

Navigation task serotonin mice

Navigation task

Task design on miniworld

Use a long T-maze with a diagonally striped wallpaper. Say the actions available to the agent are discrete, and are up, down, left, right.

Visuomotor gain

Define 3 visuomotor gain modes g :

- Slow or $g_s = 1 : 3$
 - Action a executed once
- Normal or $g_n = 1 : 1$
 - Action a executed 3x
- Fast or $g_f = 3 : 1$
 - Action a executed 9x

Wallpaper phase

Define the phase of the wallpaper as θ , where $\theta \in [0, 2\pi]$.

⚡ Skip for v0

I suggest that we skip the wallpaper phase manipulation for our first attempt to make things simpler. It would unnecessarily complicate the task generation and honestly, I'm not 100% certain that it would be possible.

Pseudocode

- Sample a random time step t_{switch}
 - $t \sim U[t_{1/4}, t_{3/4}]$
 - uniformly at random from somewhere between 1/4 and 3/4th the time it takes the agent (empirically) to get to the end of the corridor

- Sample two random task modes
 - $g_0 \sim \{g_s, g_n, g_f\}$
 - $g_{\text{switch}} \sim \{g_s, g_n, g_f\}$
- Sample two random phases
 - $\theta_0 \sim [0, 2\pi]$
 - $\theta_{\text{switch}} \sim [0, 2\pi]$
- At t_0 , start task with gain g_0 and phase θ_0
- At t_{switch} , switch gain to g_{switch} and set wallpaper phase at current location $x_{t_{\text{switch}}}$ to θ_{switch}

≡ My terrible attempt at defining episodes

- Regular episode:
 - $(g_{t=0}, \theta_{t=0})$
- Gain-manipulation episode:
 - $(g_{t=0}, \theta_{t=0}, g_{t=t_{\text{switch}}})$
- Gain and phase manipulation episode:
 - $(g_{t=0}, \theta_{x=0, t=0}, g_{t=t_{\text{switch}}}, \theta_{x=x_{t_{\text{switch}}}, t=t_{\text{switch}}})$

Moving MNIST task

⚠ Plot predictions instead of reconstruction

Instead of plotting the single trial video frames x_1, \dots, x_t and its reconstruction, i.e. $p(\hat{x}_t|z_t)$, we should plot the prediction given the history of observations and/or latents, i.e. $p(\hat{x}_t|\hat{z}_t)p(\hat{z}_t|z_{t-1}, h_{t-1}, a_{t-1})$.

In other words, use the MDN-RNN to predict the next latent and use the VAE decoder to generate an observation based on that. Doing this will demonstrate exactly what happens at the time of the glitch in a human-interpretable manner.

👉 Sort neurons in plots

For the position glitch, direction glitch, digit glitch, etc., sort the neurons (latent dimensions) by their magnitude of response to each glitch at the

time of the glitch. Maintain this sorting order across all glitches and then plot all of them together. This will help showcase the diversity of responses to the different features of the SPE in a visually appealing manner.

Task manipulations

- ☐ Speed glitch
- ☐ Bounce glitch
- ☐ Digit glitch with same digit class

For the last one, we would need to build a classifier that samples MNIST digits, classifies them into their respective digit classes, and then samples them appropriately. A little tedious, but very doable.

Non-learnable glitches: Random Pixel noise

We demonstrate the state prediction error (SPE) effect on several task manipulations, each of which evokes a varying magnitude of the SPE signal. The point we would like to make is that this prediction error is in the latent state space, which is highly structured, as opposed to the pixel space, which is kind of meaningless. If we can show that the random errors in pixel space, which are unpredictable, do not evoke a strong response as the structured errors do, that might be a nice figure to have. If we can control the magnitude of the noise (with the pixel)

🤔 Low priority?

I reckon that this point is lower priority than getting the hallway task working, but I do agree that it may be an important point to make for the neuroscientists out there.

Other points of discussion

Points to emphasize in the article

Serotonin response captures:

1. a low-D (but not 1D) latent state prediction error signal
2. only to *attended* or *relevant* features of state

3. that is modulated by learning

Points 2 & 3 are subtle. In the model, this may only be demonstrated by restricting the training data in a manner that we can or cannot define what is *relevant* for the agent in its self-supervised learning task.

References

Added to [the Zotero library](#):

- Ranade & Mainen 2009
 - shows the multiplicity of responses of DRN neurons
 - low-D but not one-D signal
- Glascher, Daw, Dayan, and O'Doherty 2010
 - State prediction errors linked to Dopamine