Dezfouli's model Replication

Teddy Landron

July 2017

1 Equations

Eq. 1.1 :

$$Q(s_t, a_t) = E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \ldots | s_t, a_t] = E\left[\sum_{i=t}^{\infty} \gamma^{i-t} r_i | s_t, a_t\right] \quad (1)$$

Eq. 1.2 :

$$\delta_t = \gamma^{d_t}(r_{t+1} + V(s_{t+1})) - Q(s_t, a_t) \quad (2)$$

Eq. 1.3 :

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \delta_t \quad (3)$$

Eq. 1.4 :

$$\delta_t^c = max(\gamma^{d_t}(r_{t+1} + V(s_{t+1})) - Q(s_t, a_t) + D(s_t), D(s_t)) \quad (4)$$

Eq. 2.1 :

$$Q(s_t, a_t) = E\left[\sum_{i=t}^{\infty} (r_i - \bar{R}_i) | s_t, a_t\right] \quad (5)$$

Eq. 2.2 :

$$\bar{R}_{t+1} = (1 - \sigma)\bar{R}_t + \sigma r_t \quad (6)$$

Eq. 2.3 :

$$\delta_t = r_t + V(s_{t+1}) - Q(s_t, a_t) - \bar{R}_t \quad (7)$$

Eq. 2.4 :

$$\delta_t^c = max(r_t + V(s_{t+1})) - Q(s_t, a_t) + D(s_t), D(s_t)) - \bar{R}_t \quad (8)$$

Eq. 2.5 :

$$r_t = \delta_t - V(s_{t+1}) + Q(s_t, a_t) + \bar{R}_t \quad (9)$$

Eq. 2.6 :

$$r_t^c = \delta_t^c - V(s_{t+1}) + Q(s_t, a_t) + \bar{R}_t \quad (10)$$

Eq. 2.7 :

$$\rho_t = \bar{R}_t + \kappa_t \quad (11)$$

Eq. 2.8 :

$$\kappa_{t+1} = (1 - \lambda)\kappa_t + \lambda N \quad (12)$$

Eq. 2.9 :

$$\kappa_{t+1} = (1 - \lambda)\kappa_t \quad (13)$$

Eq. 2.10 :

$$\delta_t^c = max(r_t + V(s_{t+1}) - Q(s_t, a_t) + [D(s_t) - \kappa_t], [D(s_t) - \kappa_t] - \bar{R}_t \quad (14)$$

2 Used equations

Equations used in the model's code:

Eq. 1.3 :

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\delta_t \tag{3}$$

Eq. 2.2 :

$$\bar{R}_{t+1} = (1 - \sigma)\bar{R}_t + \sigma r_t \tag{6}$$

Eq. 2.8 :

$$\kappa_{t+1} = (1 - \lambda)\kappa_t + \lambda N \tag{12}$$

Eq. 2.10 :

$$\delta_t^c = max(r_t + V(s_{t+1}) - Q(s_t, a_t) + [D(s_t) - \kappa_t], [D(s_t) - \kappa_t] - \bar{R}_t \tag{14}$$

3 Appendix

3.1 Appendix A

Parameter	Value
σ	0.005
$D(s_t)$	15
α	0.2
μ_N	5
σ_N	0.02
μ_{fr}	-2
σ_{fr}	0.02
μ_{sh}	-200
σ_{sh}	0.02
μ_c	2
σ_c	0.02
C_u	6
λ	0.0003
N	2
μ_s	1
σ_s	0.02
μ_l	15
σ_l	0.02
ϵ	0.1
k	7ts

Table 1: Simulation Parameters' Values