Equations

$$
\begin{aligned}
\omega_t &= \lambda X_t + (1-\lambda)\omega_{t-1} && (1)\\
\gamma_{11}(t) &= \left[1+\exp\left(\alpha\left(\frac{1}{2}-\omega_t\right)\right)\right]^{-1} && (2)\\
\gamma_{22}(t) &= \left[1+\exp\left(\beta\left(\frac{1}{2}-\omega_t\right)\right)\right]^{-1} && (3)\\
L(R_t=r|S_t,Y_{t-1}) &= \begin{cases} \frac{P(Y_{t-1}|R_t=r)}{\sum_j P(Y_{t-1}|R_t=j)}, & \text{if } S_t=1 \\ \\ \frac{1-P(Y_{t-1}|R_t=r)}{\sum_j [1-P(Y_{t-1}|R_t=j)]}, & \text{if } S_t=2 \end{cases} && (4)\\
\vartheta(R_t=r|S_t,Y_{t-1}) &= \frac{L(R_t=r|S_t,Y_{t-1})h(R_{t-1}=r)}{\sum_j L(R_t=j|S_t,Y_{t-1})h(R_{t-1}=j)} && (5)\\
h_R(t) &= (\pi)g_R + (1-\pi)\vartheta_R(t) && (6)
\end{aligned}
$$

Parameters

$\lambda \in (0,1]$

$\alpha \in (0,10]$

$\beta \in [-10,0)$

$\pi \in (0,1)$

Model

Eq. (1) is a 1-parameter smoothing function which does not differentiate between reward and punishment processing. Eq. (2) and (3) are logistic functions centered in $0-1$ with different slopes and help the model to differentiate between reward and punishment related behaviour. Eq. (4) is the likelihood computed over the rules at trial t, based on both response at $t-1$ and state switching at t. Eq. (5) is the posterior over the rules at trial t, computed by taking into account the prior distribution over the rules before state transition, that is, at trial $t-1$, and the likelihood. The posterior at trial t becomes the new prior for trial t, and the rule is sampled. However, the system can fail in relying on the posterior/prior due to distractibility. Eq. (6) represents the corrupted posterior/prior.