
HW4-1 - Policy Gradient Report

組員：陳泓均、陳欽安、詹書愷、丁昱升

Model

Dense(80*80, 10)

tanh()

Dense(10, action_space_num)

softmax()

LR=0.02

GAMMA = 0.99

Adam optimizer

Preprocessing

- $(210, 160, 3) \rightarrow (80, 80, 1)$
- resize & down-sample the image
- set background to **0**, others to **1**
- **flatten** the image

Tip - Variance Reduction

$$\nabla \bar{R}_\theta \approx \frac{1}{N} \sum_{n=1}^N \sum_{t=1}^{T_n} (\underbrace{R(t^n) - b}_{\text{Can be state-dependent}}) \nabla \log p_\theta(a_t^n | s_t^n)$$

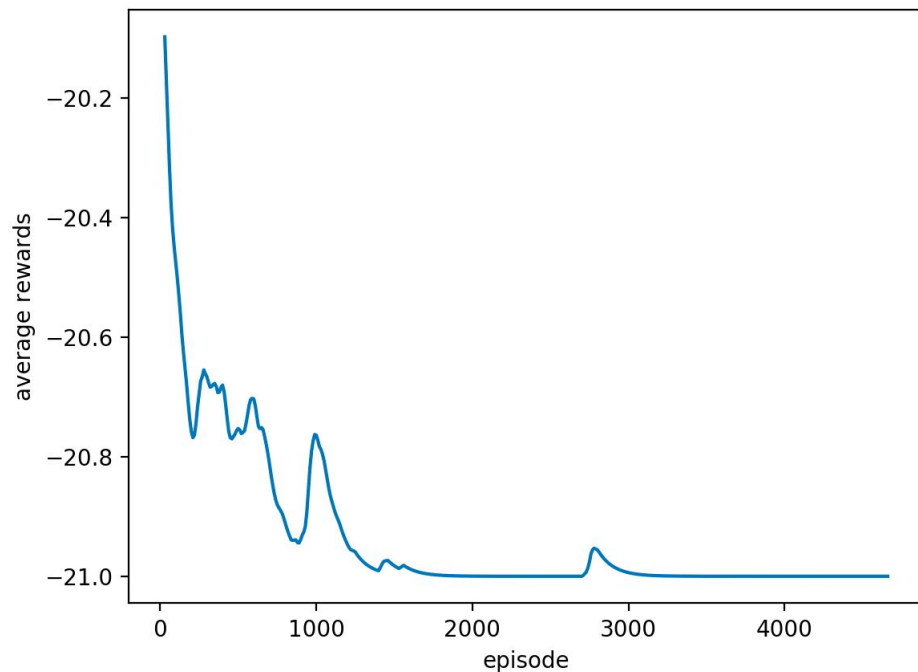
Can be state-dependent

↓

$$\sum_{t'=t}^{T_n} r_{t'}^n \rightarrow \sum_{t'=t}^{T_n} \gamma^{t'-t} r_{t'}^n$$

Add discount factor $\gamma < 1$

Learning Curve - Without Variance Reduction



Learning Curve - With variance reduction

