Reward: $r(x) = R(x) - \sum_{j=1}^{m} S_j(u_j)$; $S(x) = \frac{2}{\pi} \arctan\left(\frac{\pi}{2} x\right)$ $x; \dot{\theta}, \ddot{\theta}$

Error: $\delta(t) = r(t) + \frac{1}{\Delta t}\left[\left(1 - \frac{\Delta t}{\tau}\right) V(t) - V(t - \Delta t)\right]$

Actor: $u(t) = S\left(A\left(x(t); w^A\right) + \sigma n(t)\right)$

$0 < \kappa < \tau$

$\sigma = \sigma_0 \min\left[1, \max\left(0, \frac{V_1 - V(t)}{V_1 - V_0}\right)\right]$

$\dot{n}(t) = -n(t) + N(t)$; $N(t)$ gaussian noise

$\dot{w}_i^A = \eta^A \cdot \delta(t)\, n(t) \frac{\partial A\left(x(t); w^A\right)}{\partial w_i^A}$ gradient

Critic: $\dot{e}_i(t) = \frac{-1}{\kappa} e_i(t) + \frac{\partial V(x(t); w)}{\partial w_i}$; eligibility trace

$\dot{w}_i = \eta\, \delta(t) \cdot e_i(t)$

NGnet, $V(x; w) = \sum_{k=1}^{K} w_k b_k(x)$; $b_k(x) = \frac{a_k(x)}{\sum_{l=1}^{K} a_l(x)}$; $a_k(x) = e^{-\| s_k^T (x - c_k) \|^2}$

instead we will use mean field populations

$\frac{\partial V(x; w)}{\partial w} = \frac{\partial V}{\partial x} \cdot \frac{\partial x}{\partial w} = \sum_{k=1}^{K} w_k\, b_k'(x) \cdot b_k(x)$

$w_k\, b_k'(x)$ $\quad b_k(x)$

$a_k'(x) = 2 s_k^2 (c_k - x)\, e^{-s_k^2 (c - x)^2}$

$b_k'(x) = \frac{a_k'(x) \sum_{l=1}^{K} a_l(x) - a_k(x) \sum_{l=1}^{K} a_l'(x)}{\left[\sum_{l=1}^{K} a_l(x)\right]^2}$