.

### 3.4 From Schizophrenia to Cocaine Addiction

This section is inspired by [4]. Imagine now that we’re training some new rats on another CS-US association. Here, the US is a dose of cocaine.

[4] has suggested that addiction could be modelled using a model which is very similar to the model above. The idea is that administration of cocaine has a direct effect on dopaminergic neurons, which persists even after training and

\(^1\)Note: you will encounter negative values if you implement the expression for the values as described in the paper. This should not happen.

The expression should have been: \( V(t) = [V(t) + \alpha \delta(t)]_+ \) where \([ ]_+ \) denotes rectification (that is we take the max between the value and 0) so that the values are never negative.
interferes with the signaling of the prediction error. More precisely, administration of the drug induces an increase $D(US)$ in the prediction error signal that cannot be compensated by changes in the value during training.

Equation 1 is replaced by:

$$\delta = \max \{ r(t) + \gamma V(t + 1) - V(t) + D(t); D(t) \}$$

As before $r(S) = 0$ for all states except for $r(US) = 1$; Similarly $D(S) = 0$ except for $D(US) = 0.025$.

- Simulate the evolution of $V(S)$ with training under this new model. Illustrate.
- Can you see the link with addiction? Discuss.

References


