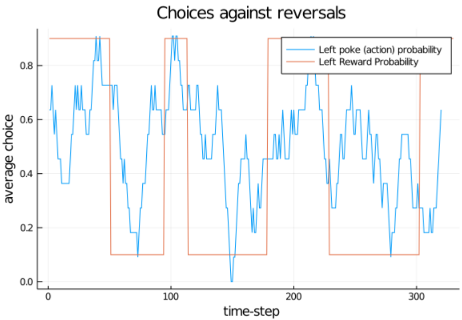
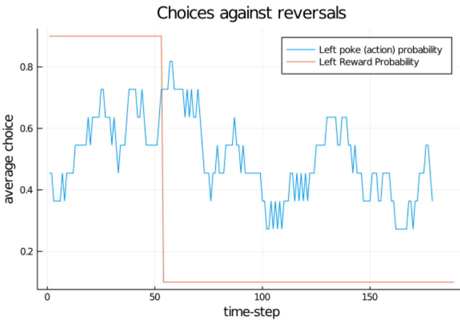
Model fitting and Comparison Results

Beren Millidge

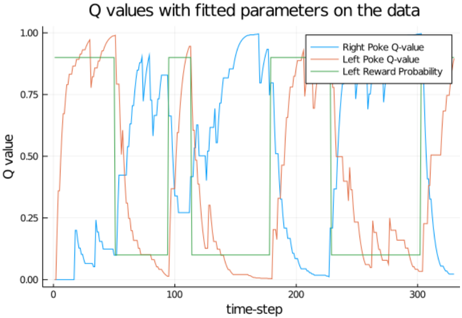
September 24, 2022

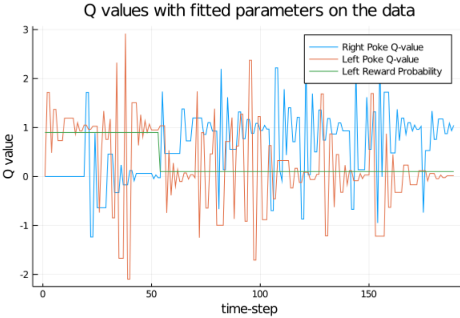
Visualizing Mouse Behaviour



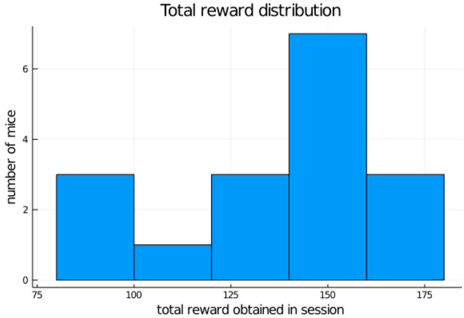


Visualizing Model Fit





Individual Differences and Model Fit



Basic Q learning Model

- Main model fitted to data here was simple Q learning model without decay. That is,

$$Q_{t+1} = Q_t + \alpha(Q_t - r_t) \tag{1}$$
$$p(a) = \sigma(\beta * Q) \tag{2}$$

Side Bias Model

- Tried fitting a model with explicit side biases to the data.

$$p(a) = \sigma(\beta * (Q + B)) \tag{3}$$

where B is a bias term consisting of two parameters $B = [B_L, B_R]$ where B_L is the left side bias and B_R is the right side bias.
- Compared with the standard Q learning model using likelihood ratio test, as the models are nested.
- No significant improvement given the side bias model $p = 0.9$

Softmax perserveration bias model

- Also tried fitting a model where the previously selected choice gets a bias in the softmax.

$$p(a) = \sigma(\beta * (Q + \gamma I[a_{t-1}])) \tag{4}$$

- Where $I[]$ is an indicator function which is 1 for the previously selected action. γ is the parameter which controls the strength of this bias.
- Compared to standard Q learning with another likelihood ratio test.
- Also not significant. Funnily enough got p=1, d=0.001 – i.e. no distinct difference in log likelihoods at all for this one.

Random Perserveration Model

- Tried fitting another perserveration model where instead of adding a softmax bias, the mouse simply chooses to repeat its previous action with some probability. i.e.

$$a_t = a_{t-1} \ if \ rand() < \gamma \tag{5}$$
$$else : p(a_t) = \sigma(\beta * Q) \tag{6}$$

- γ is a parameter which controls the probability of simply repeating the last action.
- Found this model to be highly significant vs the baseline Q learning model ($p = 0.001$).
- Not sure why this model differs so much from the softmax perserveration model bias

Different Learning rates in force vs choice

- Here I tried to fit a model with both a force trial learning rate α_{force} and a choice trial learning rate α_{choice}.
- In general optimization/parameter fitting was not super stable for this and typically exhibited bistability
- I.e. two attractors depending on initial conditions. 1.) with force and choice learning rates about the same. 2.) A significantly larger *force-choice* learning rate. I.e. the mice learn better on the force trials.
- This bistability is a pretty general feature of the optimization for all the good mice trials.
- Instability and bias towards force may be due to data artefact where the sessions always start with force trials which is also when the mice are first exposed to the task and so learning the most.
- If we cut out first 25 steps, optimization typically converges to force and choice being about the same, but not always.

Conclusions

- In general, I think we can say that the animals *do learn* from force trials pretty conclusively. But there is still a question of whether they learn *more* from force trials.
- A lot of variability between mice – some don't seem to be acting like RL agents at all while others do
- Model fit (as measured by log likelihood) correlates very closely with how good the mouse is
- Haven't found much evidence of side bias in *good* mice (but this might be why they are deemed "good" in the first place.
- Some evidence for randomly selecting to keep same action as last time.
- Optimization is usually pretty stable and gives similar results for all the good mice/sessions, except for the different force vs choice learning rates.