Version GH/3 – CRIB

EGT2
ENGINEERING TRIPOS PART IIA

---

Monday 4 May 2026 14:00 to 15:40

---

**Module 3G3 – CRIB**

**INTRODUCTION TO NEUROSCIENCE**

*Answer not more than **three** questions.*

*All questions carry the same number of marks.*

*The **approximate** percentage of marks allocated to each part of a question is indicated in the right margin.*

*Write your candidate number **not** your name on the cover sheet.*

**STATIONERY REQUIREMENTS**
Single-sided script paper

**SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM**
CUED approved calculator allowed
Engineering Data Book
Supplementary page: one extra copy of Fig. 1 (Question 1)

## 10 minutes reading time is allowed for this paper at the start of the exam.

## You may not start to read the questions printed on the subsequent pages of this question paper until instructed to do so.

## You may not remove any stationery from the Examination Room.

1 (a) Write short notes on the following:

(i) Why perception is fundamentally different from sensation. Support your answer with one example; [15%]

Answer: Sensation is merely the relaying of information about the environment through sensors (e.g. photoreceptors). Perception, on the contrary, is an active process of constructing interpretations of those sensory inputs. In essence, it is an inference problem, whereby the brain infers the state of the world based on noisy and ambiguous observations. For example, two people of the same size standing at the two far corners of the "Ames room" are strongly perceived as being of very different sizes. This shows that our brain makes up percepts that are not real, and indeed percepts that may even conflict with information gathered from stereoscopic vision in this case (any of the other examples discussed in lectures, or indeed any other suitable example, are acceptable).

(ii) Why optimal decision-making often requires taking into account one's uncertainty. In particular, what kinds of uncertainty must be considered? Support your answer with one example. [15%]

Answer: Imagine the following game. A gold coin is hidden in one of my two closed hands. You have two choices: i) pick one of my hands – if it holds the coin, the coin is yours, but if it's empty, you must give me 3 gold coins instead – or ii) opt out of the game and nothing happens. Which would you choose? You might get some visual information as to which hand holds the coin by comparing their sizes. If this clue is weak, your uncertainty about the winning hand is so great that you have near-equal probability of winning and losing, i.e. an expected negative return of -2 coins if you choose to play. You should therefore opt out. If, on the other hand, you estimate that the gold coin is 90% more likely to be in my left hand, then your return is $0.9 - 3 \times 0.1 = 0.6$ gold coins and you might want to play. More generally, optimal decision making requires incorporating uncertainty about both the state of the environment and the way motor actions affect it.

(b) This question is about the drift-diffusion model studied in lectures, whereby a decision variable $x$ evolves according to the following stochastic differential equation:

$$\mathrm{d}x = \mu \mathrm{d}t + \sigma \mathrm{d}W$$

(i) State the meaning of $\mu$, $\sigma$ and $\mathrm{d}W$ in this equation.

Answer: $\mu$ is the so-called "drift" term, describing the deterministic aspect of the evolution of the decision variable $x$. $\sigma dW$ together form the diffusion term – $dW$ is a so-called unit-variance Wiener process (white noise of variance equal to the time increment $dt$). $\sigma$ is the standard deviation of that Wiener process – it controls the amount of stochasticity in the temporal evolution of $x$. [10%]

(ii) How is this equation typically used to model animal behaviour in the classic

Shadlen-Newsome visual motion discrimination task discussed in lectures? Are there any other parameters not explicitly present in the equation?

Answer: This equation is used to model decision-making behaviour as follows: the model is said to go for choice A (say, "motion was to the left") when $x$ crosses some upper bound $+L$ from below. Likewise, the model goes for choice B (opposite judgement) when $x$ crosses $-L$ from above. (Here, parameter $L$ is redundant with the overall scale of $\mu$ and $\sigma$, so can be set to $L = 1$ without loss of generality). The initial condition $x(t = 0) = D$ is a free parameter, not shown in the equation, which models the subject's bias towards choice A ($D > 0$) or choice B ($D < 0$). The drift term $\mu$ is typically set to scale linearly with the (signed) motion coherence in the stimulus. By running a lot of simulations for a given $c$ but different, independent realisations of the Wiener process, one can extract psychometric statistics such as "fraction of A choices" and "mean reaction time". [20%]

(iii) For a 2AFC task involving accumulation of sensory evidence, Fig. 1 shows the probability of choosing one of the two alternatives ('A'; left panel) and the average reaction time (right panel) as a function of some task parameter $c$ describing stimulus difficulty. Three different parameter sets (P, Q, and R) were used, corresponding (in arbitrary order) to the three different line types (solid, dashed, dotted) in Fig. 1. These parameter sets are:

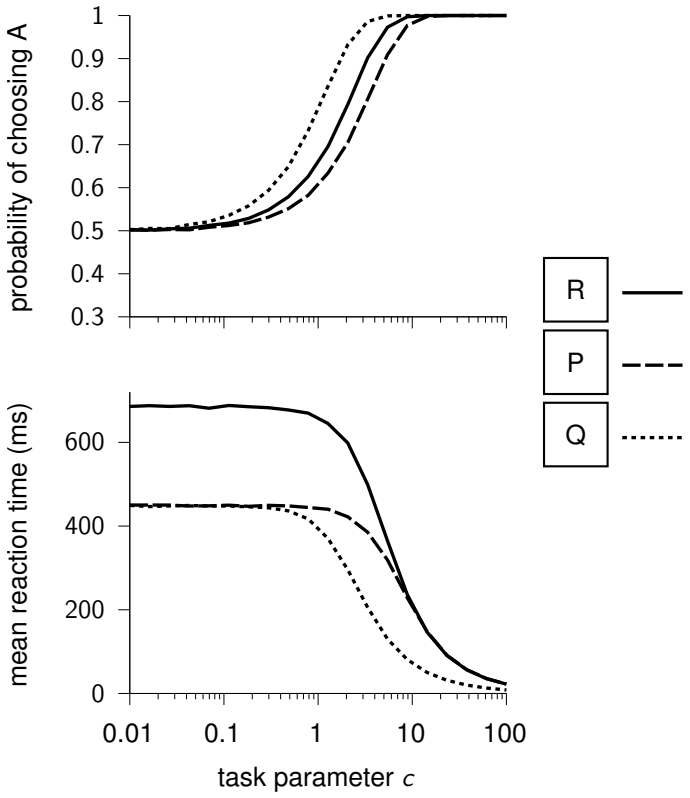
$$\text{set P}: \quad \mu = 0.2c \quad \sigma = 1.0 \tag{1}$$

$$\text{set Q}: \quad \mu = 0.6c \quad \sigma = 1.0 \tag{2}$$

$$\text{set R}: \quad \mu = 0.2c \quad \sigma = 0.8 \tag{3}$$

Note that one parameter set is deliberately missing from each graph. On the *additional copy of Fig. 1* provided at the end of this paper, complete the legend by entering the correct set ('P', 'Q', or 'R') next to each line style. In addition, draw a *qualitative* sketch of the missing curve in each panel. Provide a justification for your choices. Remember to hand in your completed copy of Fig. 1 with your answer to this question. [40%]

Answer: Solid: R / Dashed: P / Dotted: Q. Rationale: For $c = 0$, lower noise (smaller $\sigma$) should result in diffusion taking longer to reach one of the two bounds (randomly). So clearly the bottom panel is missing a curve for set $R$ – thus, Solid = R. For intermediate coherence levels that do not saturate the probabilities of choosing A or B, set $P$ (higher noise) should lead to a lower proportion of A choices than set $R$. The corresponding curve for $P$ in the top panel should therefore offset to the right relative to the solid line. This curve is missing, so Dashed = P. By elimination, Dotted = Q.


probability of choosing A
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
mean reaction time (ms)
600
400
200
0
0.01
0.1
1
10
100
task parameter c
R
P
Q



(cont.

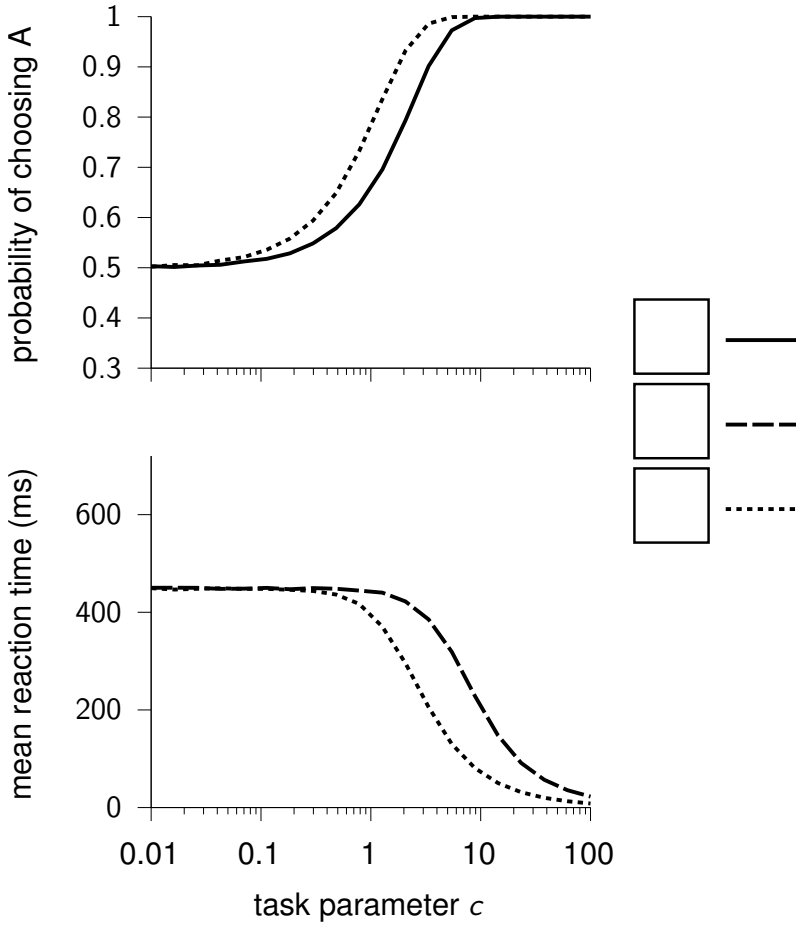

probability of choosing A
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
mean reaction time (ms)
600
400
200
0
0.01
0.1
1
10
100
task parameter c

Fig. 1

2 (a) Write short notes on the following:

(i) David Marr's three levels of understanding; [15%]

Answer: Marr proposed to study vision on three different levels:

- **Computational**. What is vision about? How can we formalise the problem that our visual systems is solving? What is the mathematical nature of the problem at hand?
- **Algorithmic/Representational**. What is the algorithm (e.g., recipe) that our brain uses for vision? What data representations does it use?
- **Implementation**. How is the algorithm implemented in biological/neural hardware (e.g. dynamics of brain networks)?

These levels of analysis are applicable more broadly to neuroscience research.

(ii) The action potential – in particular, what triggers it, and what terminates it. [15%]

Answer: When the membrane potential goes above a certain threshold (typically around −50 mV, sodium channels activate rapidly (more rapidly than potassium channels) and this drives the membrane potential even higher towards the Nernst potential of sodium. A millisecond or so later, the now elevated $V_{\mathrm{m}}$ causes sodium channels to deactivate, which allows the outward potassium curret (which has also grown with $V_{\mathrm{m}}$, and whose inactivation will follow more slowly) to dominate and to drive the membrane potential back to rest.

(b) This part is about Bayesian inference in sensory perception. A subject is shown a brief flash of light at a random location $x$ along a horizontal line. The location $x$ is drawn from a Gaussian prior $p(x) = \mathcal{N}(x; 0, \sigma_p^2)$. The subject's sensory likelihood function is $p(s|x) \propto \mathcal{N}(x; s, \sigma_s^2)$.

(i) Give two biophysical reasons why $\sigma_s$ cannot be zero.

Answer: Light arriving at the retina is subject to shot noise (Poisson noise), such that how many photons will even get a chance of elicing a photoreceptor response is a random variable from trial to trial. Ion channels in sensory neurons are subject to thermal noise such that there is stochasticity in their opening and closing; this creates variability in neuronal spiking. Likewise, synaptic vesicle release is also stochastic, introducing communication noise in neural networks at the sensory periphery. [10%]

(ii) Derive the posterior distribution $p(x|s)$. Explain how the subject should estimate $x$ to minimize the expected squared error.

Answer: $p(x|s) \propto p(x)p(s|x)$. Both $p(x)$ and $p(s|x)$ are bell-shaped functions of $x$, so too is there

product. Indeed, working with logarithms for convenience,

$$\log p(x|s) = \text{cst} - \frac{1}{2}\left(\frac{x^2}{\sigma_p^2} + \frac{(x-s)^2}{\sigma_s^2}\right) \tag{4}$$

$$= \text{cst} - \frac{1}{2}\left[\left(\frac{1}{\sigma_p^2} + \frac{1}{\sigma_s^2}\right)x^2 - \frac{2sx}{\sigma_s^2} + \text{cst'}\right] \tag{5}$$

$$= \text{cst} - \frac{1}{2}\left[\frac{(x-\mu)^2}{\sigma^2} + \text{cst''}\right] \qquad \text{(completing the square)} \tag{6}$$

and by identification of quadratic and linear terms, we get

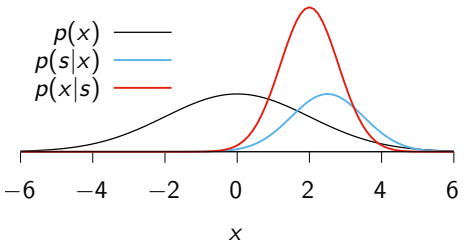
$$\sigma^2 = \frac{\sigma_p^2\sigma_s^2}{\sigma_p^2 + \sigma_s^2} \tag{7}$$

$$\text{and} \quad \mu = s\frac{\sigma^2}{\sigma_s^2} = s\frac{\sigma_p^2}{\sigma_p^2 + \sigma_s^2} \tag{8}$$

In order to minimise the expected squared error, the subject should use $\mu$ (the posterior mean, more generally) as their estimate of the stimulus location. [20%]

(iii) Assume that the subject's perception of the stimulus location conforms to Bayesian principles. Sketch the likelihood, prior distribution and posterior distribution on the same axes for $\sigma_p = 2$ cm and $\sigma_s = 1$ cm, for a trial for which the subject's perceived location turns out to be $\mu = 2$ cm.

Answer:




According to our previous calculations, with $\sigma_p = 2$ we have $\sigma^2 = 4/5$, and a posterior mean of $\mu = 2$ cm means a maximum likelihood estimate of $s = \mu \times (4+1)/4 = 5/2$ (hence where the blue curve above has its peak). (Note: the likelihood has no reason to have unit area.) [10%]

(iv) What would the subject's estimate of $x$ have been, in the same trial as in (iii), if the prior variance $\sigma_p^2$ had been halved? Justify your answer.

Answer: Had $\sigma_p$ been equal to 1 instead of 2, $\sigma^2$ would have been equal to $1/2$, so for the same ML estimate of $5/2$ ("same trial"), the posterior mean would have been $5/2 \times 1/(1+1) = 5/4 = 1.25$ cm instead. [10%]

(v) A naive subject performs 100 consecutive trials of this task. The stimulus position $x$ is drawn independently from $p(x)$ in each trial, and the subject is asked to report their estimate of it. It is found that the subject's mean squared error is lower in the second half of the trials, compared to the first half. How would the Bayesian perception hypothesis explain this?

Answer: The subject needs time to learn the correct prior $p(x)$ from repeated exposure to the task. Since the posterior mean is the MSE-optimal estimate, one would expect that a subject who learns the prior distribution becomes able to form an accurate posterior; such a subject would then exhibit a lower MSE in the second half of the trials compared to the first half. [20%]

3 (a) This part is about synaptic transmission. For each of the following manipulations, explain how pre- and postsynaptic signalling, with special regard to the PSP, changes in a glutamatergic synapse. All manipulations below keep the system in the physiological range to avoid depolarisation block or damage to cells.

(i) Increase in intracellular $Na^+$ concentration in the presynaptic cell. [10%]

Answer: The Nernst potential of $Na^+$ decreases (which means it becomes less positive) in the presynaptic cell (but remains unchanged in the postsynaptic cell). This decreases the resting membrane potential (thus pushing it farther away from the threshold for action potential initiation), thus delaying presynaptic action potential initiation, and decelerates the upstroke and decreases the amplitude of the presynaptic action potential, which together lead to delayed, decelerated and decreased vesicle release, and in turn to delayed, decelerated and decreased glutamate receptor activation (and thus depolarisation) in the postsynaptic cell. These changes together contribute to a somewhat delayed, slower and diminished PSP.

(ii) Increase in intracellular $Na^+$ concentration in the postsynaptic cell. [10%]

Answer: The Nernst potential of $Na^+$ decreases (which means it becomes less positive) in the postsynaptic cell (but remains unchanged in the presynaptic cell). This decreases the resting membrane potential (but only by a small amount) and the peak membrane potential value of the PSP (by a relatively larger amount – because the total $Na^+$ conductance is much larger near the peak than at rest) even though glutamate receptor activation is unchanged. These changes together contribute to a diminished PSP (but no delay or slowing).

(iii) Increase in extracellular $Na^+$ concentration. [10%]

Answer: The Nernst potential of $Na^+$ increases (which means it becomes more positive) in *both* the pre- and postsynaptic cells. This increases the resting membrane potential (thus pushing it closer to the threshold for action potential initiation), thus advancing presynaptic action potential initiation, and accelerates the upstroke (once the membrane potential exceeds the threshold) and increases the amplitude of the action potential in the presynaptic cell, which together lead to advanced, accelerated and increased vesicle release, and in turn to advanced, accelerated and increased glutamate receptor activation in the postsynaptic cell. At the same time, the increase in $Na^+$ Nernst potential in the postsynaptic cell increases the resting membrane potential (but only by a small amount) and the peak membrane potential value of the PSP (by a relatively larger amount – because the total $Na^+$ conductance is much larger near the peak than at rest) even if glutamate receptor activation is unchanged. These pre- and postsynaptic changes together contribute to an advanced, faster rising and larger PSP.

(iv) Increase in intracellular $K^+$ concentration in the presynaptic cell. [10%]

Answer: The Nernst potential of $K^+$ decreases (which means it becomes more negative) in the presynaptic cell (but remains unchanged in the postsynaptic cell). This decreases the resting membrane potential (thus pushing it farther away from the threshold for action potential initiation), thus delaying presynaptic action potential initiation, decelerates the upstroke, accelerates the downstroke, and decreases the amplitude of the presynaptic action potential, which together lead to delayed, decelerated, shorter lasting, and decreased vesicle release, and in turn to delayed, decelerated, shorter lasting, and decreased glutamate receptor activation (and thus depolarisation) in the postsynaptic cell. These changes together contribute to a somewhat delayed, slower rising, but shorter lasting, and diminished PSP.

(v) Increase in intracellular $K^+$ concentration in the postsynaptic cell. [10%]

Answer: The Nernst potential of $K^+$ decreases (which means it becomes more negative) in the postsynaptic cell (but remains unchanged in the presynaptic cell). This decreases the resting membrane potential (but only by a small amount) and the peak membrane potential value of the PSP (by a relatively larger amount – because the total $K^+$ conductance is much larger near the peak than at rest) even though glutamate receptor activation is unchanged. These changes together contribute to a diminished PSP (but no delay, slowing, or change in duration).

(vi) Increase in extracellular $K^+$ concentration. [10%]

Answer: The Nernst potential of $K^+$ increases (which means it becomes less negative) in *both* the pre- and postsynaptic cells. This increases the resting membrane potential (thus pushing it closer to the threshold for action potential initiation), thus advancing presynaptic action potential initiation, and accelerates the upstroke (once the membrane potential exceeds the threshold), decelerates the downstroke, and increases the amplitude of the action potential in the presynaptic cell, which together lead to advanced, accelerated, longer-lasting, and increased vesicle release, and in turn to advanced, accelerated, longer-lasting, and increased glutamate receptor activation in the postsynaptic cell. At the same time, the increase in $K^+$ Nernst potential in the postsynaptic cell increases the resting membrane potential (but only by a small amount) and the peak membrane potential value of the PSP (by a relatively larger amount – because the total $K^+$ conductance is much larger near the peak than at rest) even if glutamate receptor activation is unchanged. These pre- and postsynaptic changes together contribute to an advanced, faster rising, nut longer lasting, and larger PSP.

(b) This part is about conditioning in the Aplysia. For each of the pharmacological manipulations below, explain with reasons the effects it has on the gill withdrawal reflex, and its cellular mechanisms, before and after conditioning. Each pharmacological agent is applied throughout the experiment, CS+ is siphon touch, CS- is mantle shelf touch, and

US is tail shock.

(i) Application of a selective serotonin re-uptake inhibitor in the synapse that connects the facilitating interneuron (of the tail pathway) to the sensory neuron (of the siphon pathway). [10%]

Answer: This increases the concentration of serotonin in the synapse once the presynaptic cell emitted an action potential and thus released serotonin into this synapse. In turn, this increases serotonin receptor activation in the postsynaptic cell (the sensory neuron of the CS+ pathway), and thus amplifies the whole cascade of reactions underlying conditioning. Before conditioning, we never stimulate the tail, and thus the facilitating interneuron in this pathway does not become active, and so no changes happen. During conditioning, we stimulate the tail, and thus activate the facilitating interneuron. As a result, after conditioning, the gill withdrawer reflex will be even stronger for CS+ than after "normal" conditioning (without this manipulation), but unchanged in the CS- pathway.

(ii) Application of AP5 in the (siphon) sensory neuron-motor neuron synapse. [10%]

Answer: AP5 blocks NMDA receptors and thus diminishes the postsynaptic component of conditioning. As NMDA receptors only get activated during training, and only in the CS+ pathway, the gill withdrawal reflex does not change before training but is less strong than after "normal" conditioning for CS+, and unchanged for CS-.

(iii) Application of activated $Ca^{2+}$ / calmodulin kinase in the motor neuron. (You can assume that the $Ca^{2+}$ / calmodulin kinase plays a similar role as in hippocampal neurons.) [10%]

Answer: The activated $Ca^{2+}$ / calmodulin kinase phosphorylates AMPA receptors such that their maximal conductance is increased. This effect does not require the coincidence of pre- and postsynaptic neuron activation, and because the CS+ and CS- pathways converge onto the same motor neuron, it is also shared between these pathways. This means that the gill withdrawal reflex is amplified both before and after conditioning, and for both CS+ and CS-.

(iv) Application of lidocaine in the (siphon) sensory neuron. [10%]

Answer: Lidocaine blocks $Na^+$ channels and thus action potentials in the CS+ sensory neuron. This effect does not require the coincidence of pre- and postsynaptic neuron activation. Thus, the gill withdrawal reflex is entirely diminished for CS+, but remains unchanged for CS, both before and after conditioning.

4 (a) The following questions are about LTP.

(i) Describe the effect of an adenylyl cyclase blocker on LTP. [10%]

Answer: Blocks the induction of late LTP but leaves early LTP unaffected. As a result, existing synapses become stronger after LTP but no new synapses are formed.

(ii) Describe the synaptic tagging problem. [10%]

Answer: The expression of late LTP requires *de novo* protein synthesis in the nucleus, and the newly synthesised proteins to be then transported back to (or near) the synapse in which induction happened. The synaptic tagging problem means that there must be a way for the inducing synapse to be tagged following induction so that newly synthesised synaptic building blocks "find their way" back to it selectively.

(iii) Explain with reasons if the amplitude and latency of presynaptic stimulation-induced action potentials change following LTP. [10%]

Answer: The amplitude of action potentials does not change considerably. (The amplitude of the population spike increases but this question is about the action potential – which is necessarily about individual cells.) The latency of action potentials decreases because EPSP amplitudes have increased, and therefore cells will reach their firing threshold sooner.

(iv) Explain what LTD is. [10%]

Answer: LTD is long-term depression, expressed as a decrease of efficacy in the synapse after the appropriate stimulation protocol has been used to induce it.

(b) The following questions are about signals in a classical conditioning experiment with partial reinforcement in a rat. Explain with reasons what is expected for each of the signals below, in the given context, following the times of CS and US delivery.

(i) Reward prediction errors according to the Rescorla–Wagner rule after extensive training, when the learning rate is small or large, respectively. [15%]

Answer: The Rescorla–Wagner (RW) rule predicts positive reward prediction errors on US+ trials and negative reward prediction errors on US- trials. This is because the reward prediction error is the difference between the amount of actual reward on a given trials (0 on US- trials and 1 on US+ trials) and predicted reward, which in turn is around the average reinforcement rate (somewhere between 0 and 1 due to partial reinforcement). A large learning rate means that reward predictions converge faster to be around the average reinforcement rate, but also exhibit larger fluctuations around this

average. This means that reward prediction errors will also fluctuate more with a larger learning rate. Critically, according to the RW rule, reward prediction errors are only computed at the end of the trial (following the time of US delivery, whether it was presented or omitted), so they are undefined (or 0) following the delivery of the CS. Alternatively, if the RW rule is simply applied in each time step of a trial, then the reward prediction error would always be negative following the CS (as on US- trials following the delivery time of the US) because the prediction has already been computed (as it only requires the CS) but no reward has been delivered yet, and so it would simply be the negative prediction. As a result, it would also fluctuate more for a larger learning rate about this negative value.

(ii) Temporal difference prediction errors (their peak value) according to the temporal difference learning rule, when the learning rate is small or large, respectively. [15%]

Answer: In this paradigm, only a single CS is presented in each trial. In such paradigms, the temporal difference rule "inherits" its main predictions for what happens after US delivery time from the RW rule. Thus, those peak temporal difference prediction errors are the same as reward predictiomn errors above. Critically, at the time of CS delivery, the temporal difference prediction errors are the same as the predictions would be according to the RW rule, and so they are always positive, such that their magnitude is around the average reinforcement rate, with larger fluctuations for a larger learning rate.

(iii) The maximum of the across-trial average (peak) dopamine signal as a function of the reinforcement rate, $r$, in a modified version of the temporal difference prediction error model of dopamine signalling. In this version, dopamine (DA) is a piecewise linear function of the temporal difference prediction error ($\delta$) with a non-zero baseline ($\alpha > 0$) and a larger sensitivity for positive than for negative prediction errors:

$$\mathrm{DA} = \begin{cases} 5\alpha\delta + \alpha & \text{if } \delta \geq 0 \\ \alpha\delta + \alpha & \text{otherwise.} \end{cases}$$

Express your answers in mathematical form as a function of $\alpha$. [30%]

Answer: The temporal difference prediction error after the CS is the same as the predicted reward would be in the RW model (see above), which in turn is $\delta = r$, and this is the same in all trials. Thus, following the CS delivery time, the average dopamine signal is $\overline{\mathrm{DA}}_{\mathrm{CS}} = 5\,\alpha\,r + \alpha$, of which the maximum is at $r = 1$, when it takes the value $\overline{\mathrm{DA}}^{*}_{\mathrm{CS}} = 6\,\alpha$.
Following US delivery time, $\delta = 1 - r$, so dopamine is $\mathrm{DA} = 5\,\alpha\,(1 - r) + \alpha$ in US+ trials, and $\delta = -r$, so dopamine is $\mathrm{DA} = -\alpha\,r + \alpha$ on US- trials. The rate of US+ and US- trials is $r$ and $1 - r$, respectively, by definition. The average dopamine signal across all trials is thus the weighted average

of dopamine signals in US+ and US- trials, weighted by the respective rates:

$$\overline{\mathrm{DA}}_{\mathrm{US}} = [5\,\alpha\ (1-r)+\alpha]\ r + [-\alpha\, r+\alpha]\ (1-r) \tag{9}$$

$$= 5\,\alpha\ (1-r)\ r - \alpha\, r\ (1-r) + \alpha\, r + \alpha\ (1-r) \tag{10}$$

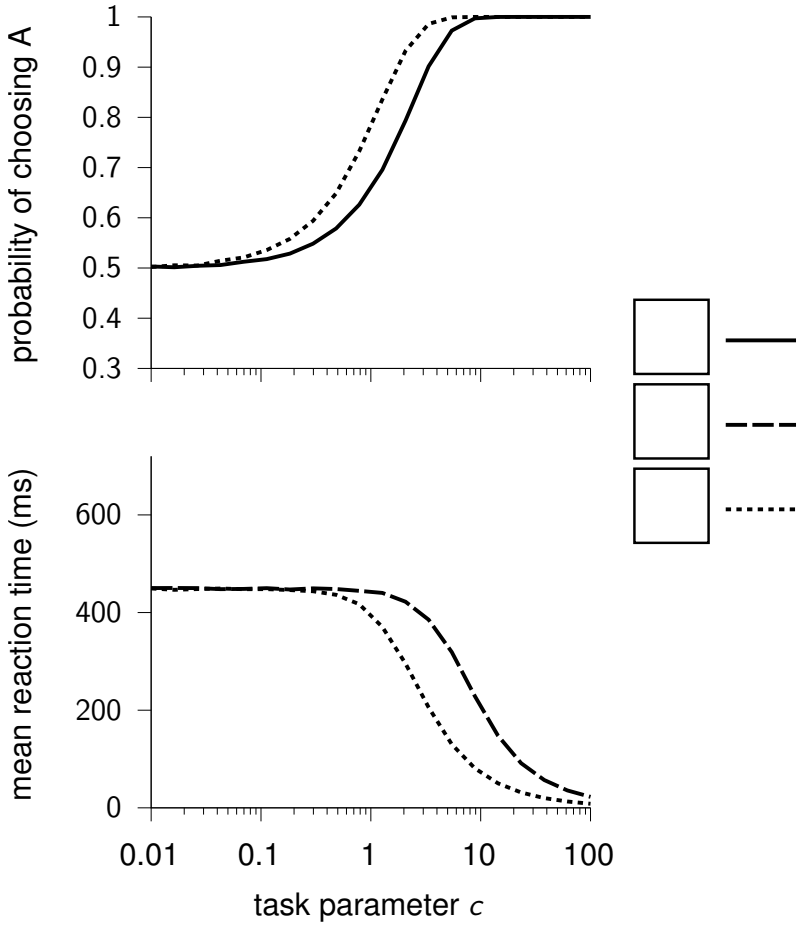
$$= 4\,\alpha\ (1-r)\ r + \alpha \tag{11}$$

(Critically, it is the dopamine signal rather than the temporal difference prediction error signal that needs to be averaged across trials. Because the former is a non-linear function of the latter, the average dopamine signal is not the same as the dopamine signal that would correspond to the average temporal difference prediction error.) It is easy to see that the maximum of Eq. 11 is at $r = 0.5$, and so the maximum average dopamine signal is at $\overline{\mathrm{DA}}^{*}_{\mathrm{US}} = 2\,\alpha$

**END OF PAPER**

Version GH/3 – CRIB Candidate Number:

EGT2
ENGINEERING TRIPOS PART IIA
Monday 4 May 2026, Module 3G3 – CRIB, Question 1.




Additional copy of Fig. 1. This should be annotated as specified in Question 1, and handed in with your answers to that question.