Ahmet Akman 2442366
Middle East Technical University
Electrical and Electronics Engineering

May 24, 2024

HOMEWORK 3 — Report

1 Questions

1.1 Agent:

1.2 Environment:

1.3 Reward:

1.4 Policy:

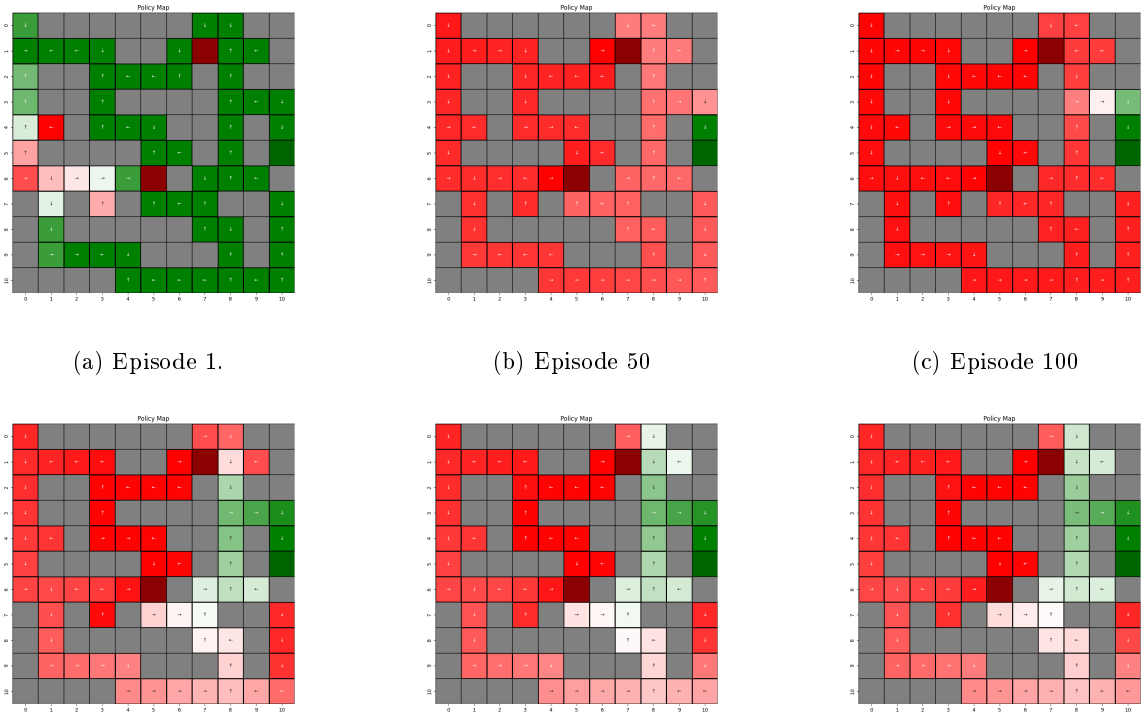
1.5 Exploration:

1.6 Exploitation:

2 Experimental Work

2.1 Temporal Difference Learning Default Parameters

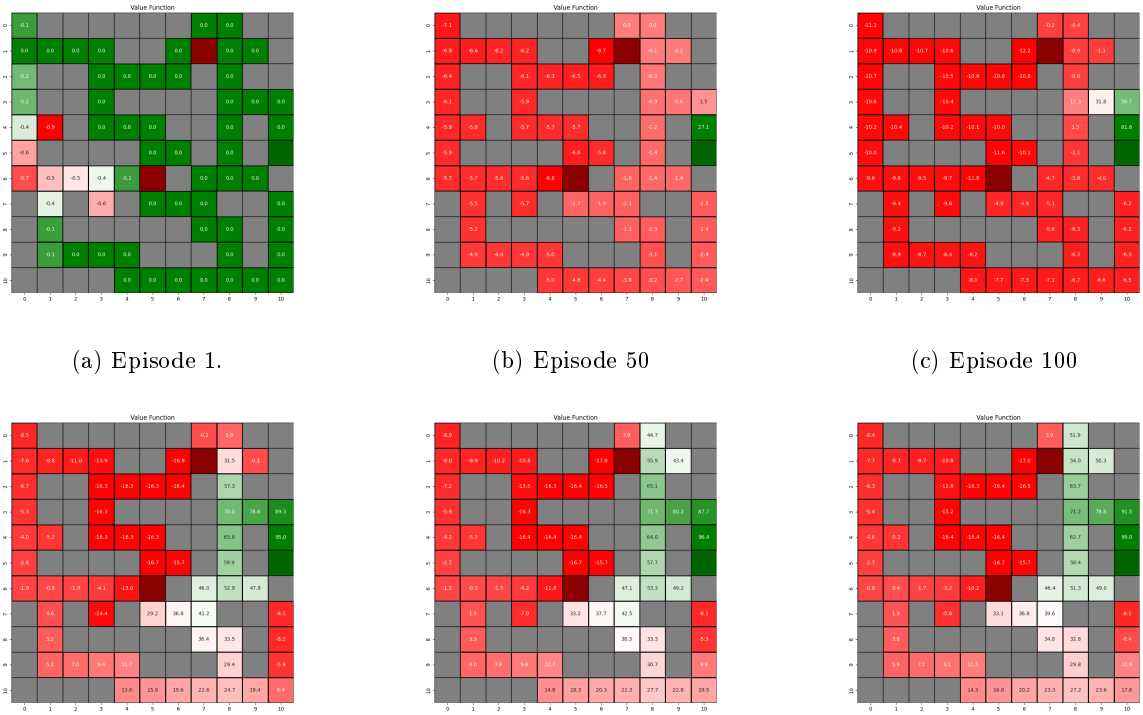
2.2 Q-Learning Default Parameters



(a) Episode 1. (b) Episode 50 (c) Episode 100

(d) Episode 1000. (e) Episode 5000 (f) Episode 10000

Figure 1: Evolution of policy maps throughout episodes.



(a) Episode 1. (b) Episode 50 (c) Episode 100

(d) Episode 1000. (e) Episode 5000 (f) Episode 10000

Figure 2: Evolution of value function throughout episodes.

Appendix

The code set used throughout this homework is provided as follows.

Submitted by Ahmet Akman 2442366 on May 24, 2024.