Notes on Reinforcement Learning and Deep Reinforcement Learning

compiled by D.Gueorguiev, 3/20/25

# Introductory Notes

What is Reinforcement Learning: branch of machine learning concerned with making decisions and taking sequences of actions based on some current state thereby maximizing some reward objective over time.

action

Environment

Agent

state, reward

Figure 1: Feedback loop between the Agent and the Environment in RL

The Agent and the Environment interact with each other on discrete timesteps creating a feedback loop depicted in Figure 1. The Agent has a goal of maximizing the cumulative reward while interacting with the Environment.

Observations in RL:

Robotics: camera images, joint angles

Actions in RL:

Robotics: joint torques

Rewards in RL:

Robotics: stay balanced, navigate to target locations

# References

[1] [Deep Reinforcement Learning: Lecture 1: Intro, Episodic Reinforcement Learning and Markov Decision Processes, John Schulman, OpenAI, Berkeley (MLSS Cadiz, 2016)](https://youtu.be/aUrX-rP_ss4?si=cxFW4A2UwqBzVS_S)

[2] [Deep Reinforcement Learning: Lecture 2: More on Cross-Entropy Method and Value Functions, John Schulman, OpenAI, Berkeley (MLSS Cadiz, 2016)](https://youtu.be/oPGVsoBonLM?si=NMt3SSkhK1o22Uww)

[3] [Deep Reinforcement Learning: Lecture 3: More on Episodic Reinforcement Learning and Policy Gradients, John Schulman, OpenAI, Berkeley (MLSS Cadiz, 2016)](https://youtu.be/rO7Dx8pSJQw?si=h5v2bh-Se-CDV93I)

[4] [Deep Reinforcement Learning: Lecture 4: Performance of Policies, Policy Approximations, Parametrized Policies, Asynchronous Methods, John Schulman, OpenAI, Berkeley (MLSS Cadiz, 2016)](https://youtu.be/gb5Q2XL5c8A?si=6fYCJE75Kcs93Ldr)

[5] [Reinforcement Learning Course: Lecture 1: Intro to Reinforcement Learning, David Silver, DeepMind x UCL, 2015](https://youtu.be/2pWv7GOvuf0?si=4EW5Iv50whZLmrqE)

[6] [Reinforcement Learning Course: Lecture 2: Markov Decision Processes, David Silver, DeepMind x UCL, 2015](https://youtu.be/lfHX2hHRMVQ?si=wuSa7hl5sIcNBg7r)

[7] [Reinforcement Learning Course: Lecture 3: Planning by Dynamic Programming, David Silver, DeepMind x UCL, 2015](https://youtu.be/Nd1-UUMVfz4?si=8W0QmfwGtYJhSswb)

[8] [Reinforcement Learning Course: Lecture 4: Model-Free Prediction, David Silver, DeepMind x UCL, 2015](https://youtu.be/PnHCvfgC_ZA?si=pMy18AnRs8K0M6vX)

[9] [Reinforcement Learning Course: Lecture 5: Model-Free Control, David Silver, DeepMind x UCL, 2015](https://youtu.be/0g4j2k_Ggc4?si=wfbcjR3q_ui0exaU)

[10] [Reinforcement Learning Course: Lecture 6: Value-Function Approximation, David Silver, DeepMind x UCL, 2015](https://youtu.be/UoPei5o4fps?si=mhU3k64hfpL4r2kJ)

[11] [Reinforcement Learning Course: Lecture 7: Policy Gradient Methods, David Silver, DeepMind x UCL, 2015](https://youtu.be/KHZVXao4qXs?si=9YTXC1M--8bAdZpH)

[12] [Reinforcement Learning Course: Lecture 8: Integrating Learning and Planning, David Silver, DeepMind x UCL, 2015](https://youtu.be/ItMutbeOHtc?si=CiYa80AdWfH84TeT)

[13] [Reinforcement Learning Course: Lecture 9: Exploration and Exploitation, David Silver, DeepMind x UCL, 2015](https://youtu.be/sGuiWX07sKw?si=FMhQ1UKGpDOaTJLw)

[14] [Reinforcement Learning Course: Lecture 10: Classic Games, David Silver, DeepMind x UCL, 2015](https://youtu.be/kZ_AUmFcZtk?si=cVKj56KPGNrPabX6)

[15] [Learning Tetris Using the Noisy Cross-Entropy Method, Istvan Szita, Andras Loerincz, 2006](https://github.com/dimitarpg13/reinforcement_learning_and_game_theory/blob/main/articles/ReinforcementLearning/Learning_Tetris_Using_the_Noisy_Cross-Entropy_Method_Szita_2006.pdf)

[16] [Approximate Dynamic Programming Finally Performs Well In The Game of Tetris, Victor Gabillon et al, INRIA, 2013](https://github.com/dimitarpg13/reinforcement_learning_and_game_theory/blob/main/articles/ReinforcementLearning/NIPS-2013-approximate-dynamic-programming-finally-performs-well-in-the-game-of-tetris-Paper.pdf)

[17] [A Tutorial On The Cross-Entropy Method, Pieter-Tjerk de Boer et al, MIT, 2003](https://github.com/dimitarpg13/reinforcement_learning_and_game_theory/blob/main/articles/ReinforcementLearning/A_Tutorial_on_the_Cross-Entropy_Method_deBoer_2003.pdf)