A Coupled Fokker-Planck Model for Collective Decision-Making

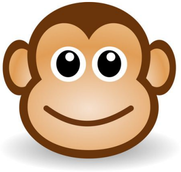
Kyle Fitzgerald

CU Boulder

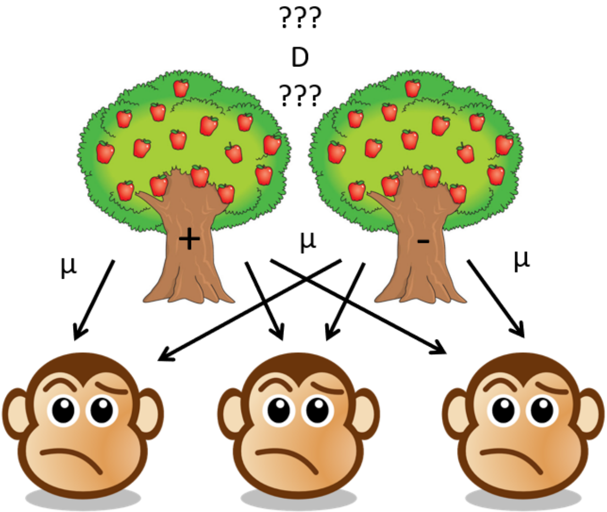
December 25, 2021

Background

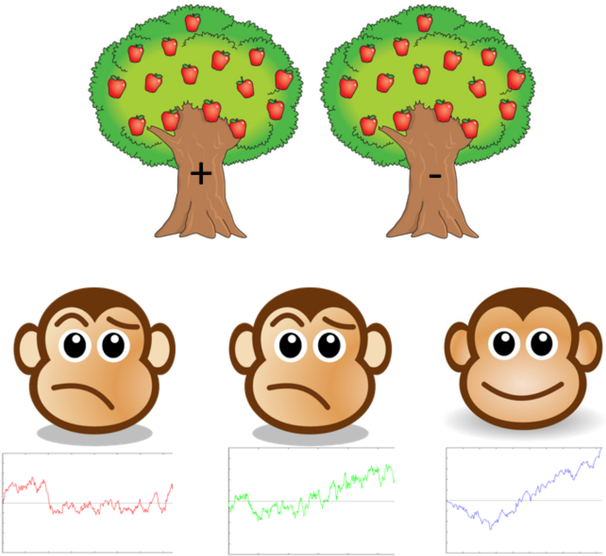
Imagine that you are a monkey.



You are looking to scope out possible foraging spots with companions. For simplicity, suppose we break this endeavor into a sequence of decisions between two trees.



How do **you** make a decision, using both your **own** evidence along with knowledge about companions' previous decisions?



Big Question: How to Maximize Collective Reward Rate?

- Where do you set your own **decision thresholds**?
- How much of a **kick** to your own decision state do you take from others' decisions? Their decisions are **quicker**, but are they **reliable**?
- How do these considerations depend on the **amount of reward** you expect from each choice, assuming it is the **correct choice**?
- Can a group with appropriately-chosen parameters achieve a **higher average reward rate** than individuals?

First, a Single-Agent Model

$$dX = \mu\, dt + \sqrt{2D}\, dW(t), \qquad X(0) = 0$$

$$\Updownarrow$$

$$\frac{\partial c}{\partial t} = -\mu \frac{\partial c}{\partial x} + D \frac{\partial^2 c}{\partial x^2}, \qquad c(-L, t) = c(H, t) = 0$$

- $X(t)$: current decision state (decision thresholds set at $X = H$ and $X = -L$)
- μ: environmental drift
- D: strength of environmental noise/uncertainty
- $c(x, t)$: probability density function

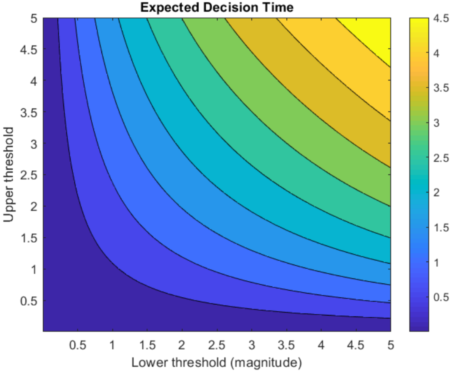
Single-Agent Reward Rate

$$RR = \frac{\langle r \rangle}{\langle T \rangle + T_I}$$

- Assume a $\mu = \pm 1$ environment with probability $\frac{1}{2}$
- $\langle T \rangle$: expected decision time
- $\langle r \rangle$: expected amount of reward per decision
- T_I: transient time between sequential decisions
  - Needs to be large enough!
  - We use $T_I = 5$

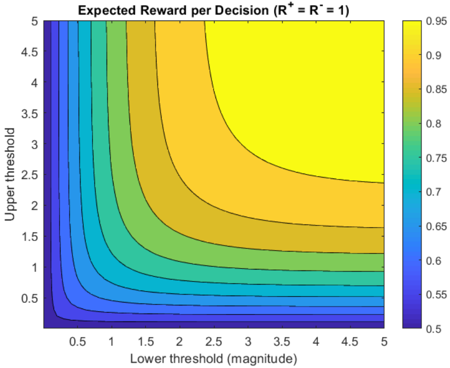
Single-Agent Expected Time

$$\langle T \rangle = \langle T \mid \mu = 1 \rangle P(\mu = 1) + \langle T \mid \mu = -1 \rangle P(\mu = -1)$$
$$= \frac{1}{2} \frac{(1 - e^L)(1 - e^H)(L + H)}{e^{L+H} - 1}$$

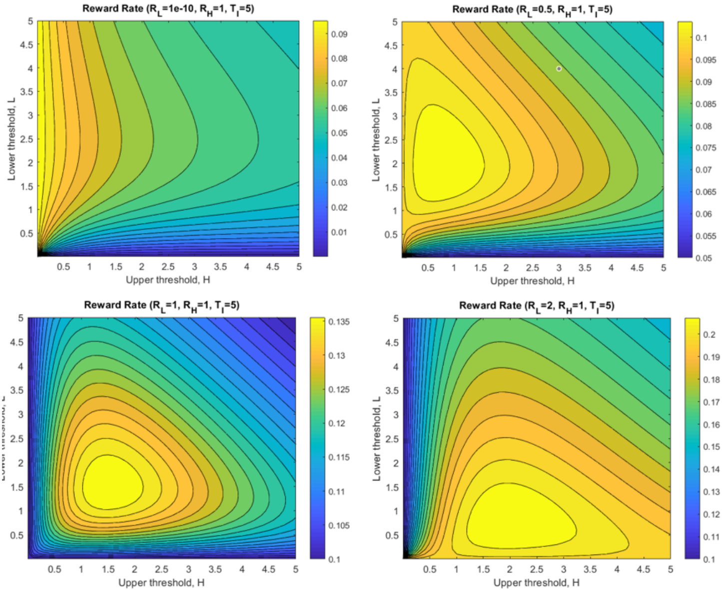


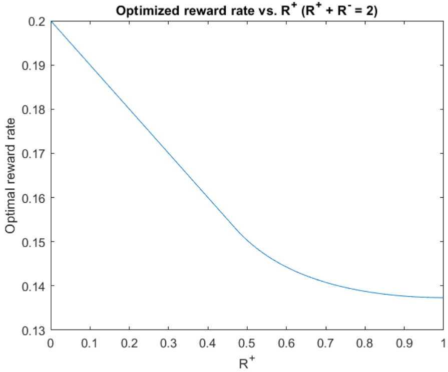
Single-Agent Expected Reward

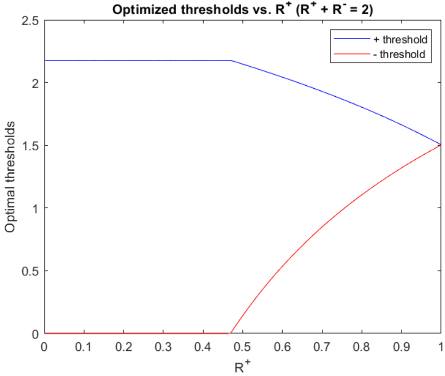
$$
\begin{aligned}
\langle r \rangle &= R^+ \cdot P(\text{decides at } H \,|\, \mu = 1) P(\mu = 1) \\
&\quad + R^- \cdot P(\text{decides at } -L \,|\, \mu = -1) P(\mu = -1) \\
&= \frac{1}{2} \frac{R^+(e^L - 1)e^H + R^-(e^H - 1)e^L}{e^{L+H} - 1}
\end{aligned}
$$



Putting it All Together







Notice how **immediate decision** is optimal for a range of reward ratios.

Coupled Two-Agent Drift-Diffusion Model

$$
\begin{aligned}
dX_1 &= \mu_1\,dt + \sqrt{2D}\,dW_1(t) + G_2 q_\pm \delta(t - t_2), \qquad X_1(0) = 0,\\
dX_2 &= \mu_2\,dt + \sqrt{2D}\,dW_2(t) + G_1 q_\pm \delta(t - t_1), \qquad X_2(0) = 0
\end{aligned}
$$

- t_i: the time at which agent i decides
- G_i: sign of the threshold at which agent i makes decision
- $q_\pm$: "kick" size in $\pm$ direction after other agent's decision

- Assume an undecided agent is forced to decide at the nearest threshold *immediately* after first decision
- Assume $\mu_1 = \mu_2$ in any given environment

Reward Considerations

For simplicity, we try to minimize the number of parameters in the reward rate optimization problem with the following reward methods:

- "Agent-symmetric" rewards
  - $R_1^+ = R_2^+$ and $R_1^- = R_2^-$
  - Optimal strategy gives $H_1 = H_2$ and $L_1 = L_2$
- "Agent-asymmetric" rewards
  - $R_1^+ = R_2^-$ and $R_1^- = R_2^+$
  - Optimal strategy gives $H_1 = L_2$ and $L_1 = H_2$

The Agent-Symmetric Case: Reward Rate

$$RR = \frac{2\langle r \rangle}{\langle T \rangle_F + T_I}$$

- $\langle T \rangle_F$: expected time for the **first** agent to decide

$$\begin{aligned}\langle T \rangle_F = &\langle T_i \,|\, \mu = 1 \text{ and } t_i < t_j \rangle P(\mu = 1) \\ &+ \langle T_i \,|\, \mu = -1 \text{ and } t_i < t_j \rangle P(\mu = -1)\end{aligned}$$

- $\langle r \rangle$: a bit more complicated than before!

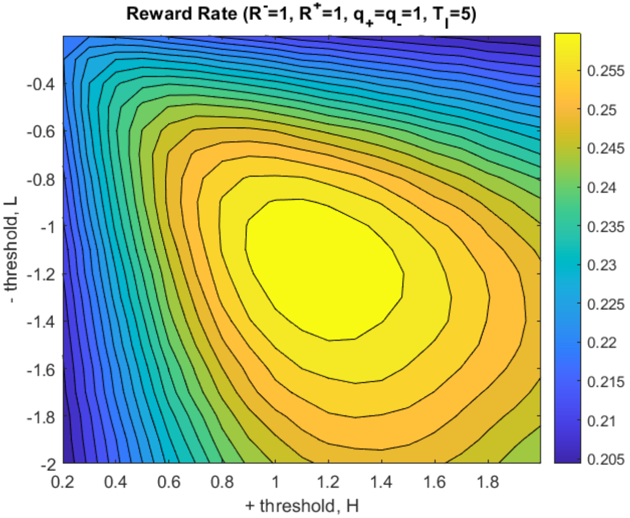
$$\begin{aligned}\langle r \rangle = R^{+} \cdot P(i \text{ decides at } H \,|\, \mu = 1)P(\mu = 1) \\ + R^{-} \cdot P(i \text{ decides at } -L \,|\, \mu = -1)P(\mu = -1)\end{aligned}$$

where

$$\begin{aligned}&P(i \text{ decides at } x \,|\, \mu) \\ &\quad = P(i \text{ decides at } x \,|\, \mu \text{ and } t_i < t_j)P(t_i < t_j \,|\, \mu) \\ &\quad + P(i \text{ inst. crosses } x \,|\, \mu \text{ and } t_i > t_j)\big(1 - P(t_i < t_j \,|\, \mu)\big).\end{aligned}$$

By symmetry, we have $P(t_i < t_j \,|\, \mu = 1) = \frac{1}{2}$.

Two-Agent Reward Rate Example



Future Work

- Generate figures similar to those the single-agent plots above, keeping $R^+ + R^- = 2$
  - Is there still a nontrivial region of rewards for which agents should decide immediately?
  - Does the two-agent model outperform the single-agent model, i.e. should you listen to your companions under this model?
- Allow diffusion after first decision instead of enforcing "immediate interrogation."

Thank you!