

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

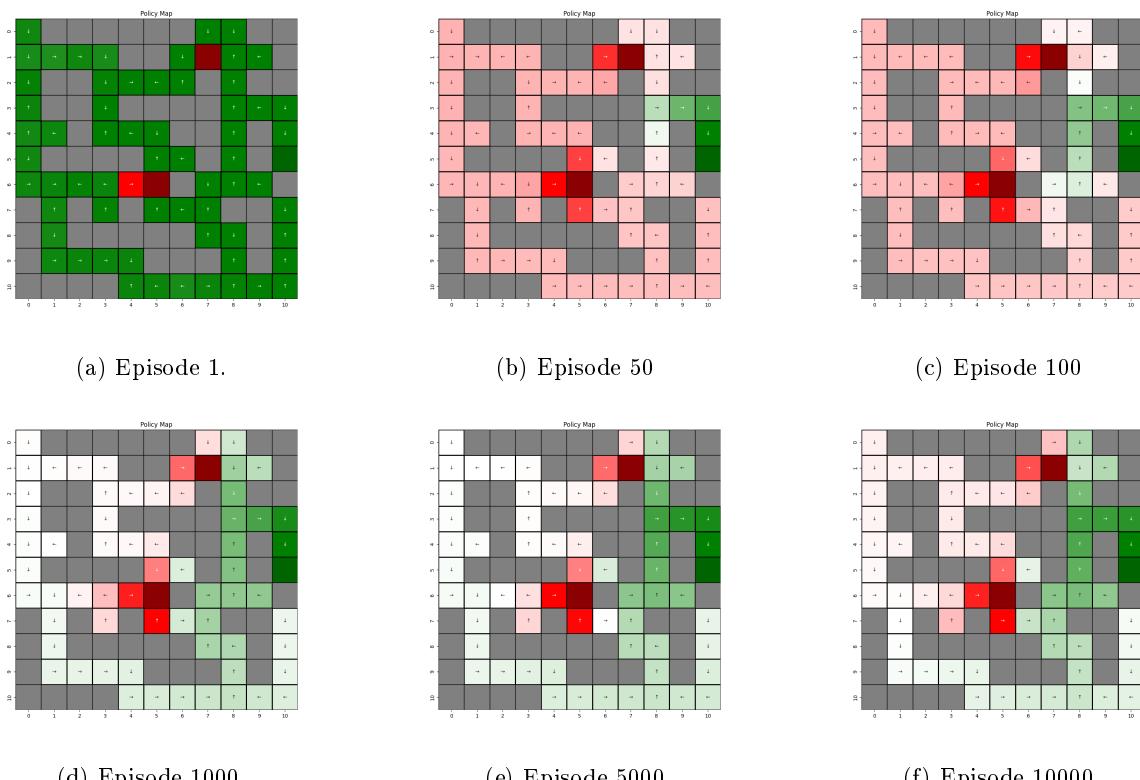


Figure 1: Evolution of policy maps throughout episodes.

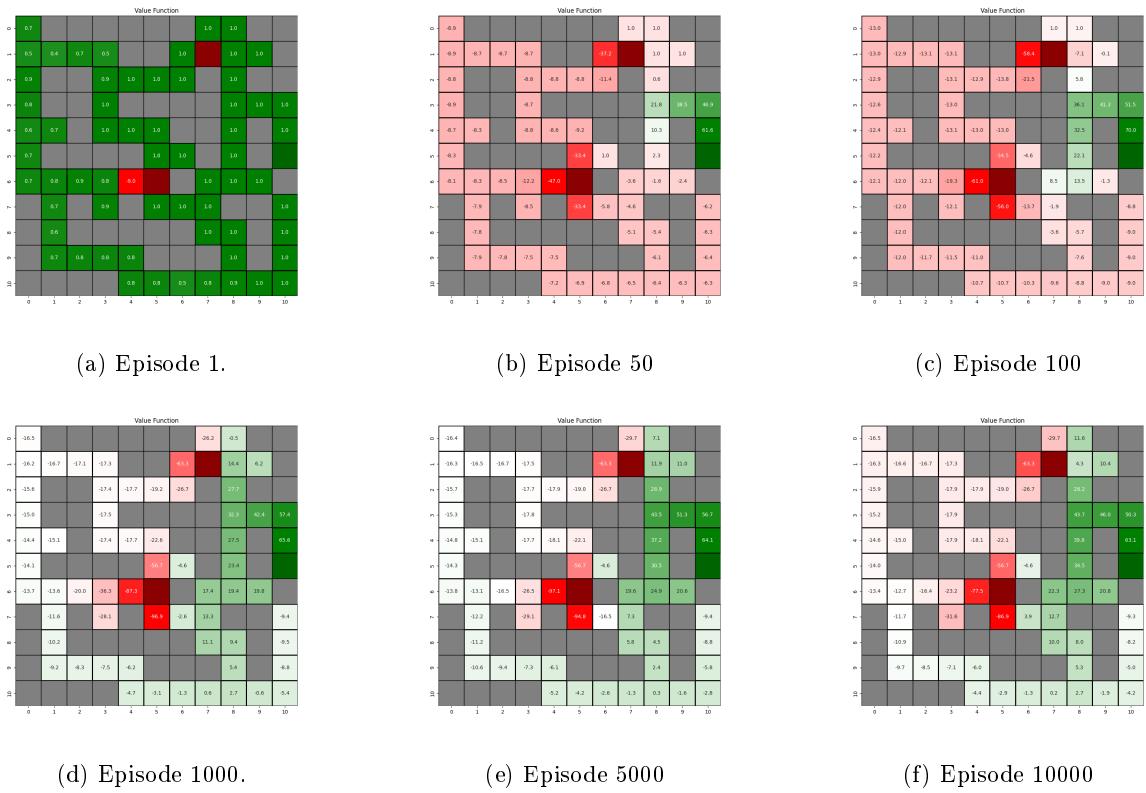


Figure 2: Evolution of value function throughout episodes.

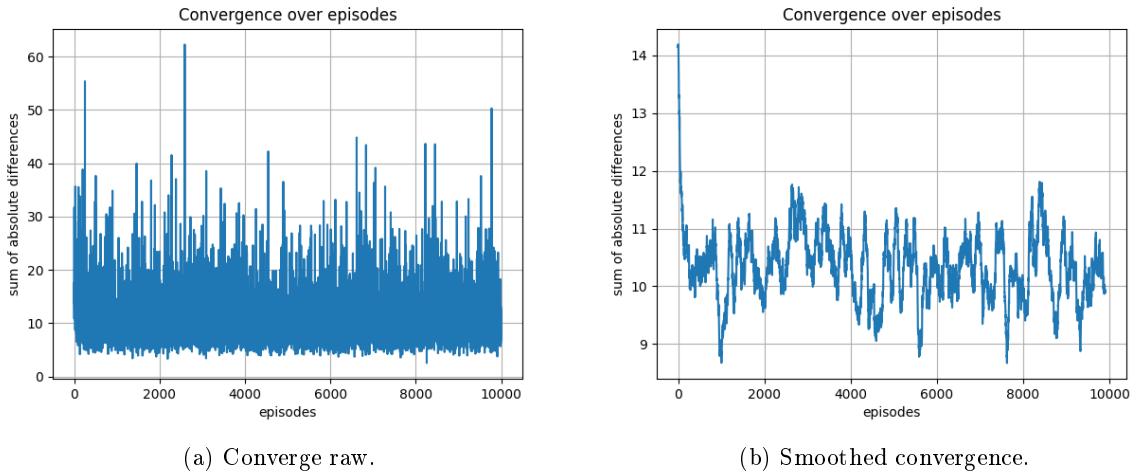


Figure 3: Convergence of value function.

## 2.2 Q-Learning Default Parameters

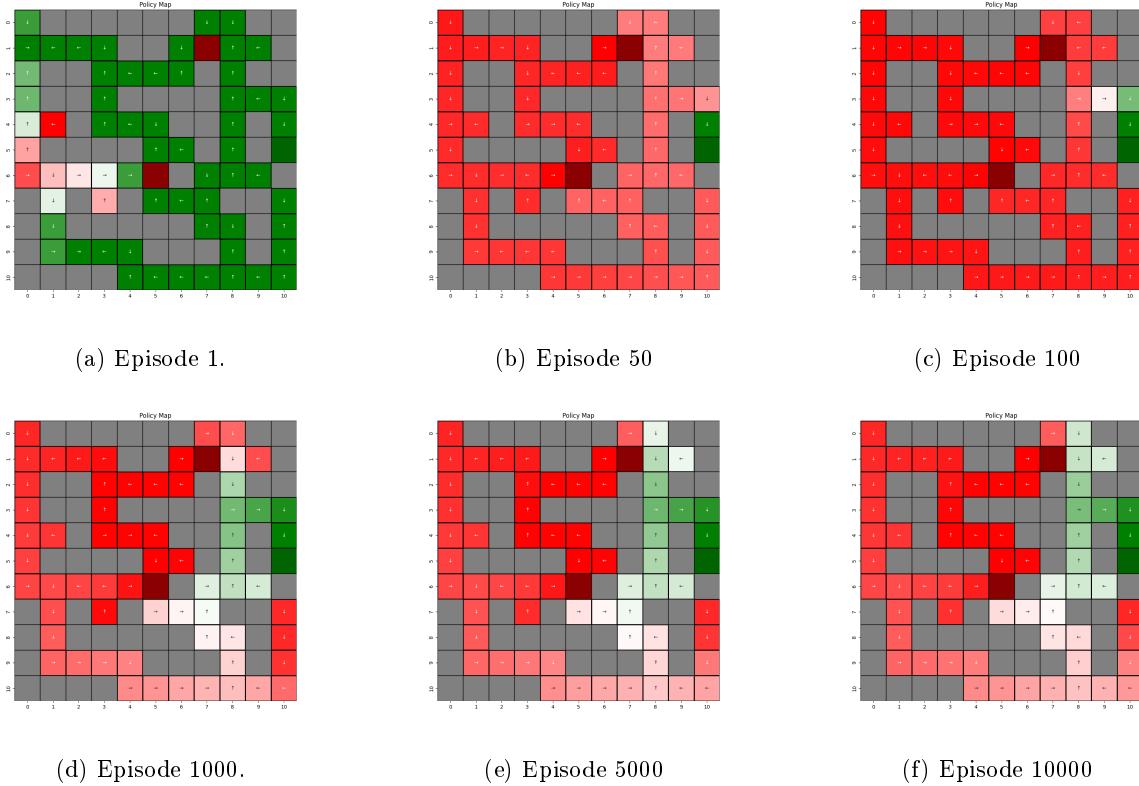


Figure 4: Evolution of policy maps throughout episodes.

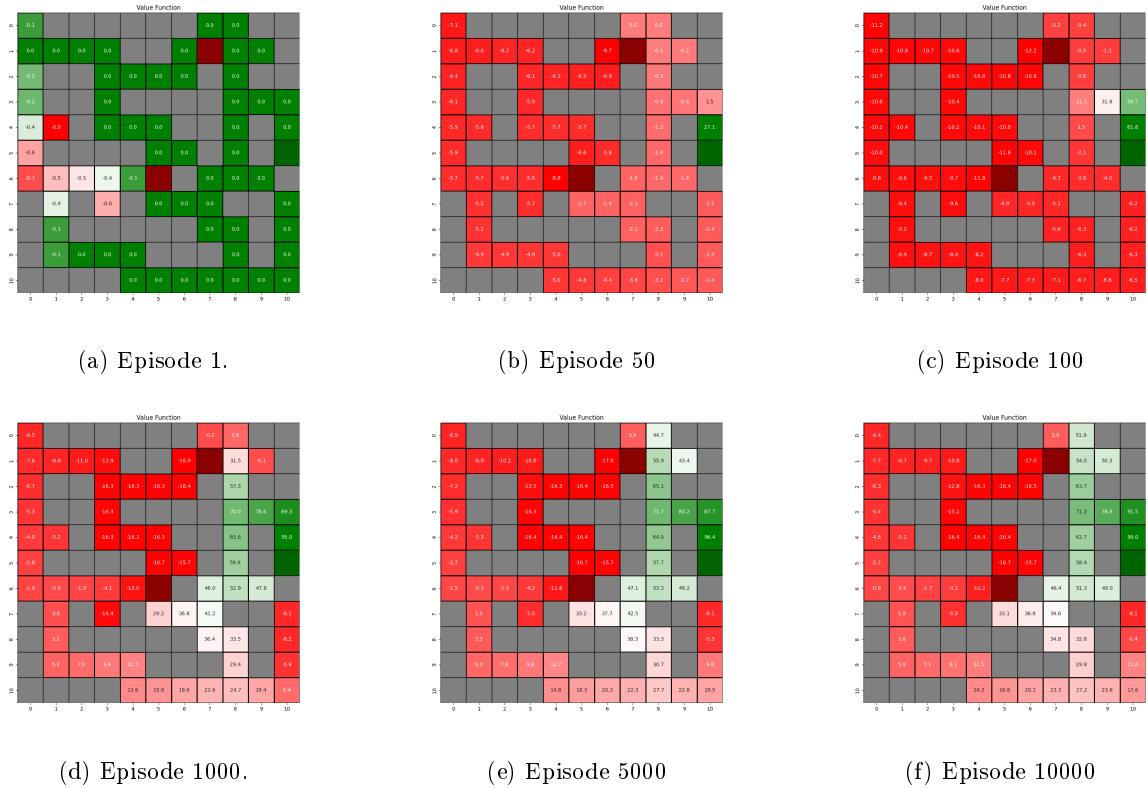


Figure 5: Evolution of value function throughout episodes.

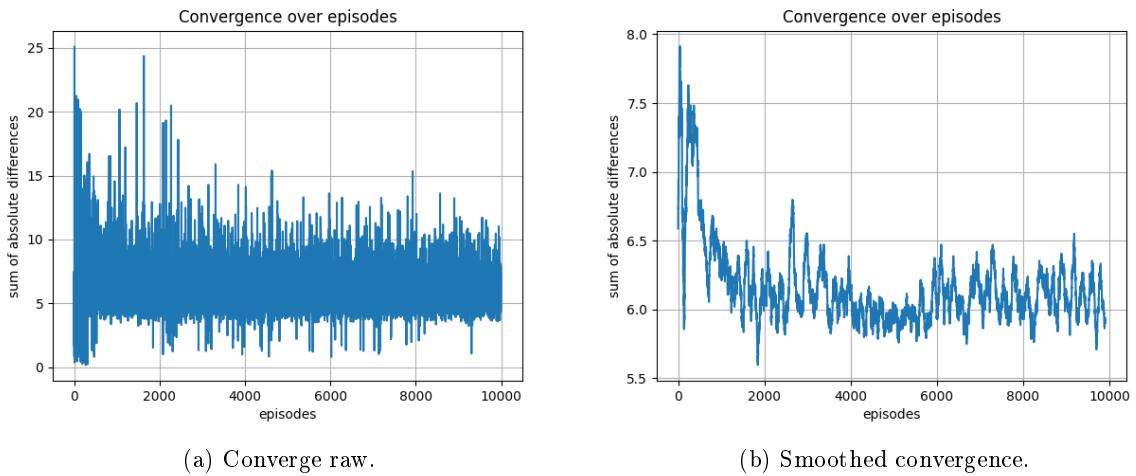


Figure 6: Convergence of value function.

### 2.3 Effect of Alpha in Temporal Difference Learning

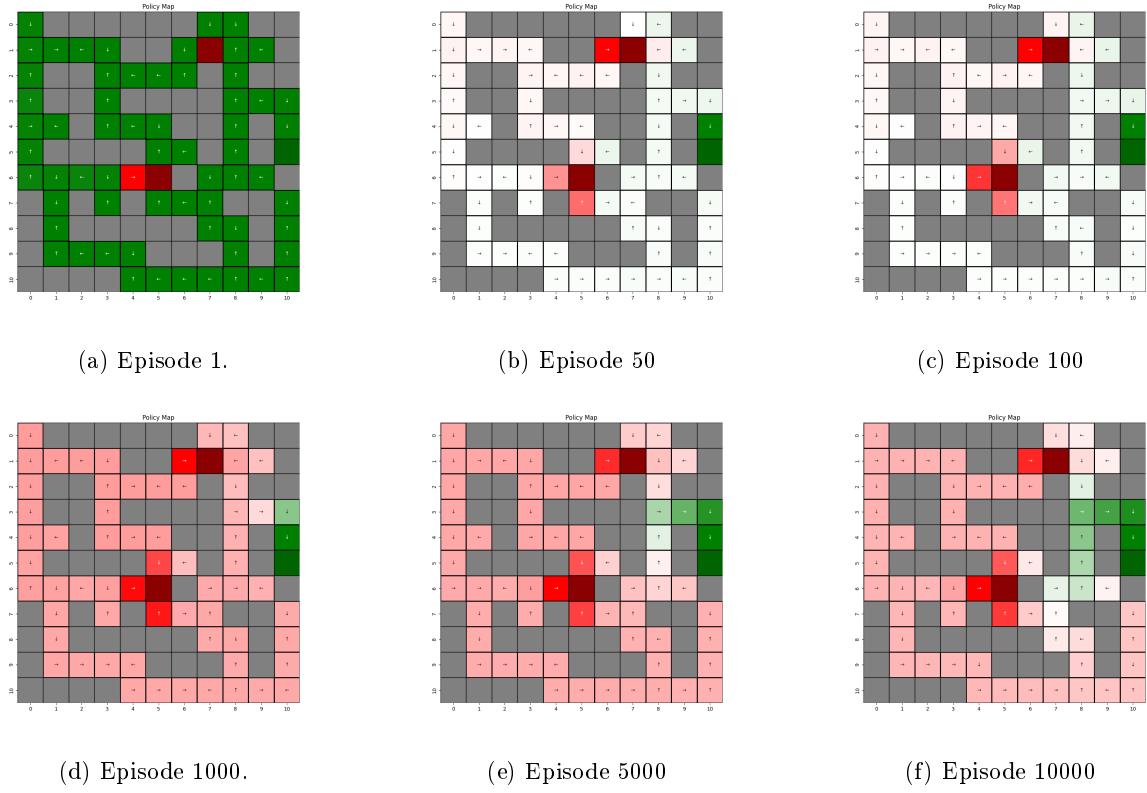


Figure 7: Evolution of policy maps throughout episodes.

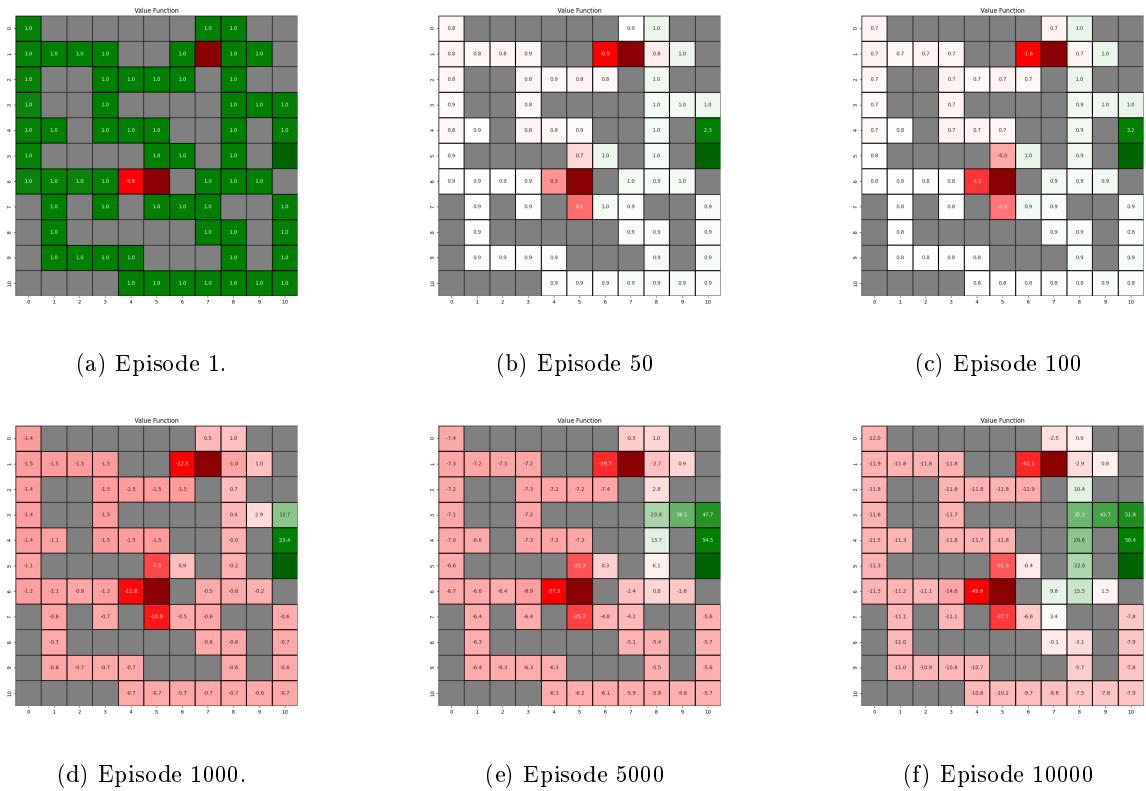


Figure 8: Evolution of value function throughout episodes.

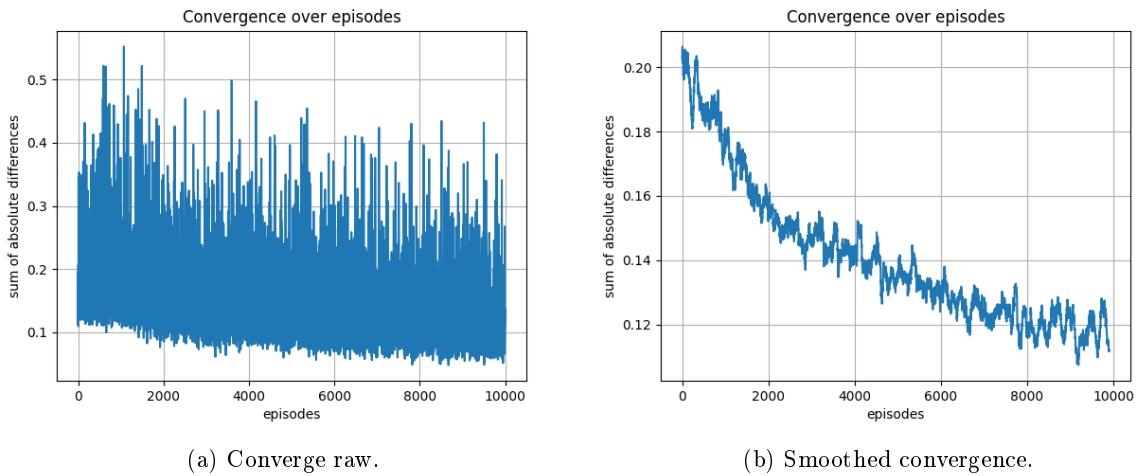


Figure 9: Convergence of value function.

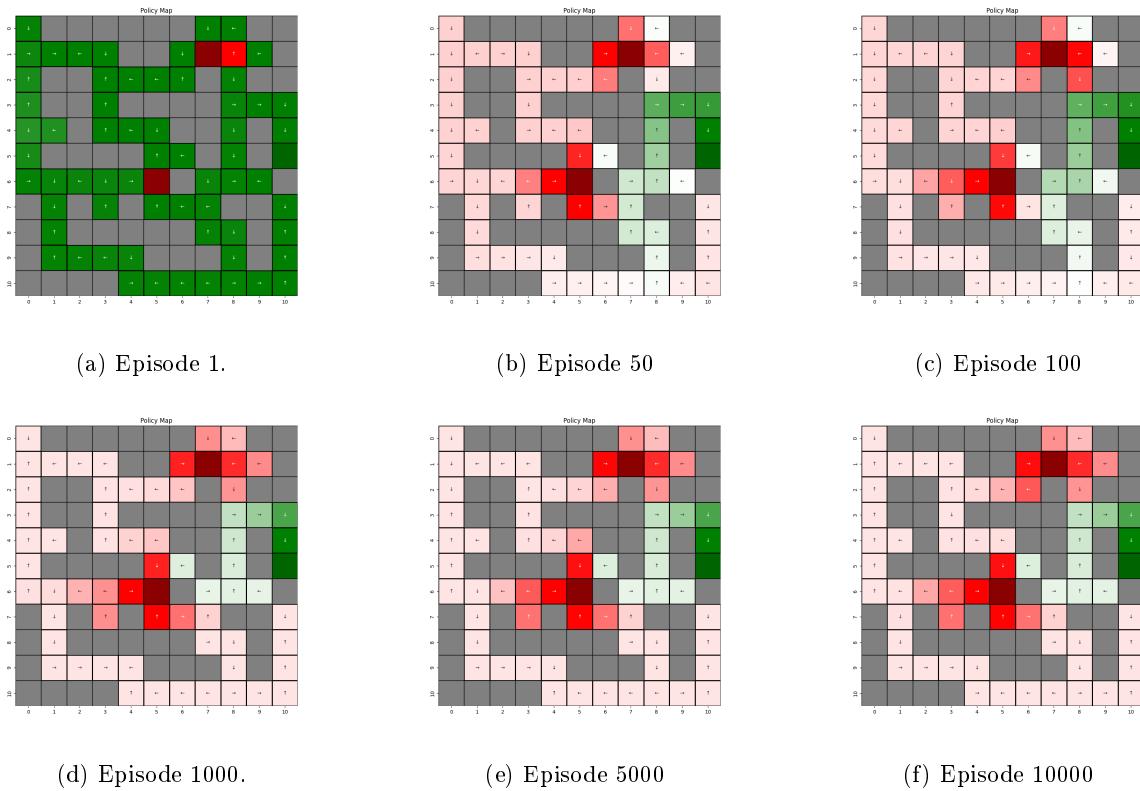


Figure 10: Evolution of policy maps throughout episodes.

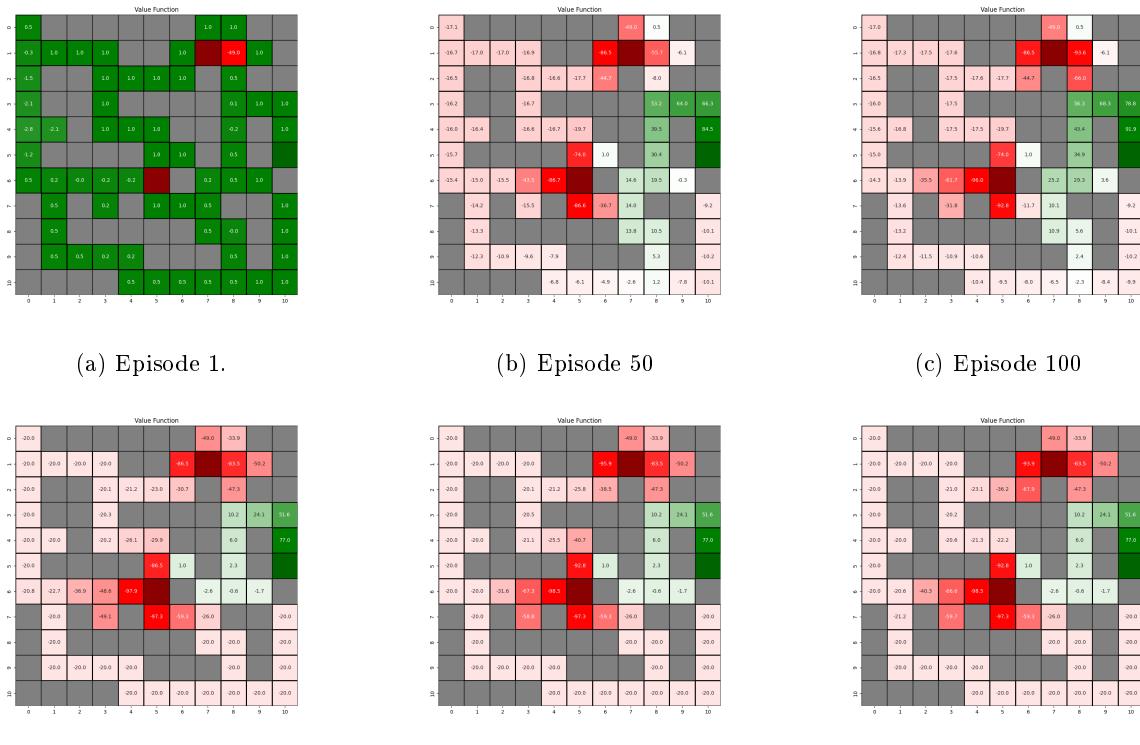


Figure 11: Evolution of value function throughout episodes.

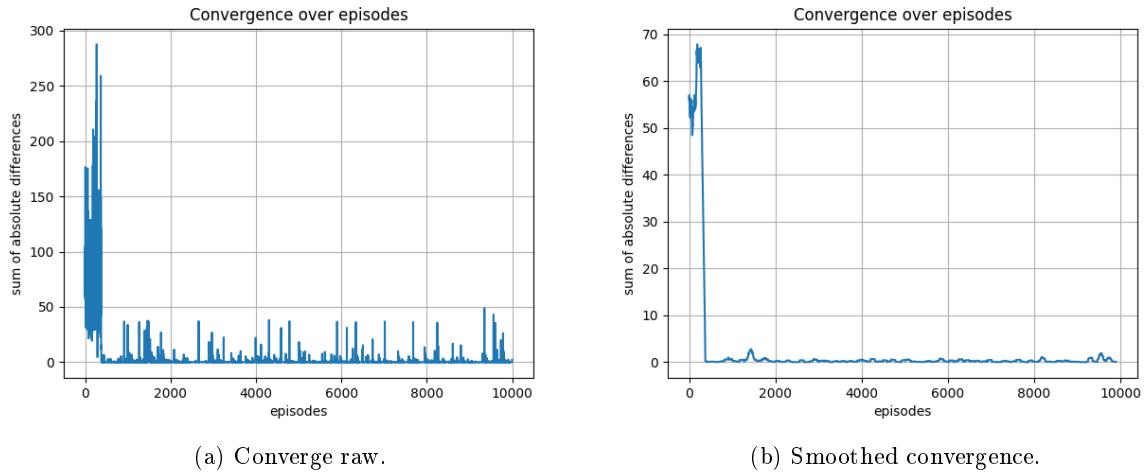


Figure 12: Converge of value function.

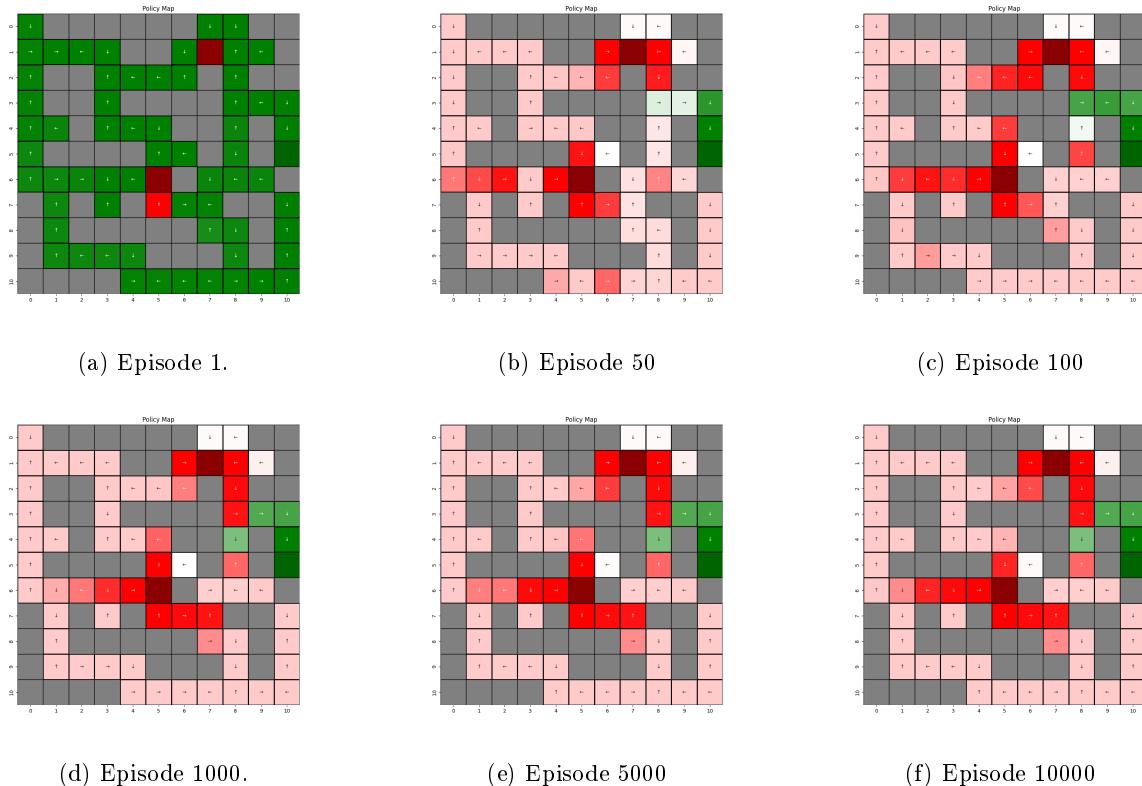


Figure 13: Evolution of policy maps throughout episodes.

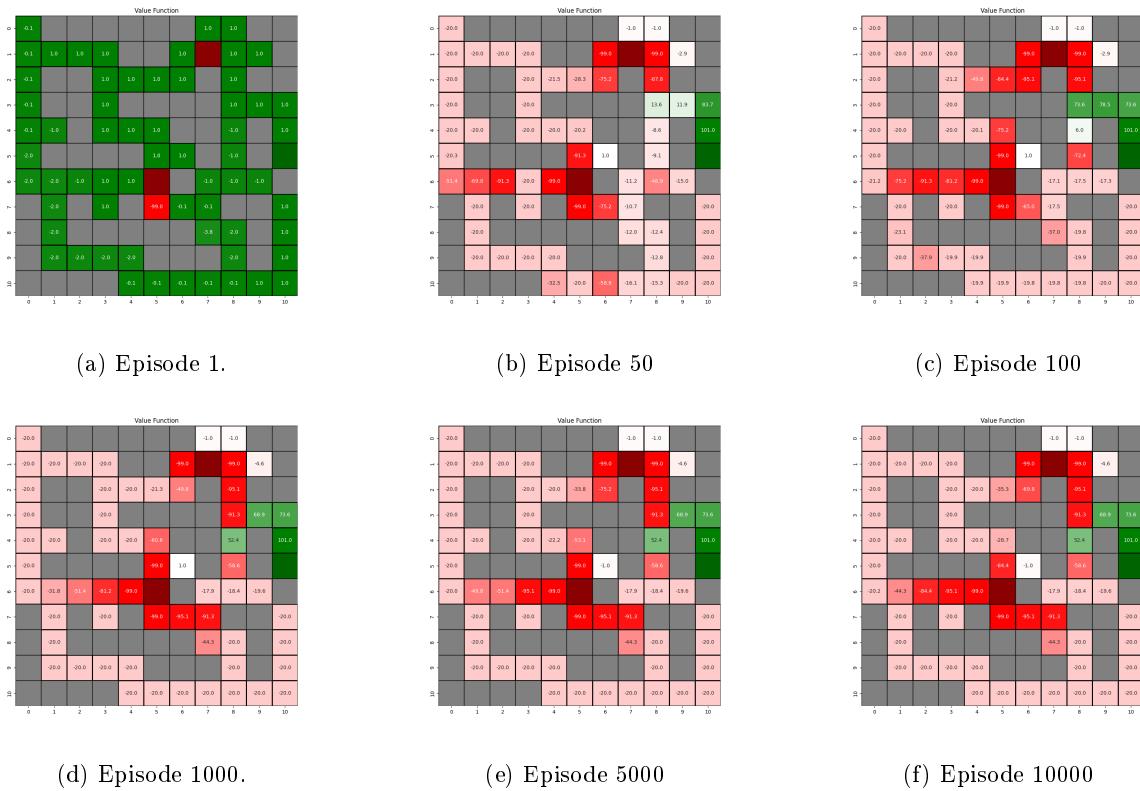


Figure 14: Evolution of value function throughout episodes.

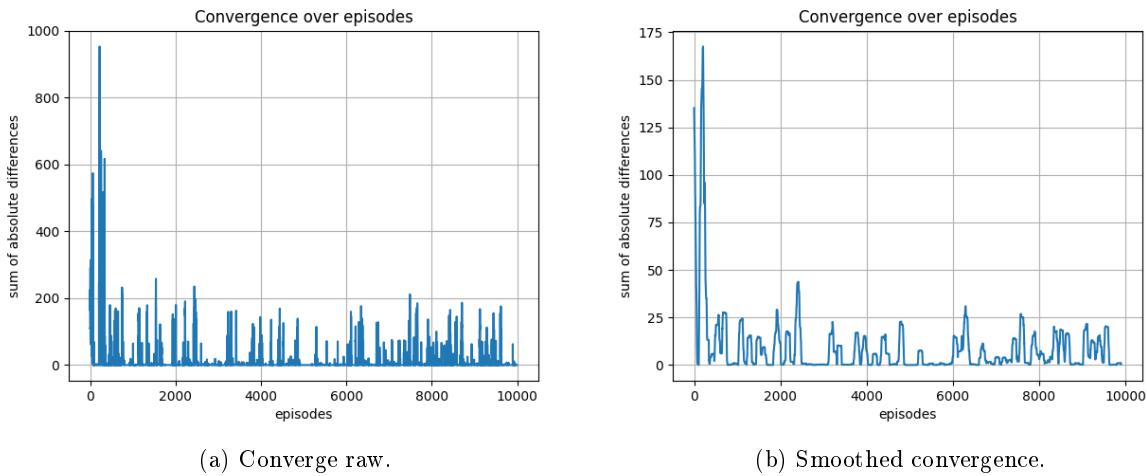


Figure 15: Convergence of value function.

## 2.4 Effect of Alpha in Q-Learning

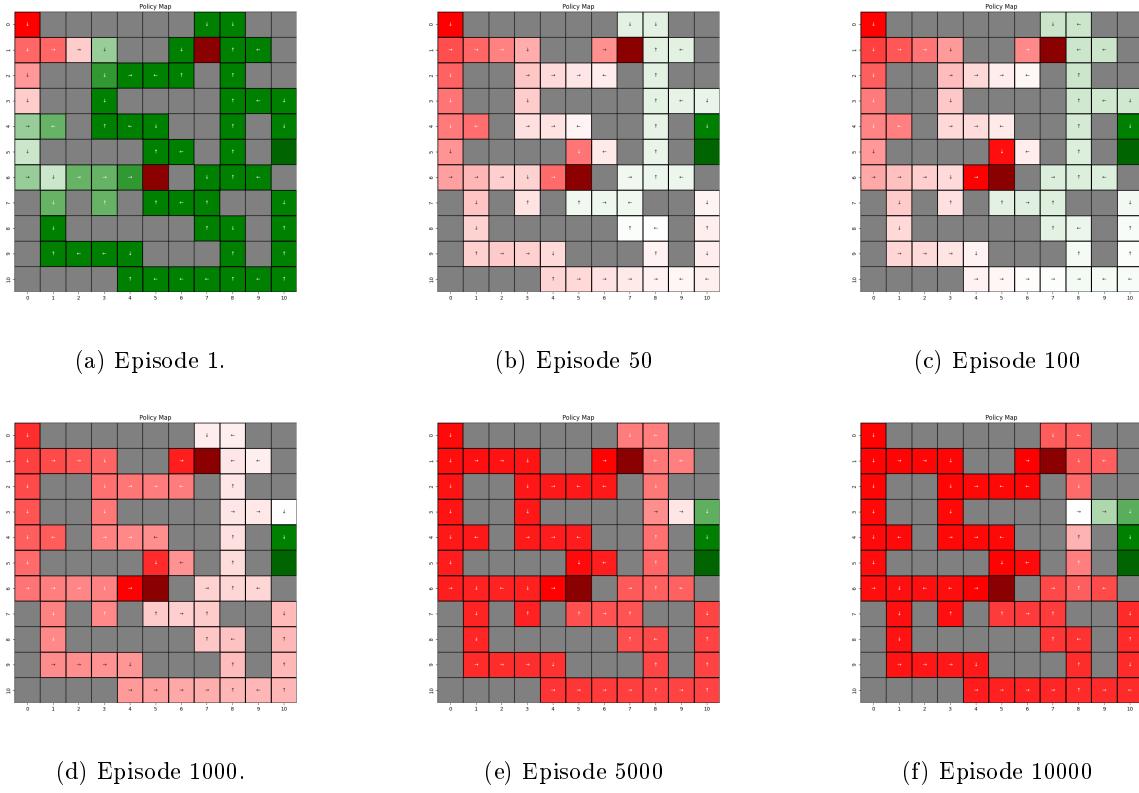


Figure 16: Evolution of policy maps throughout episodes.

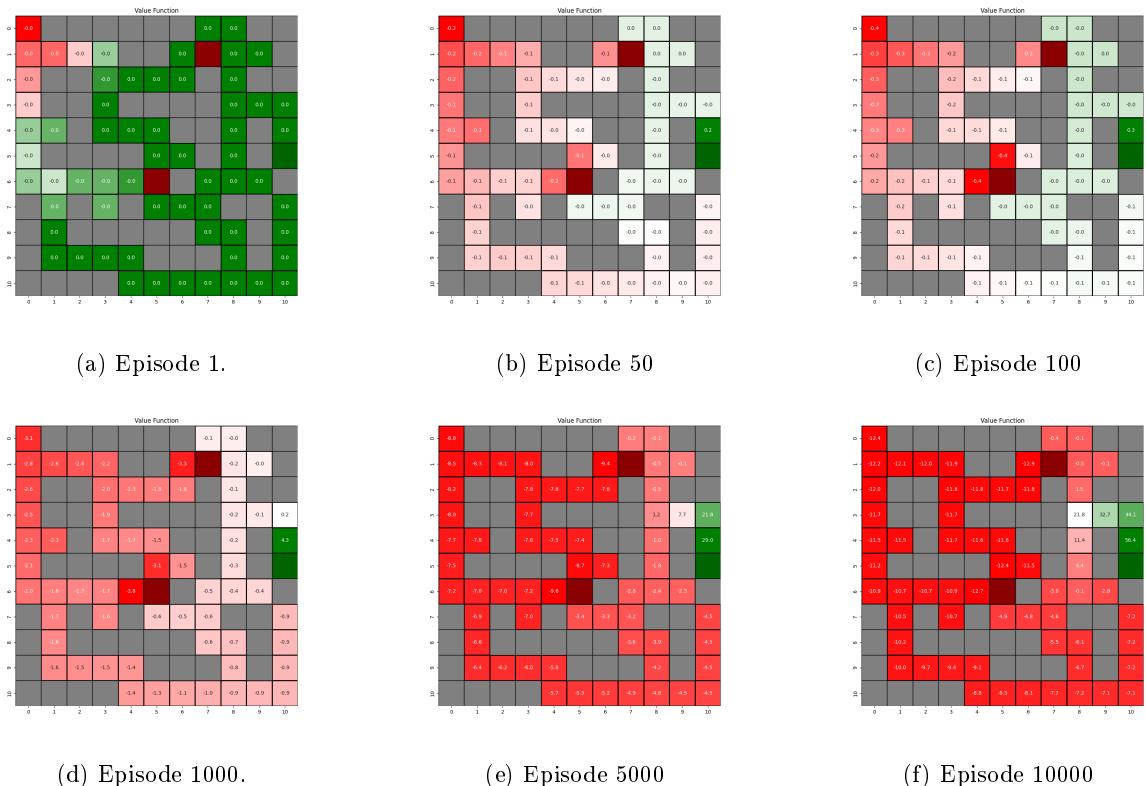


Figure 17: Evolution of value function throughout episodes.

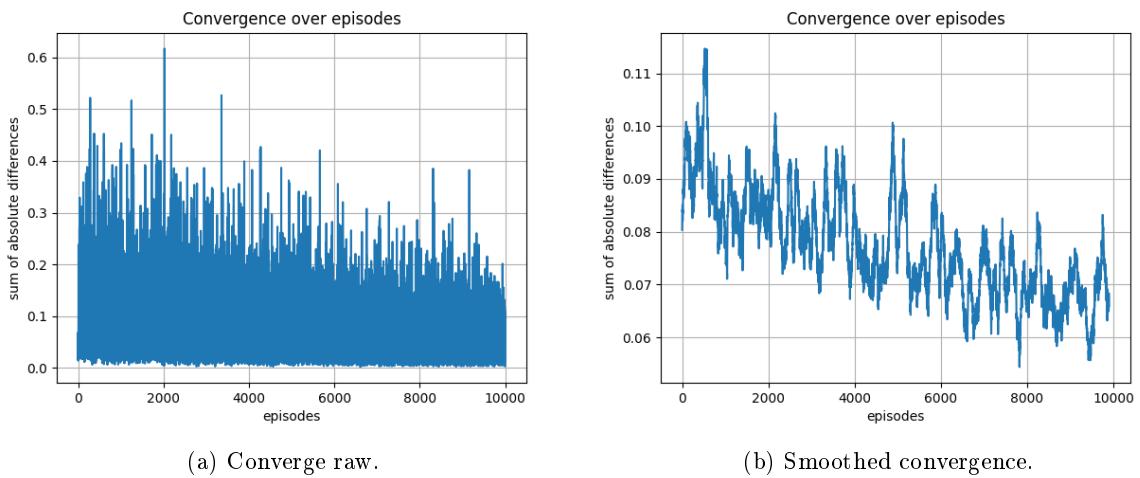


Figure 18: Convergence of value function.

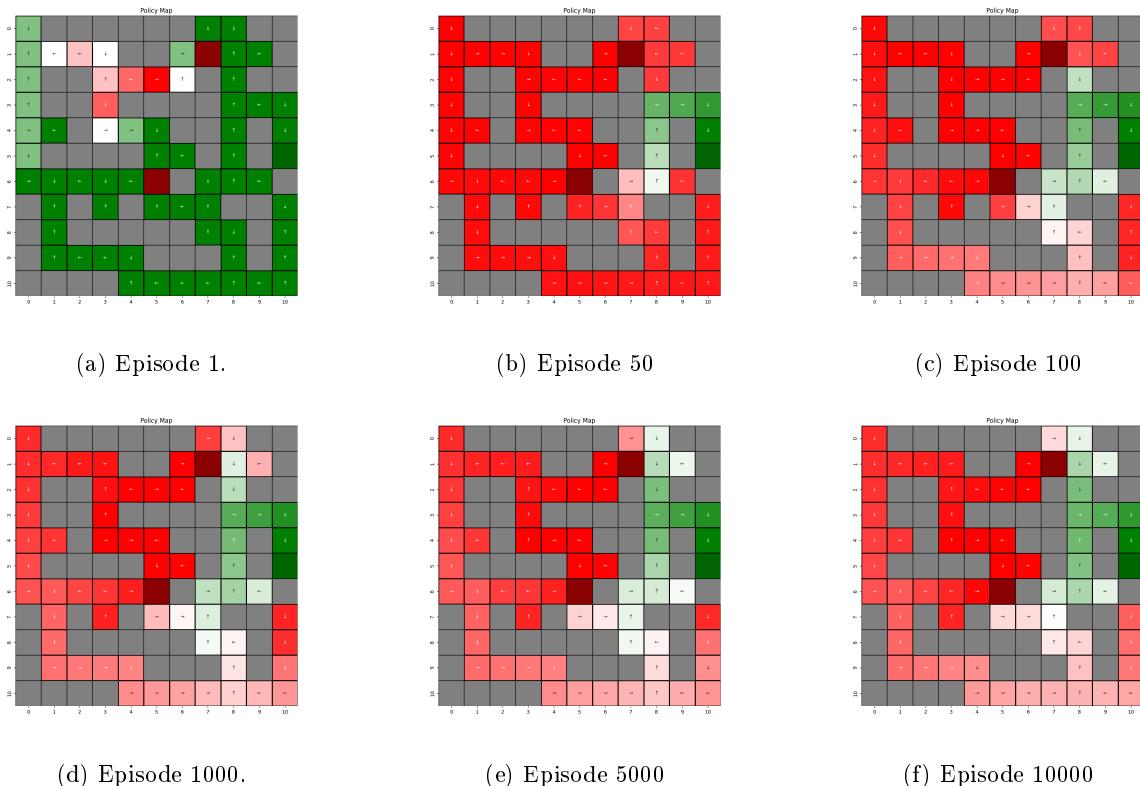


Figure 19: Evolution of policy maps throughout episodes.

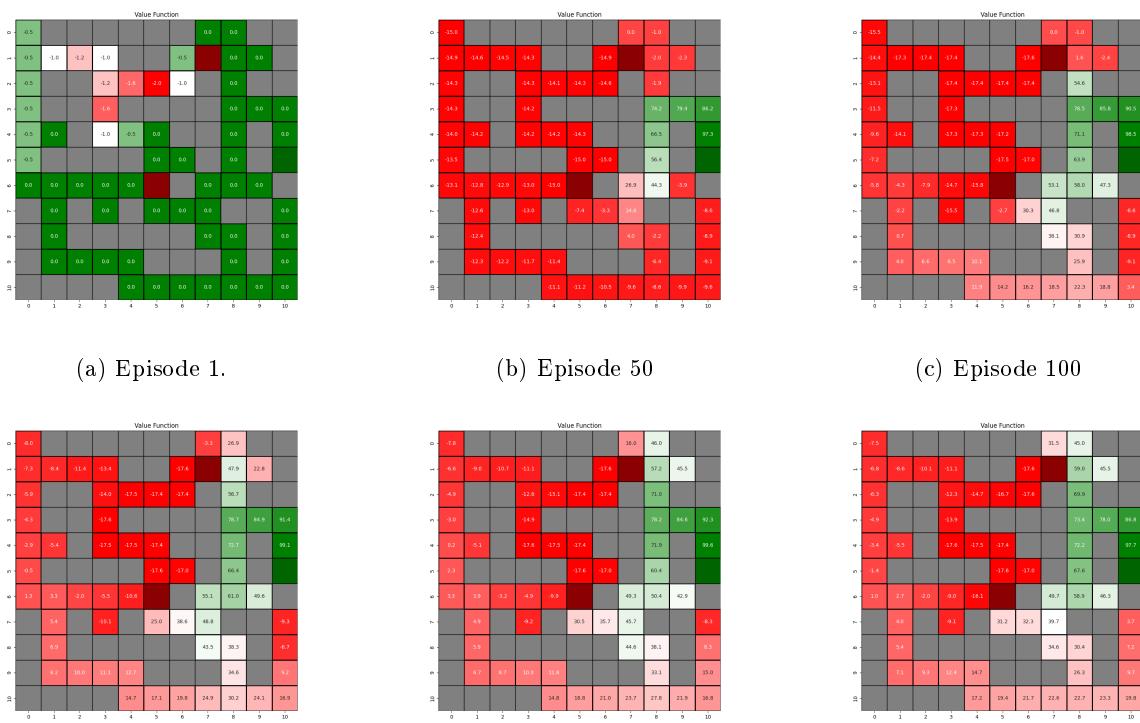


Figure 20: Evolution of value function throughout episodes.

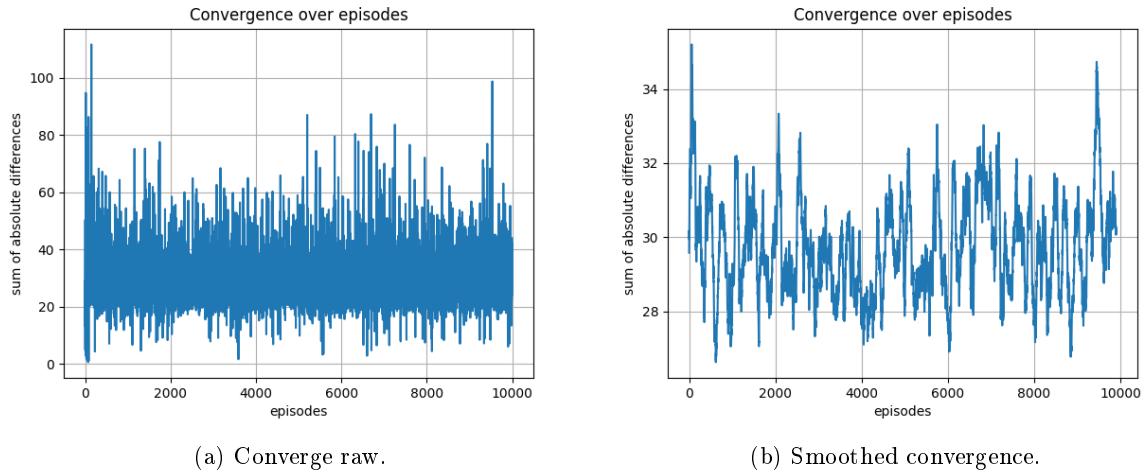


Figure 21: Converge of value function.

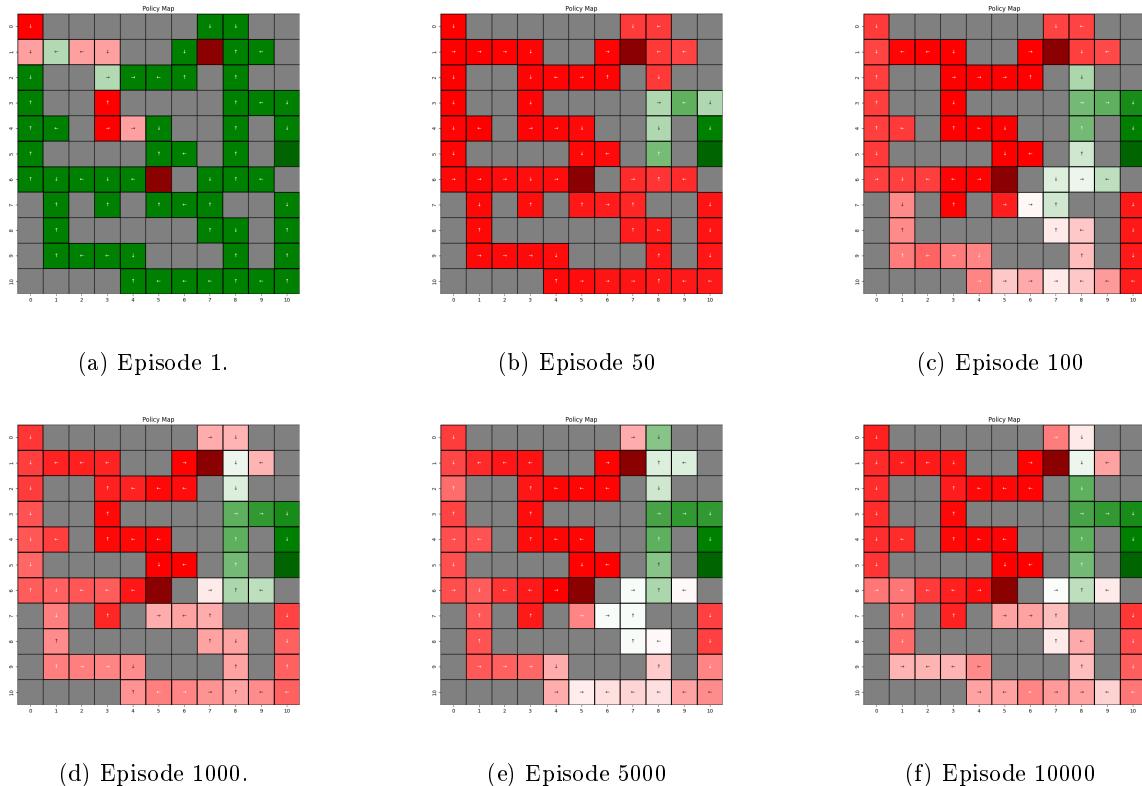


Figure 22: Evolution of policy maps throughout episodes.

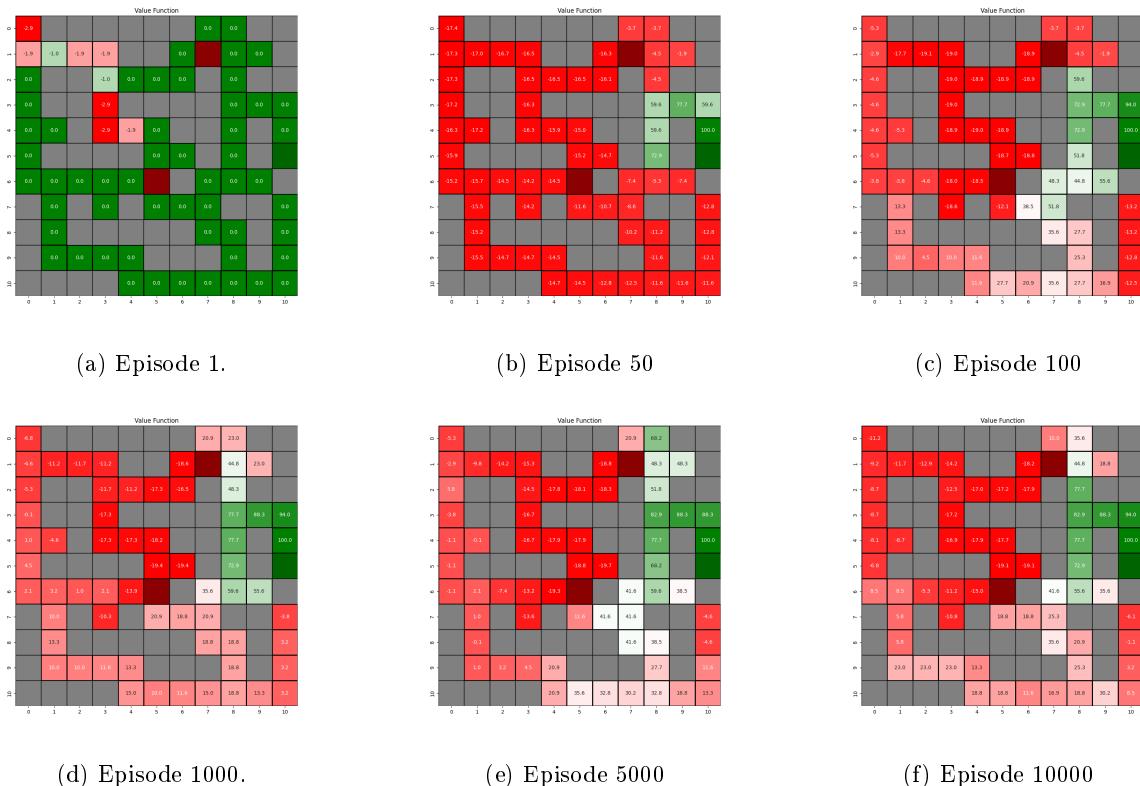


Figure 23: Evolution of value function throughout episodes.

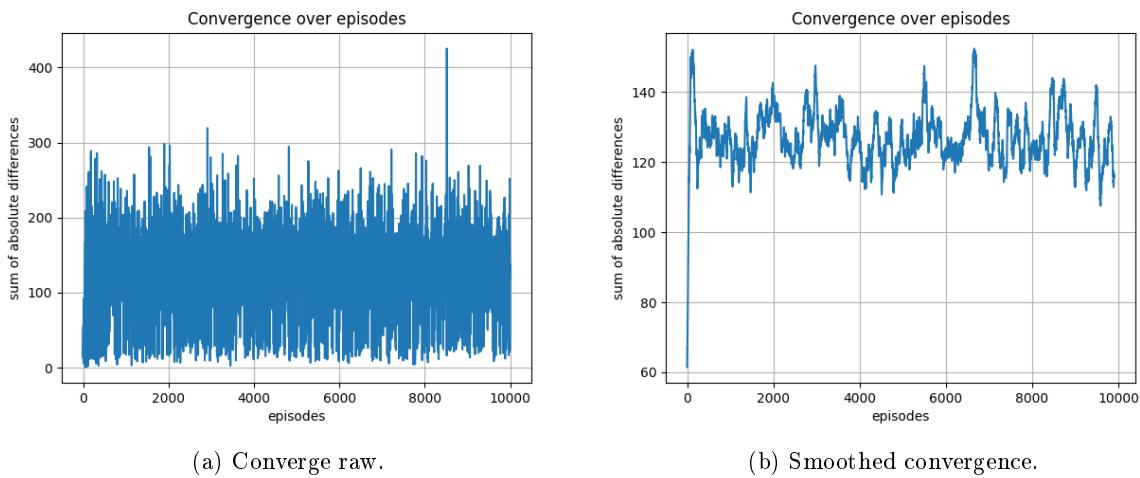


Figure 24: Convergence of value function.

## 2.5 Effect of Gamma in Temporal Difference Learning

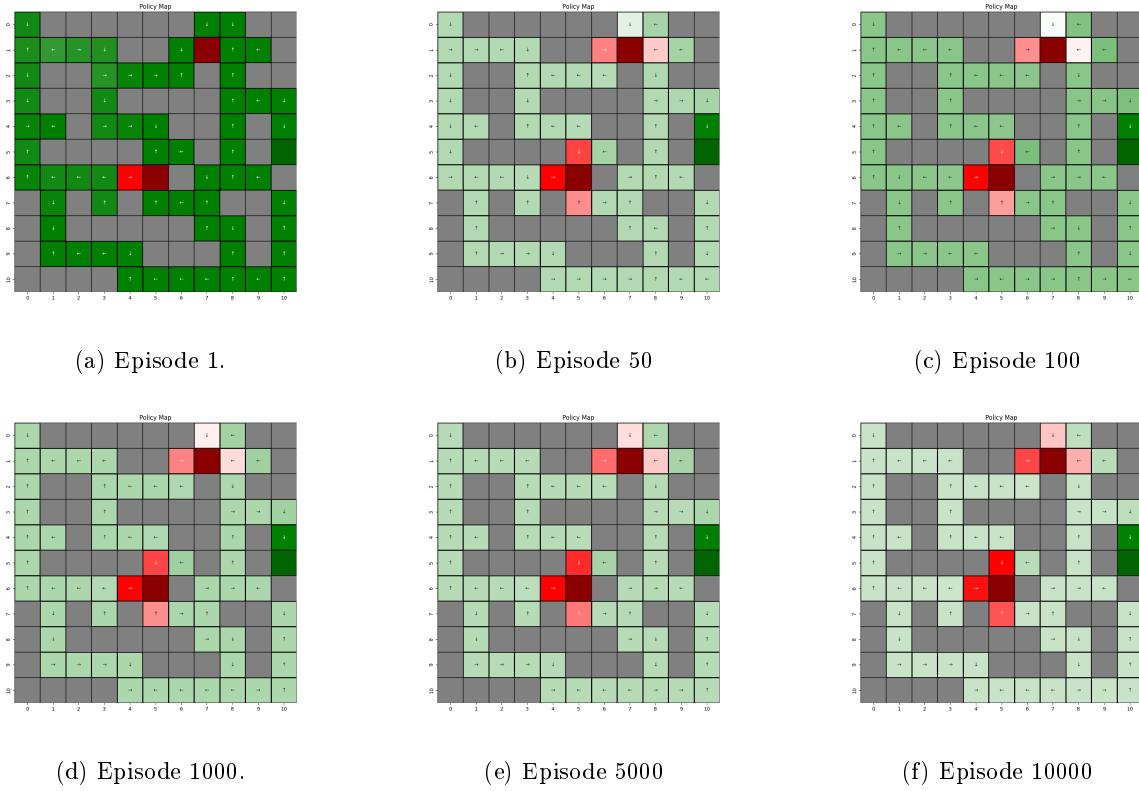


Figure 25: Evolution of policy maps throughout episodes.

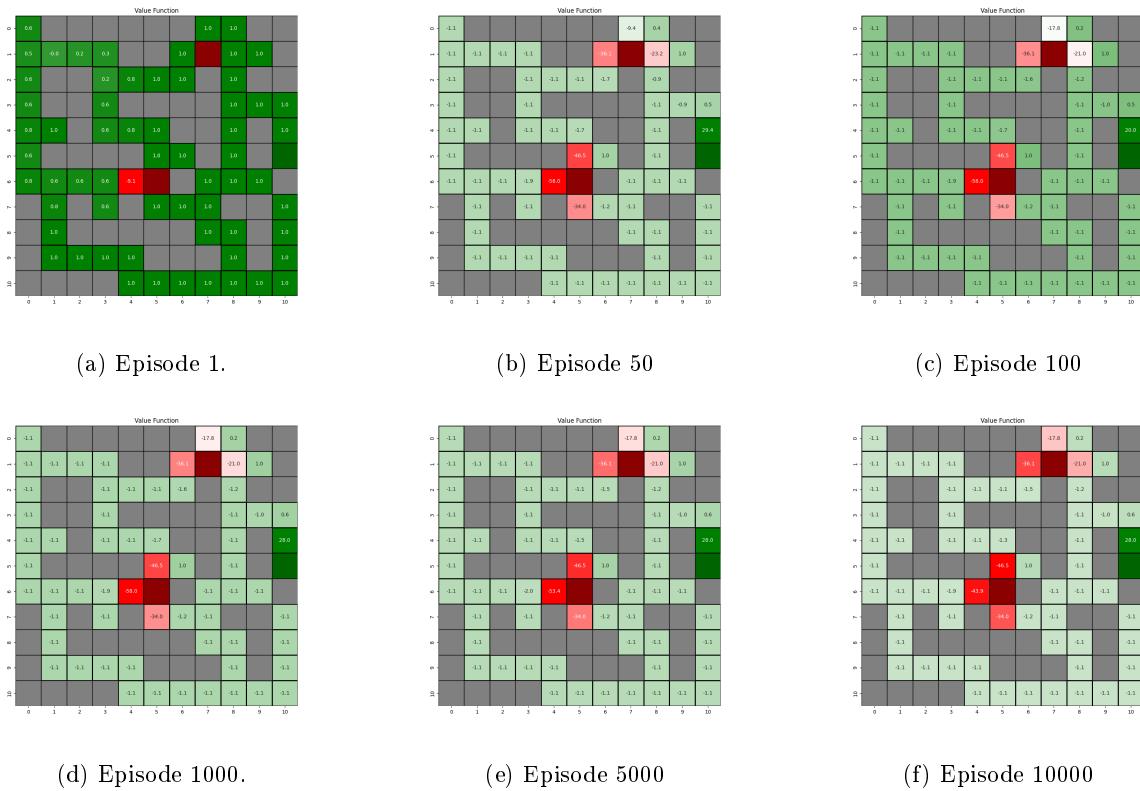


Figure 26: Evolution of value function throughout episodes.

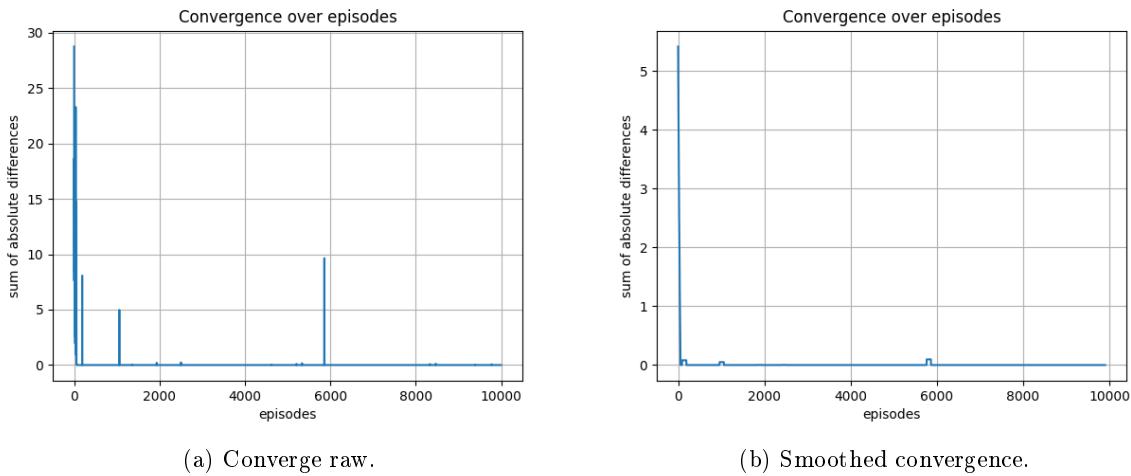


Figure 27: Convergence of value function.

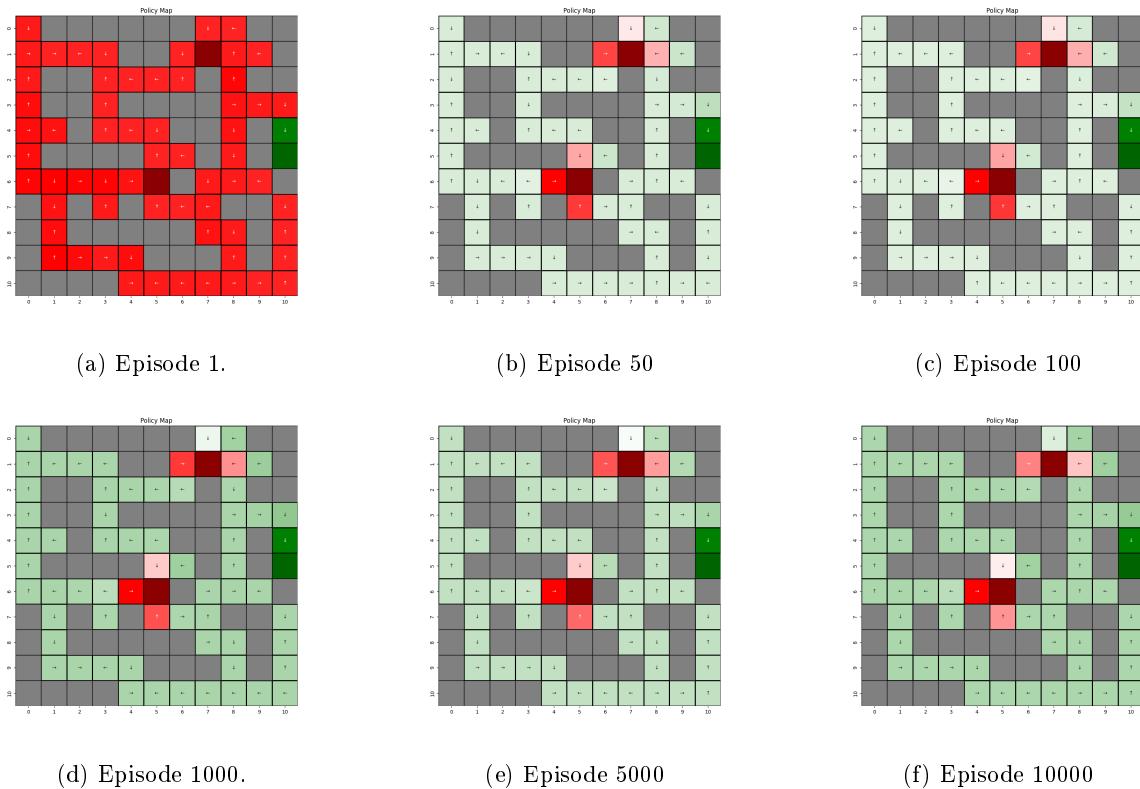


Figure 28: Evolution of policy maps throughout episodes.

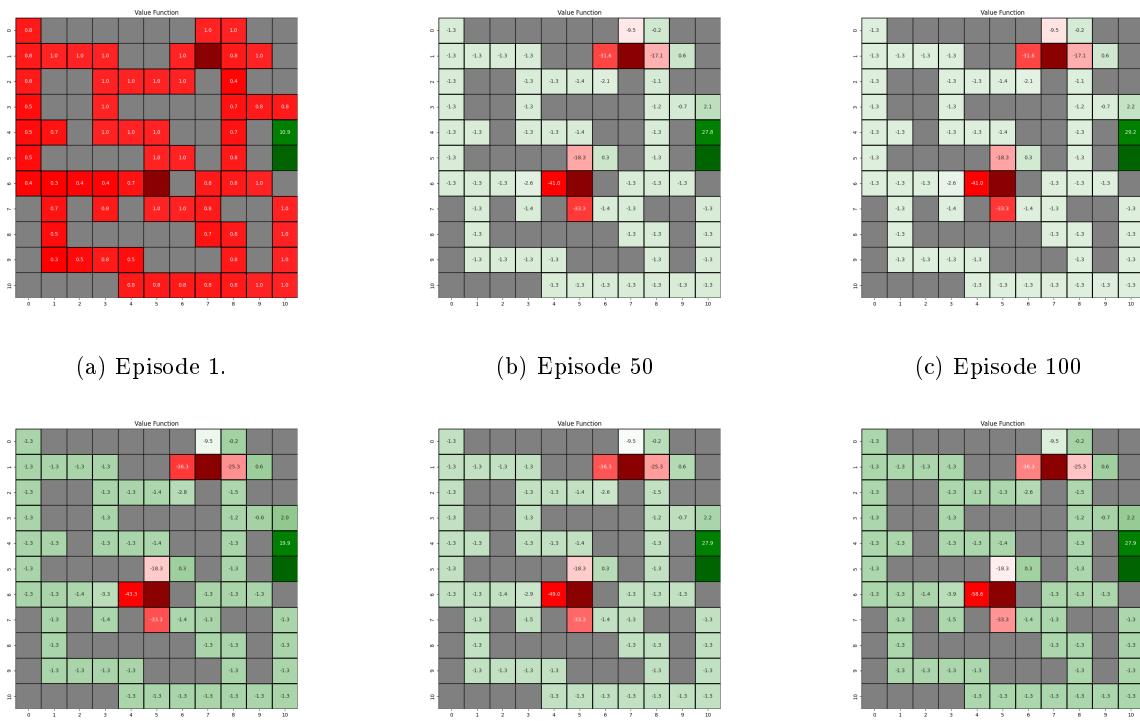


Figure 29: Evolution of value function throughout episodes.

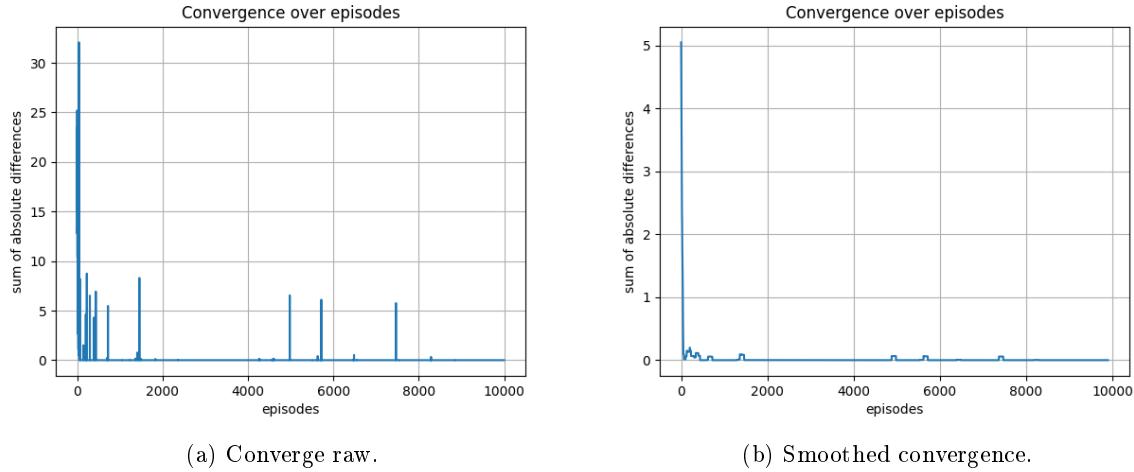


Figure 30: Converge of value function.

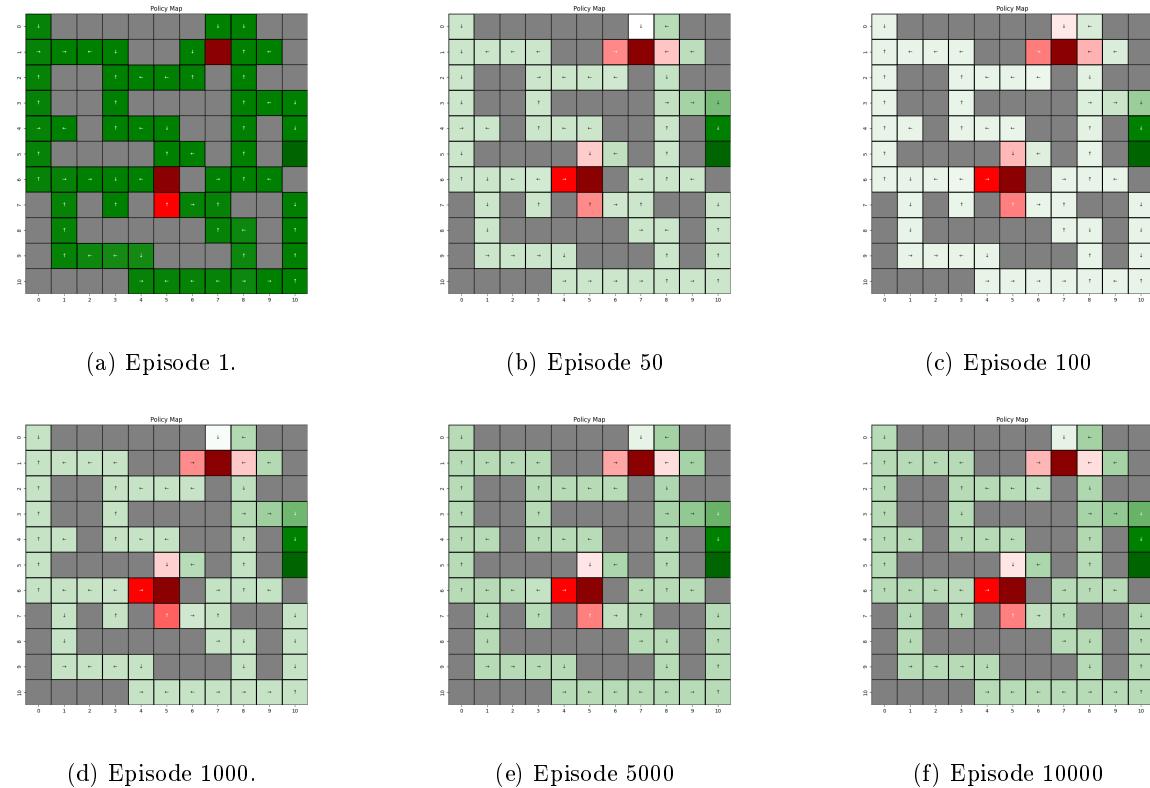


Figure 31: Evolution of policy maps throughout episodes.

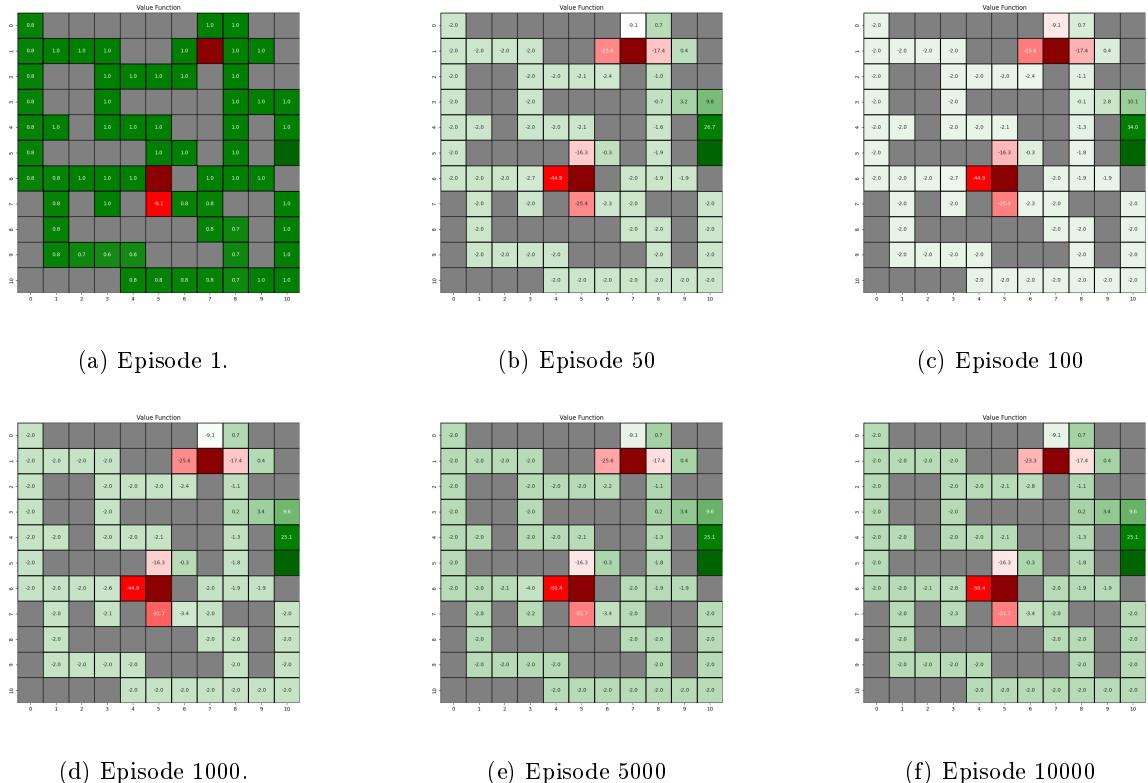


Figure 32: Evolution of value function throughout episodes.

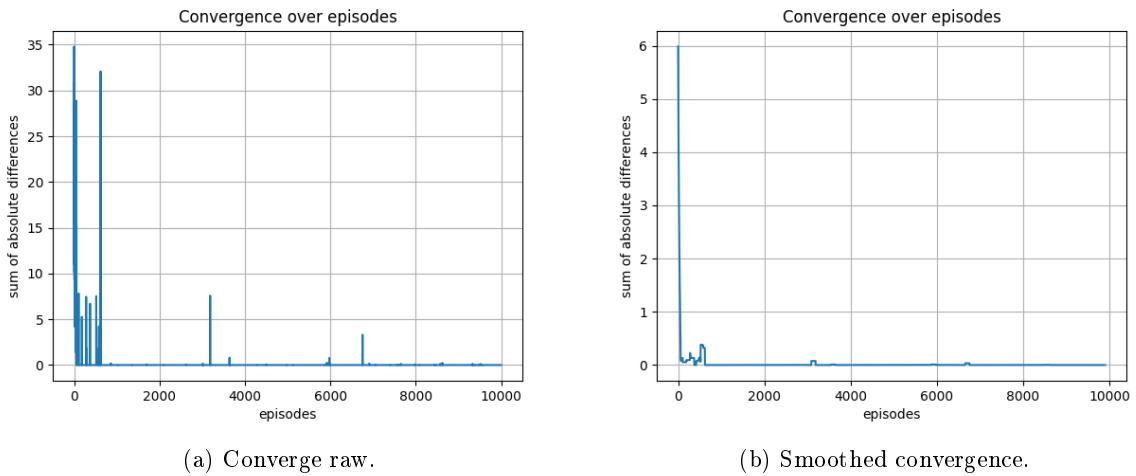


Figure 33: Convergence of value function.

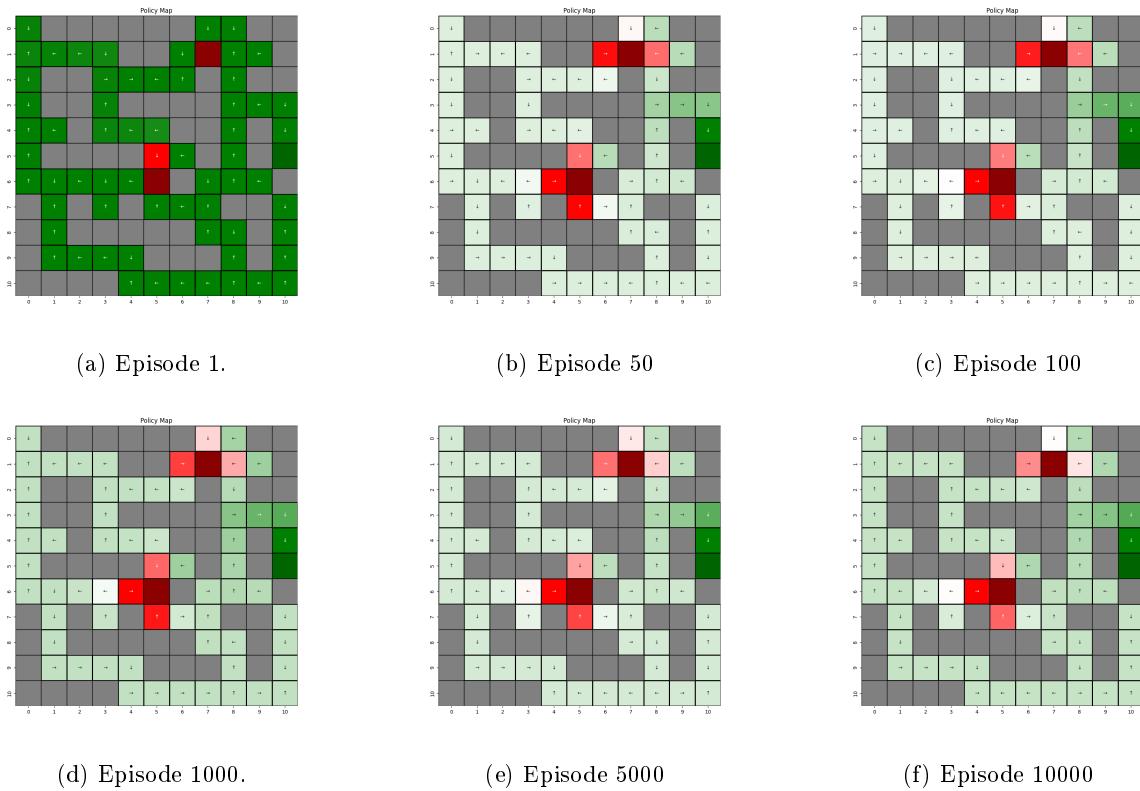


Figure 34: Evolution of policy maps throughout episodes.

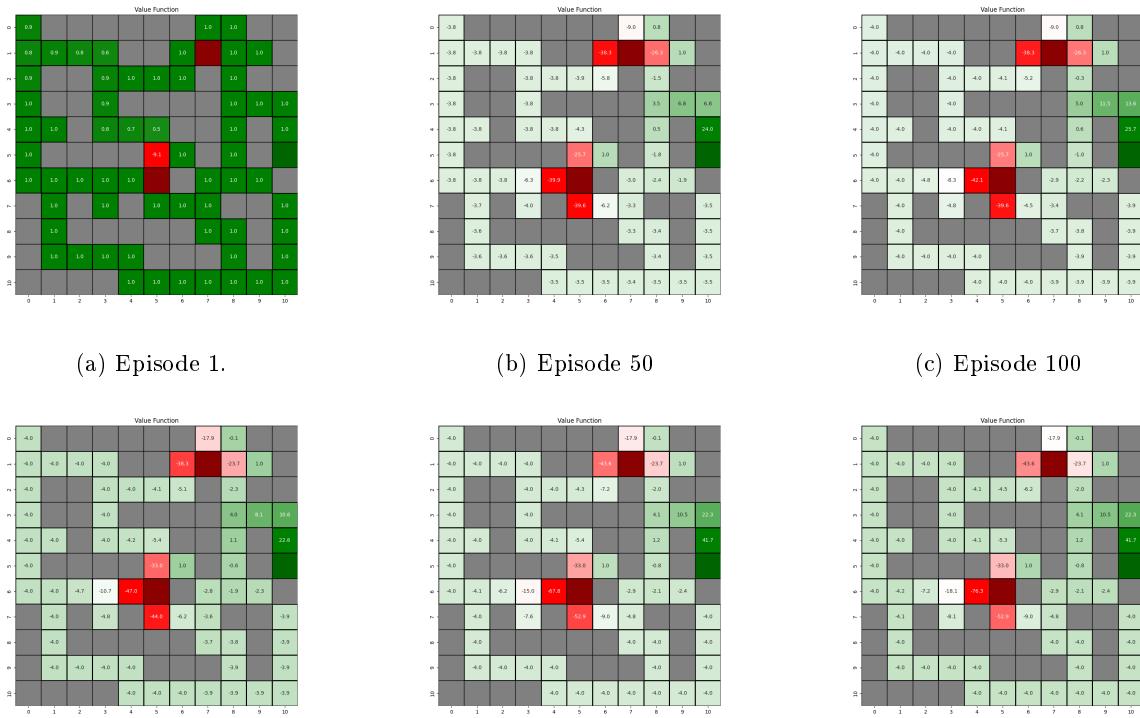


Figure 35: Evolution of value function throughout episodes.

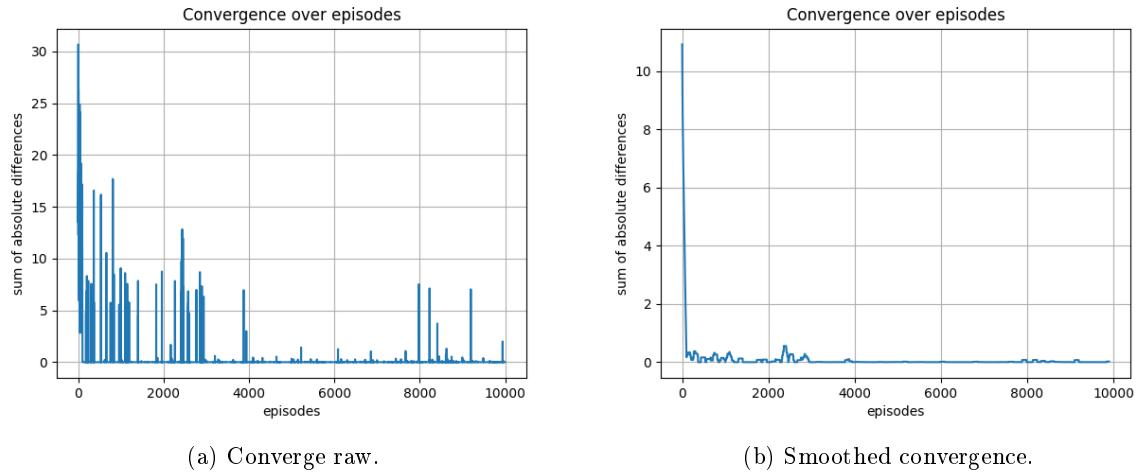


Figure 36: Converge of value function.

## 2.6 Effect of Gamma in Q-Learning

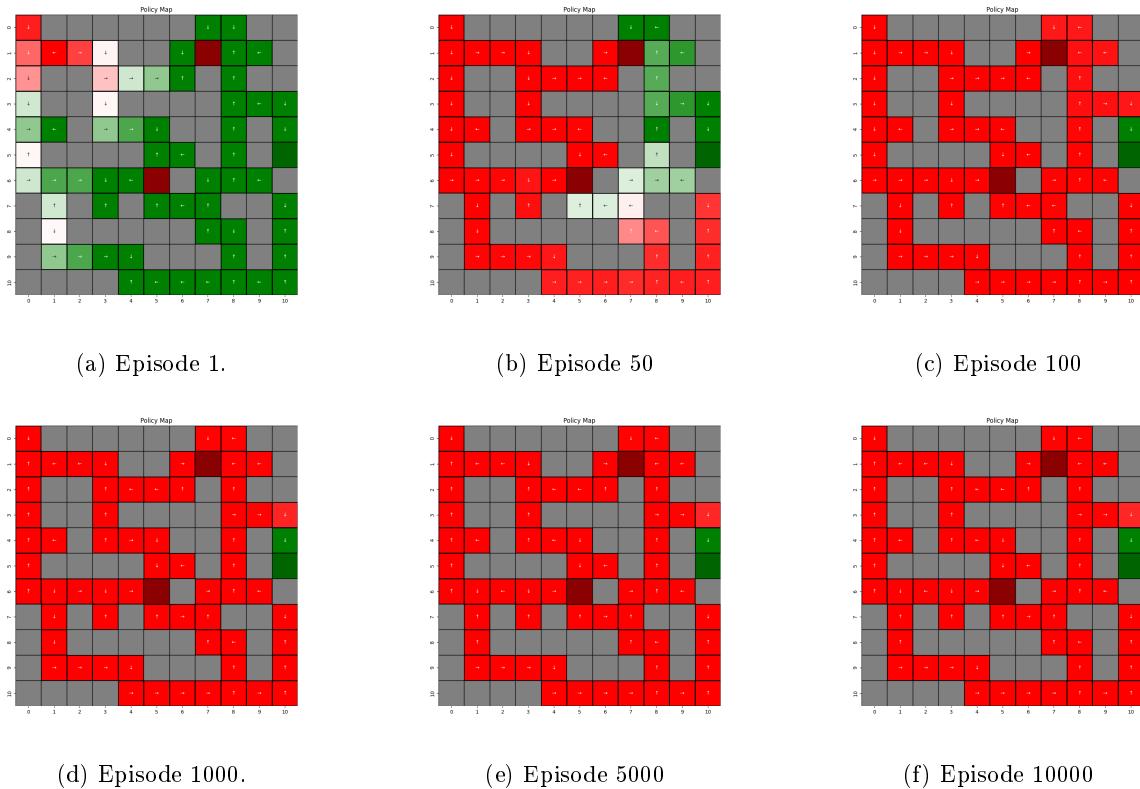


Figure 37: Evolution of policy maps throughout episodes.

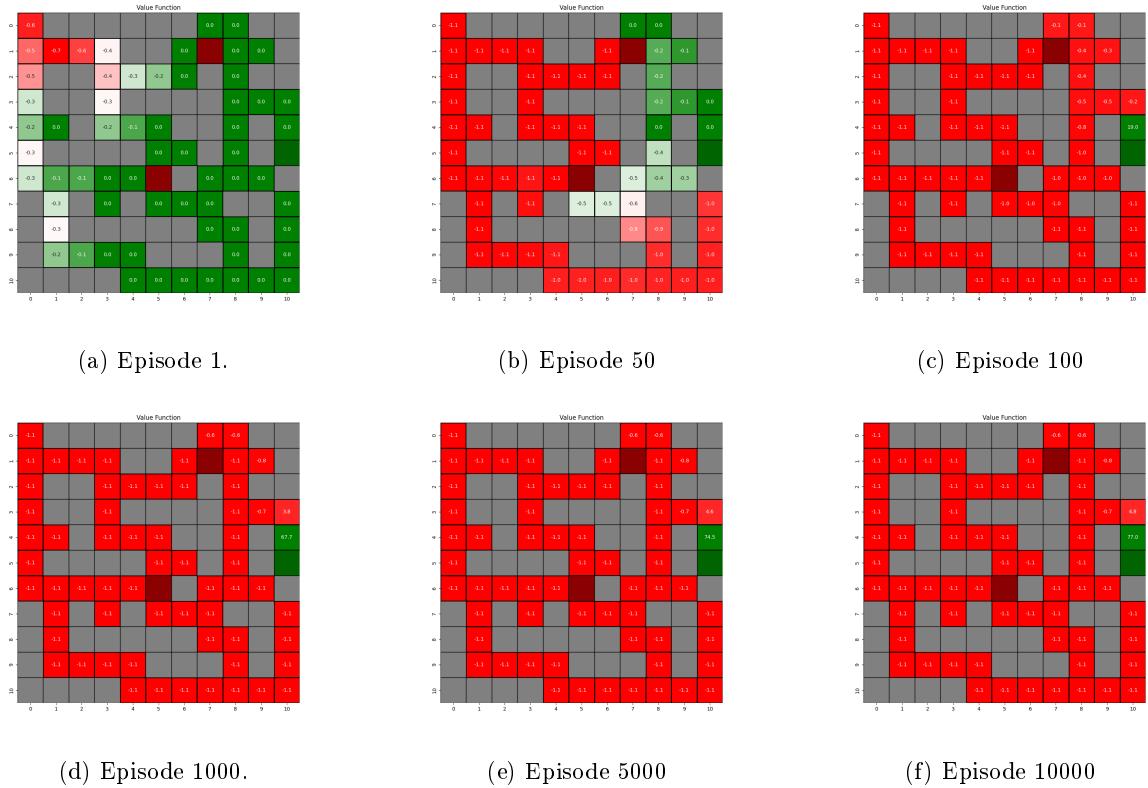


Figure 38: Evolution of value function throughout episodes.

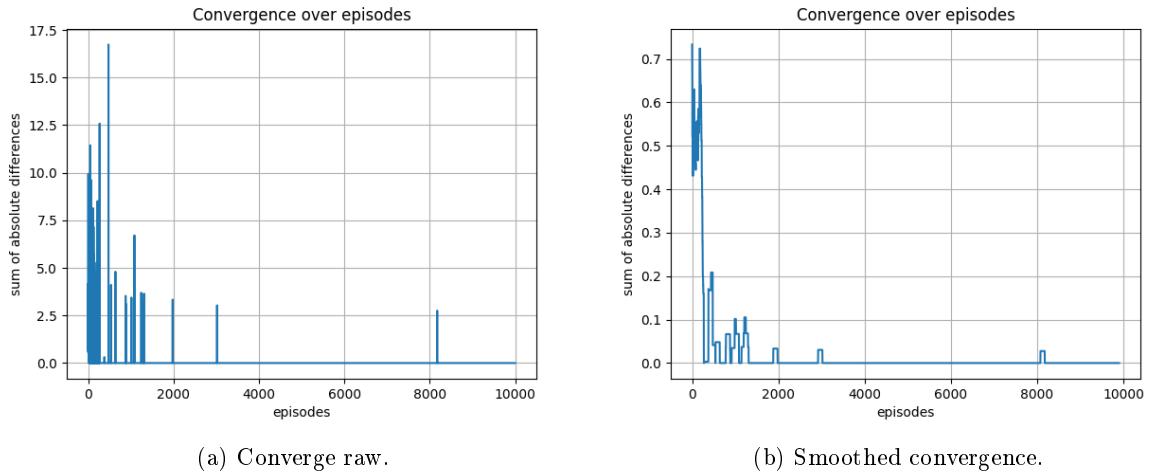


Figure 39: Converge of value function.

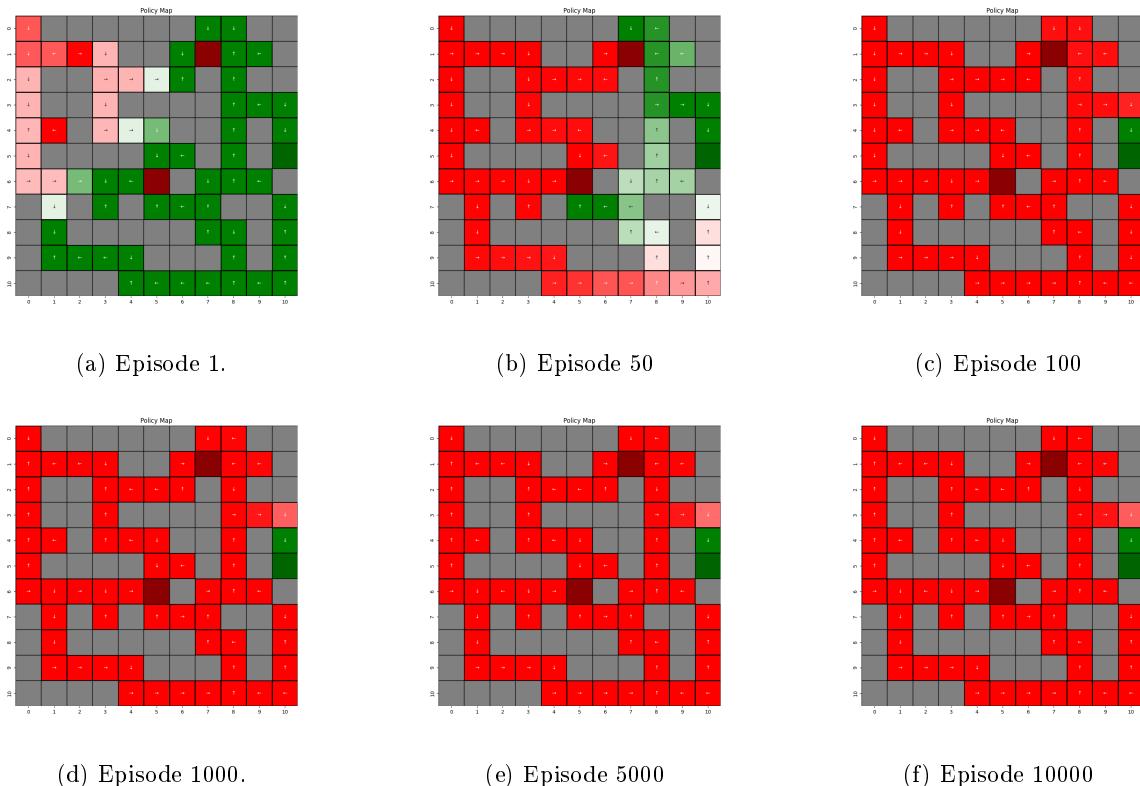


Figure 40: Evolution of policy maps throughout episodes.

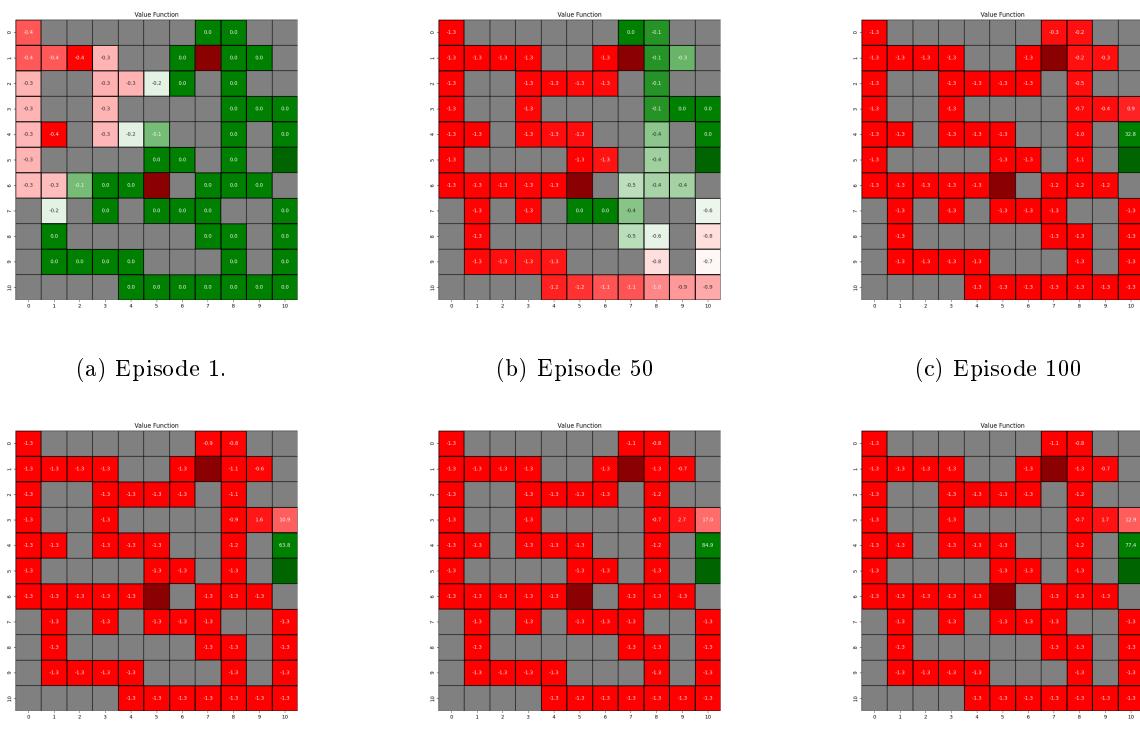


Figure 41: Evolution of value function throughout episodes.

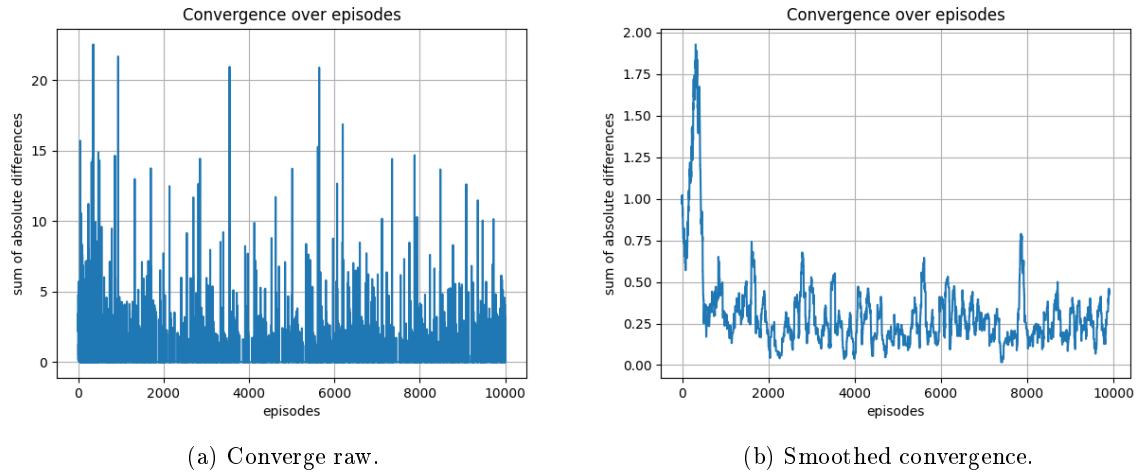


Figure 42: Converge of value function.

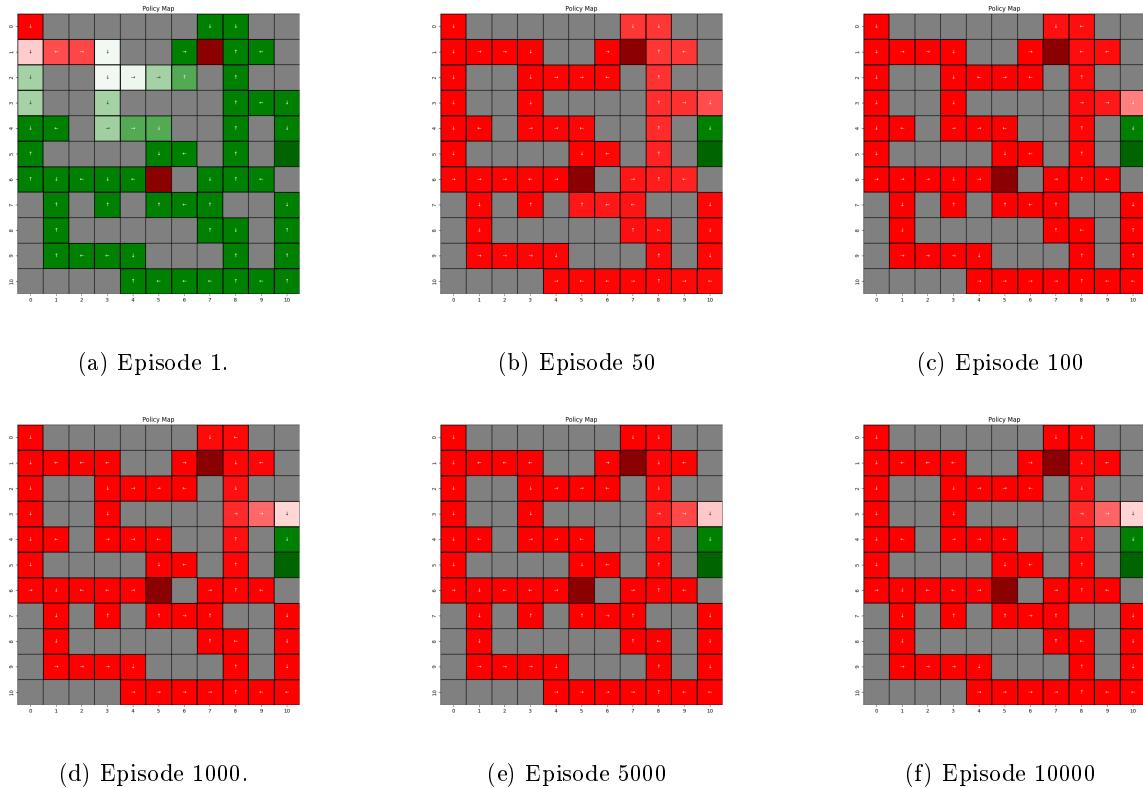


Figure 43: Evolution of policy maps throughout episodes.

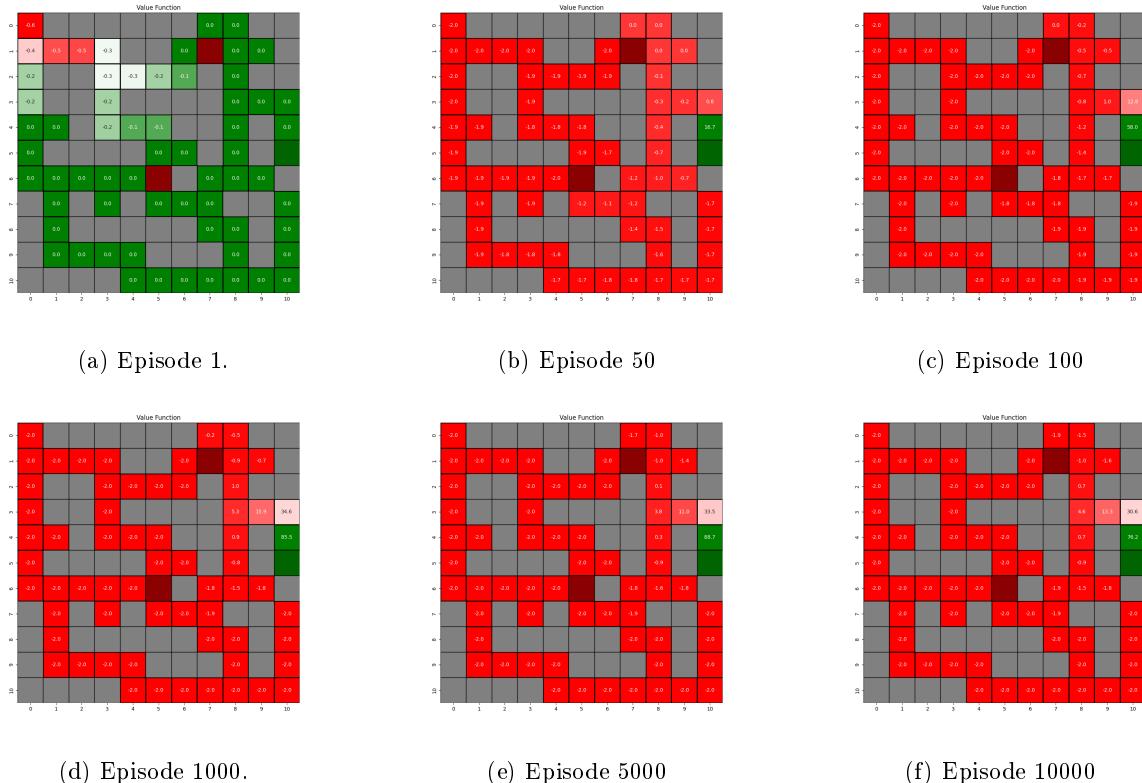


Figure 44: Evolution of value function throughout episodes.

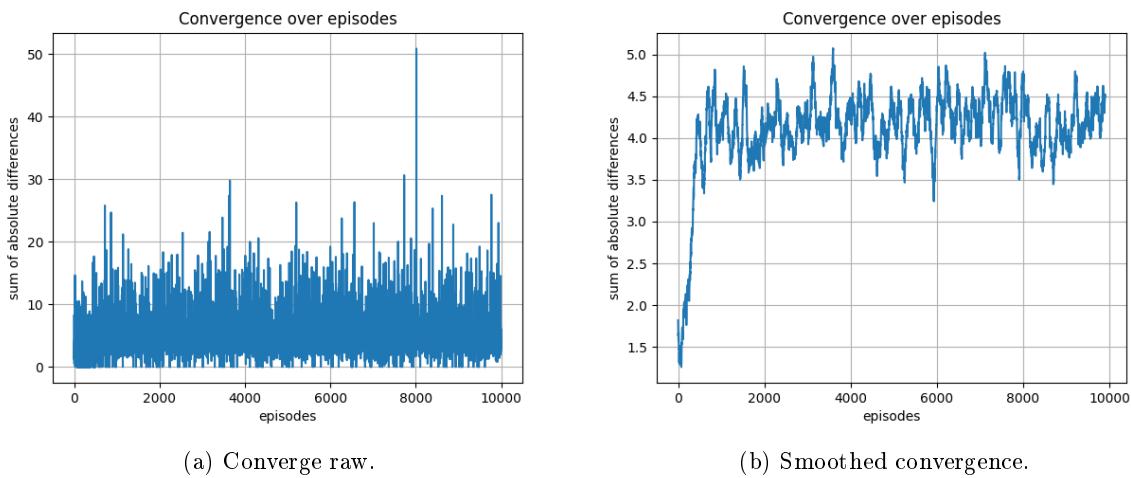


Figure 45: Convergence of value function.

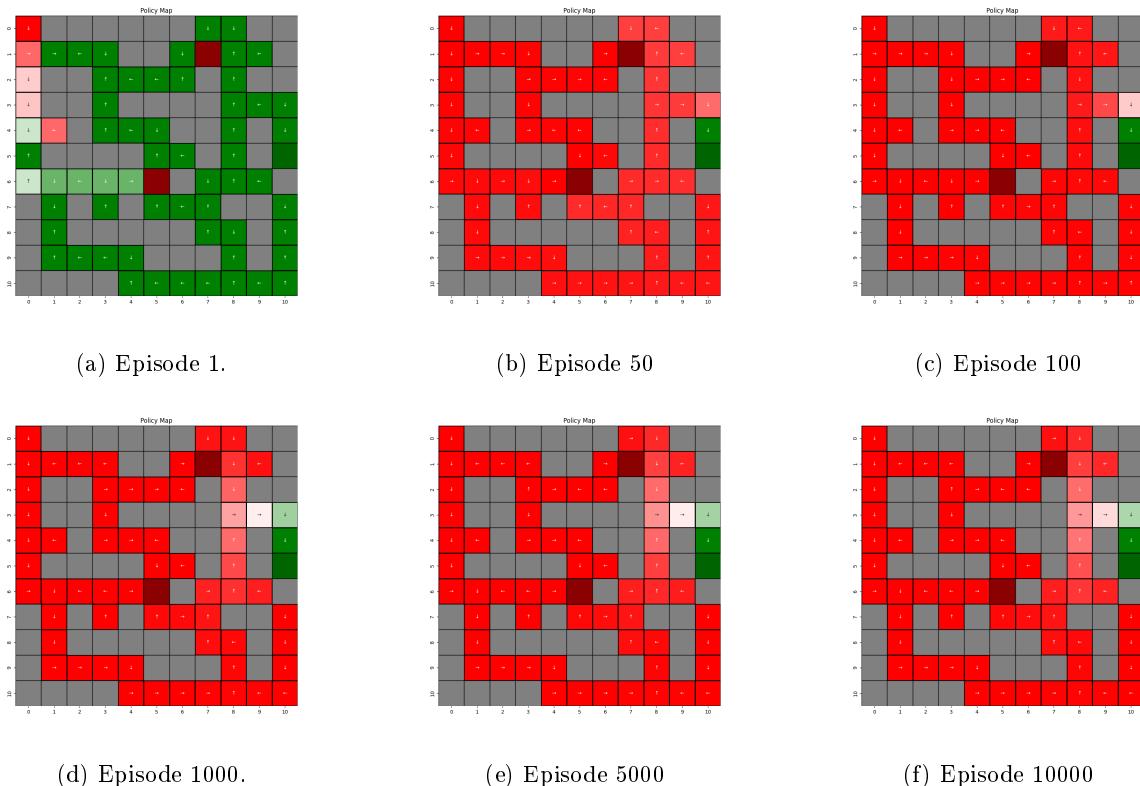


Figure 46: Evolution of policy maps throughout episodes.

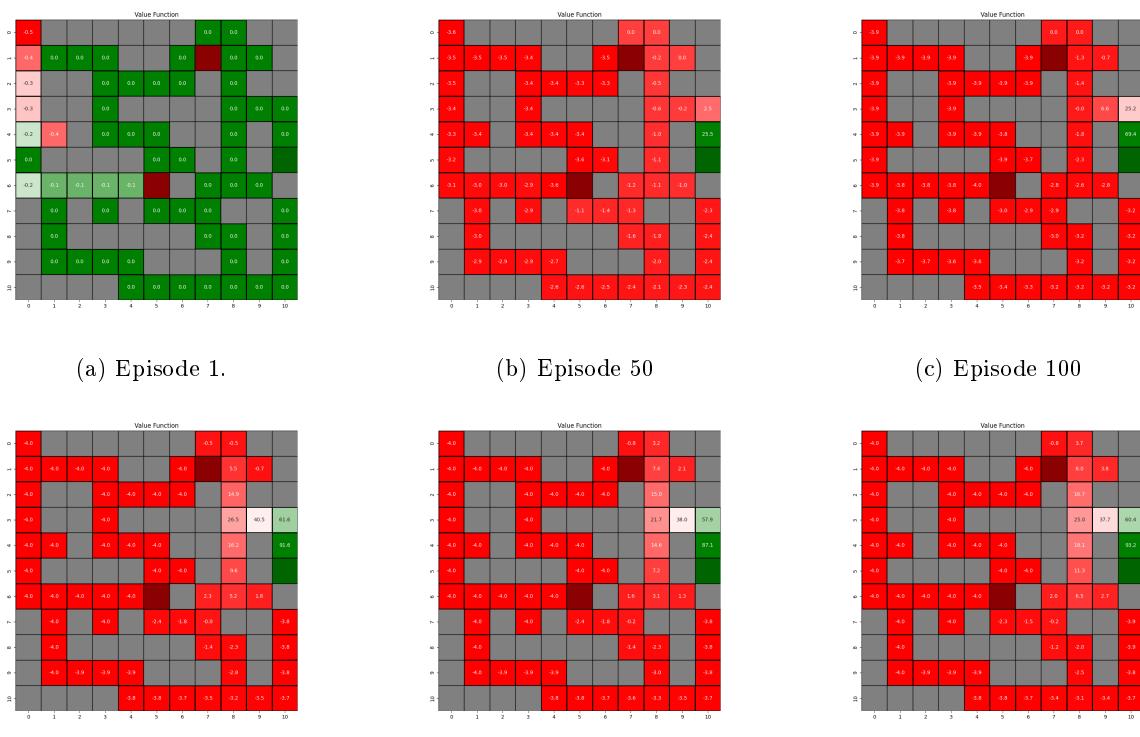


Figure 47: Evolution of value function throughout episodes.

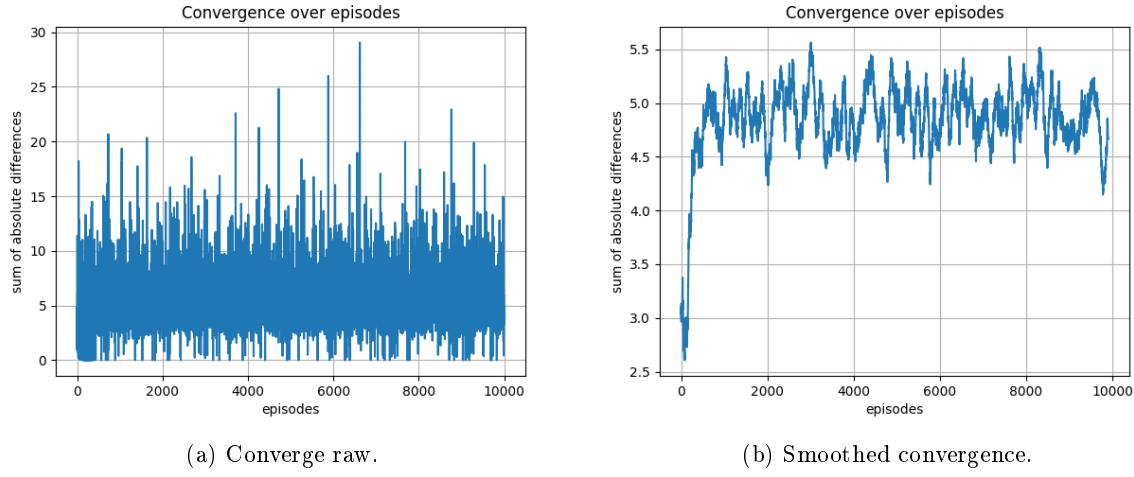


Figure 48: Converge of value function.

## 2.7 Effect of Epsilon in Temporal Difference Learning

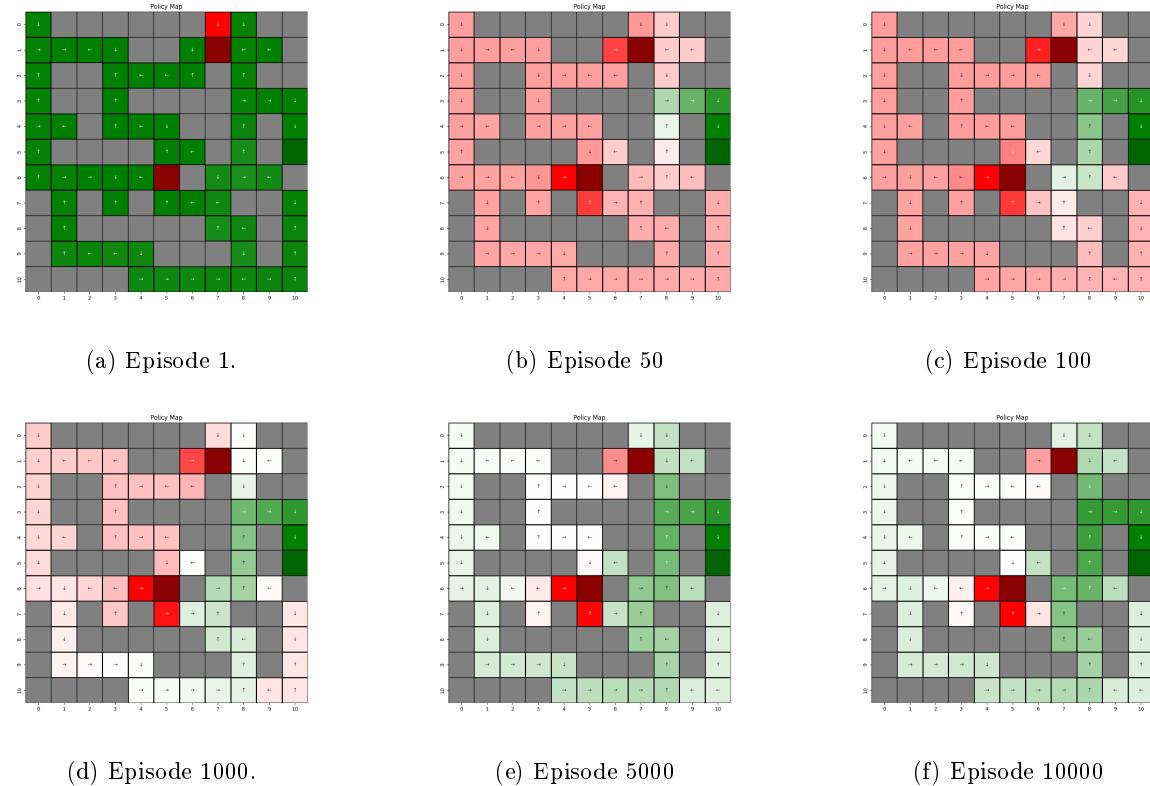


Figure 49: Evolution of policy maps throughout episodes.

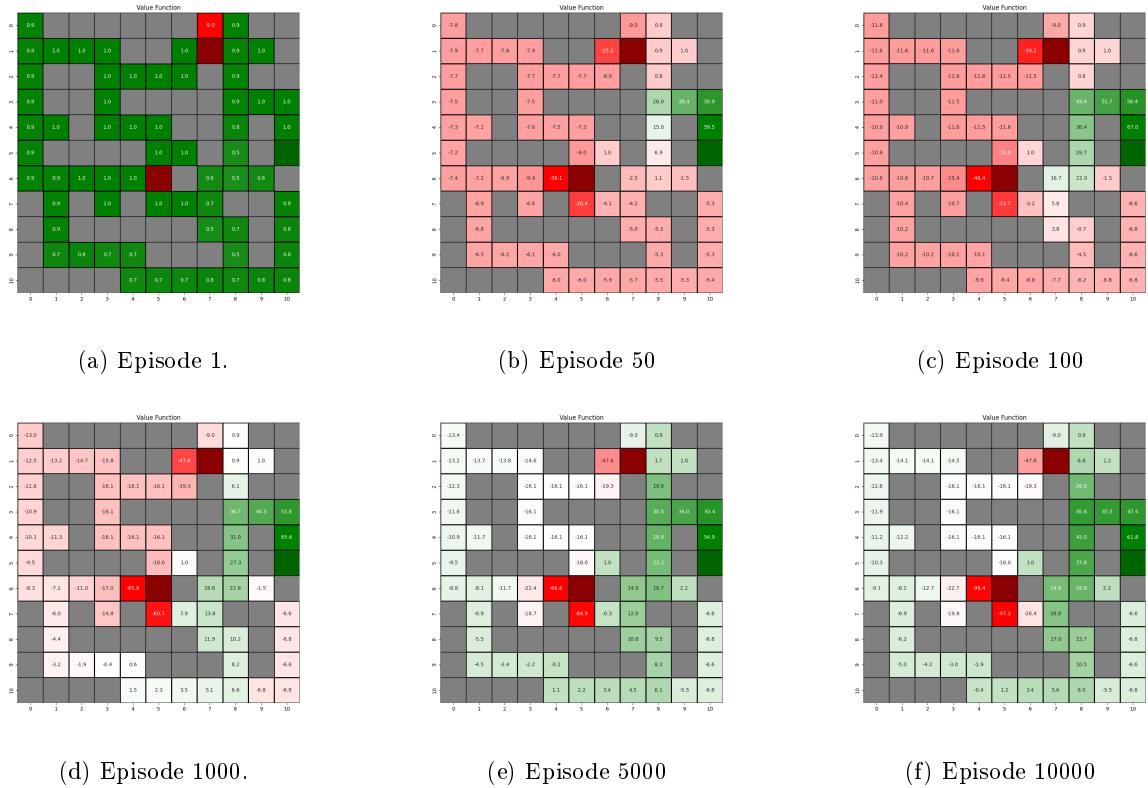


Figure 50: Evolution of value function throughout episodes.

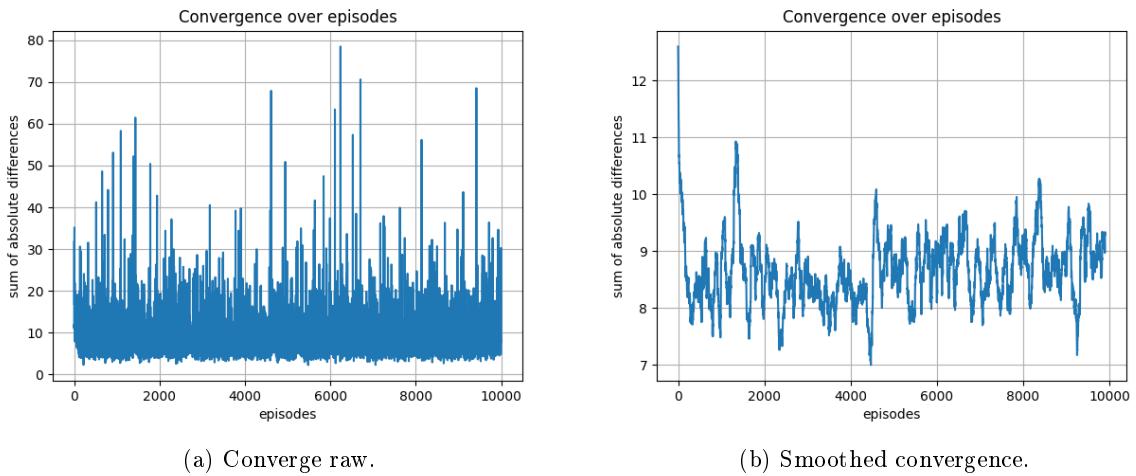


Figure 51: Convergence of value function.

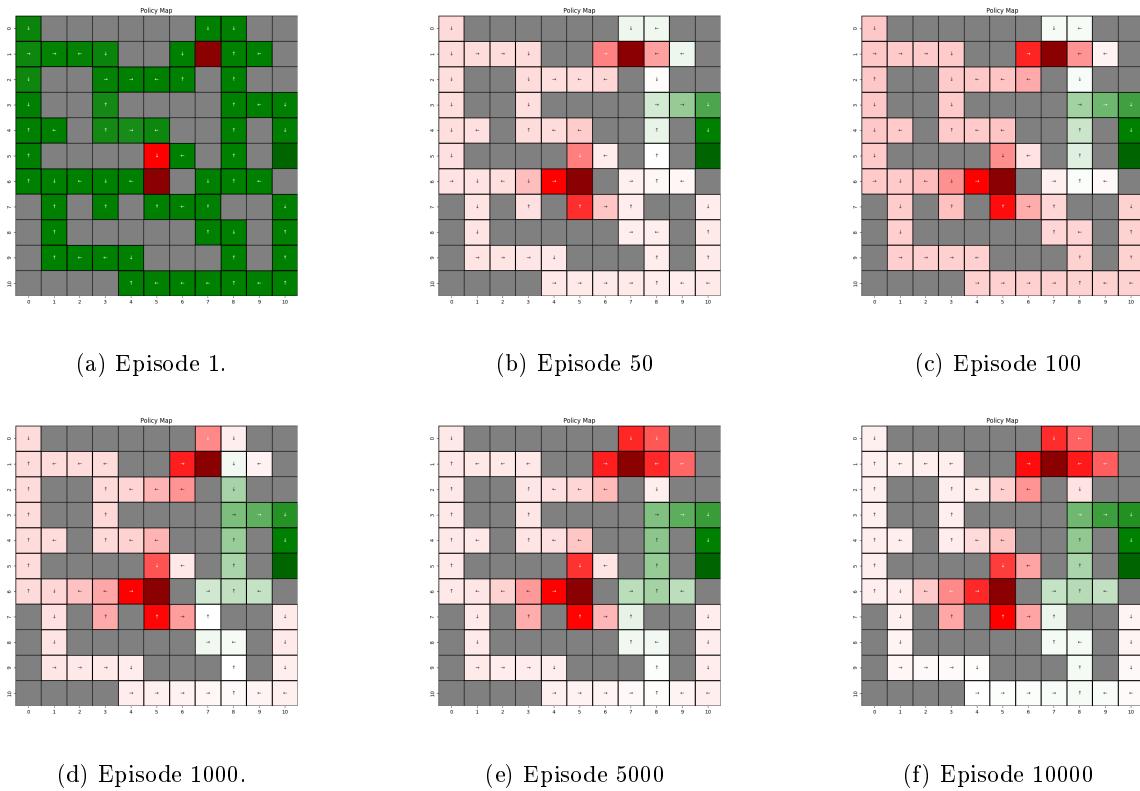


Figure 52: Evolution of policy maps throughout episodes.

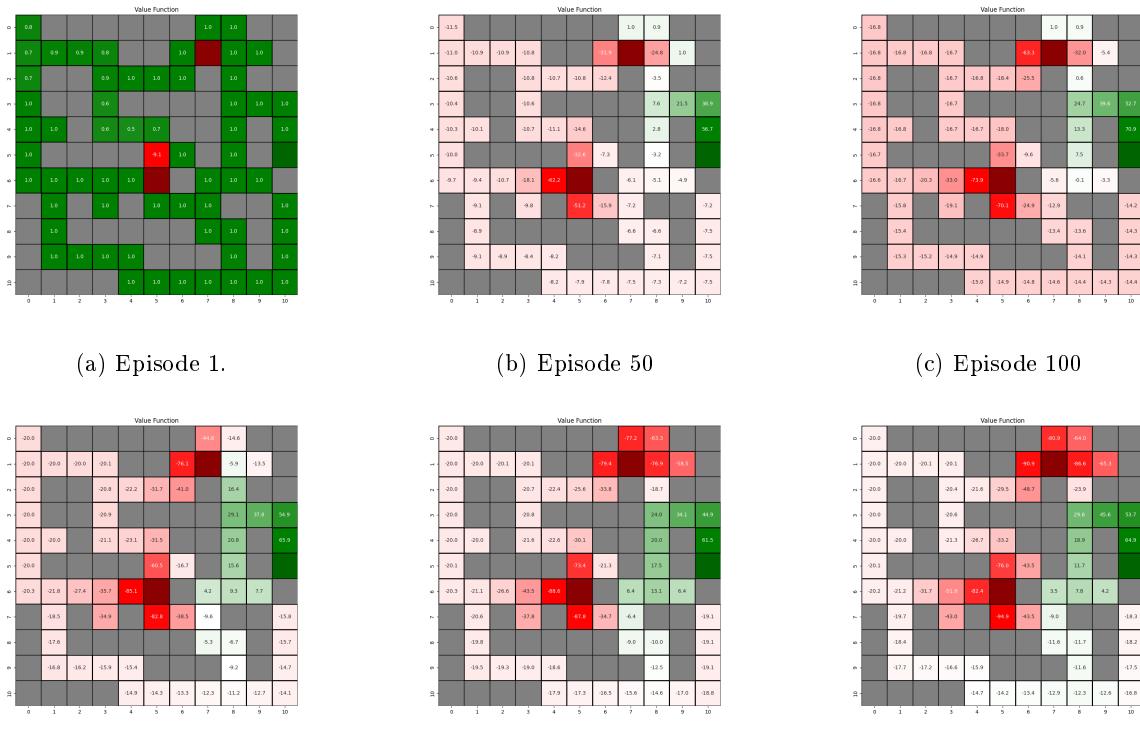


Figure 53: Evolution of value function throughout episodes.

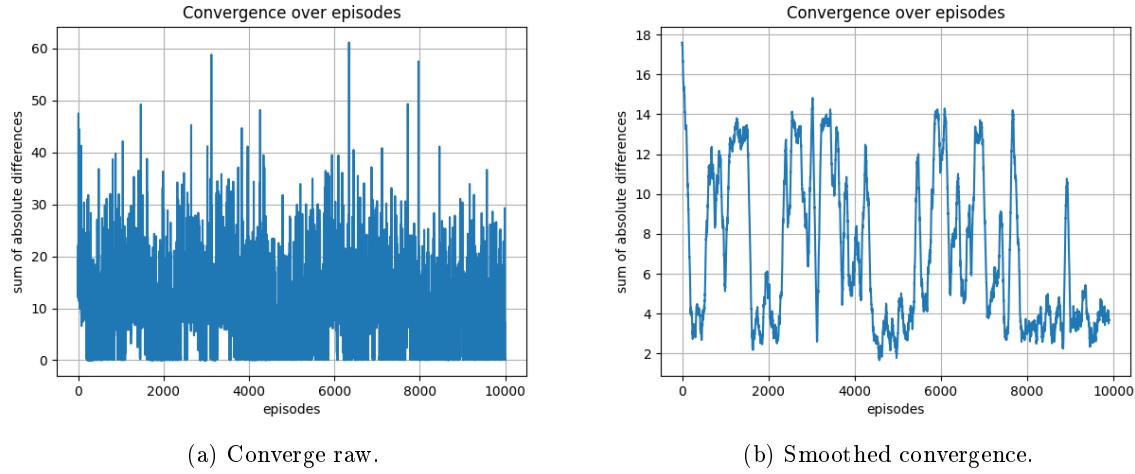


Figure 54: Converge of value function.

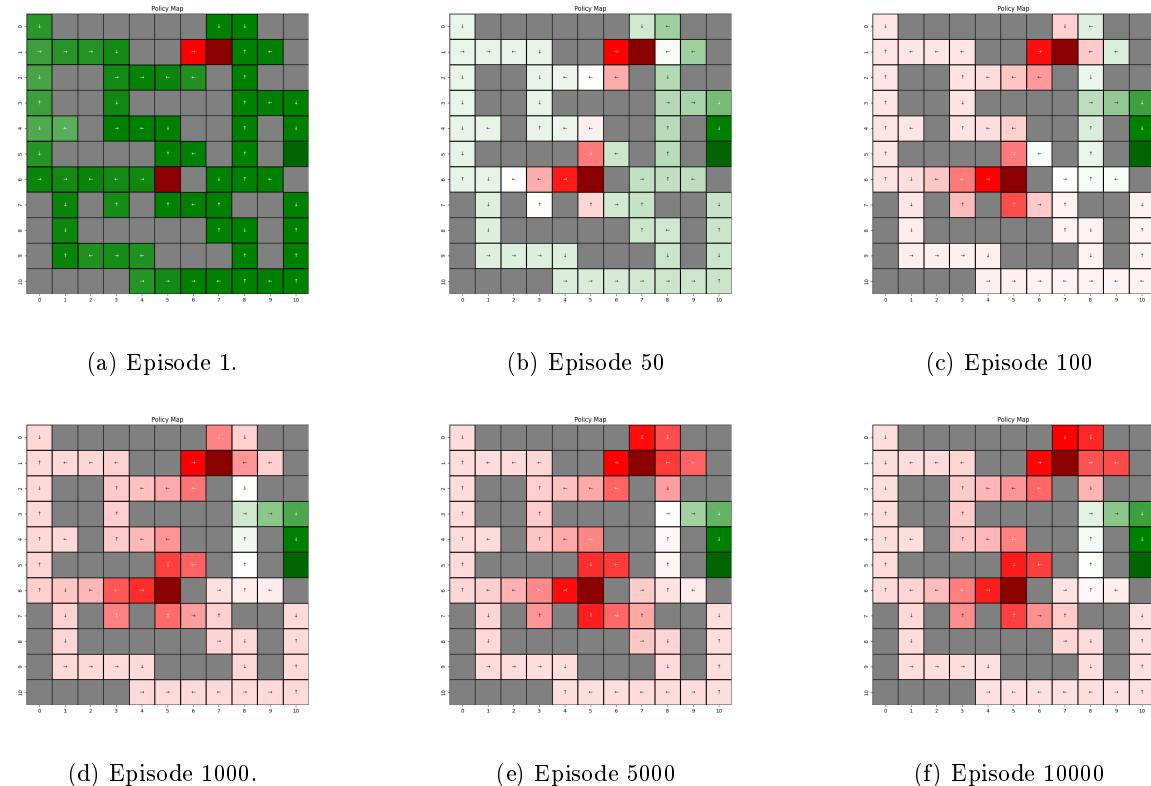
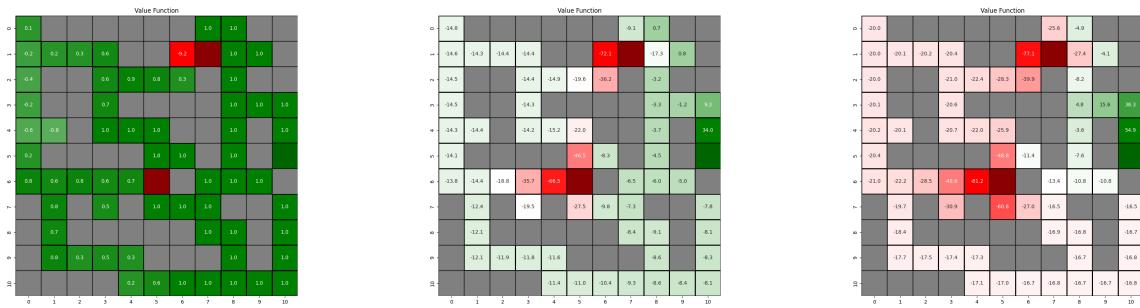


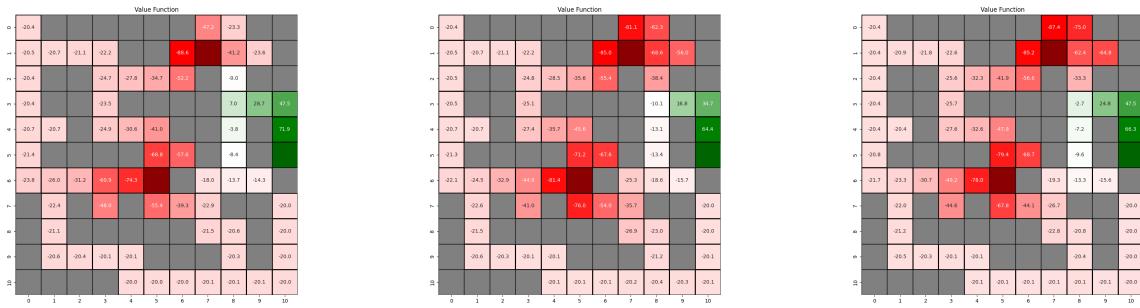
Figure 55: Evolution of policy maps throughout episodes.



(a) Episode 1.

(b) Episode 50

(c) Episode 100



(d) Episode 1000.

(e) Episode 5000

(f) Episode 10000

Figure 56: Evolution of value function throughout episodes.

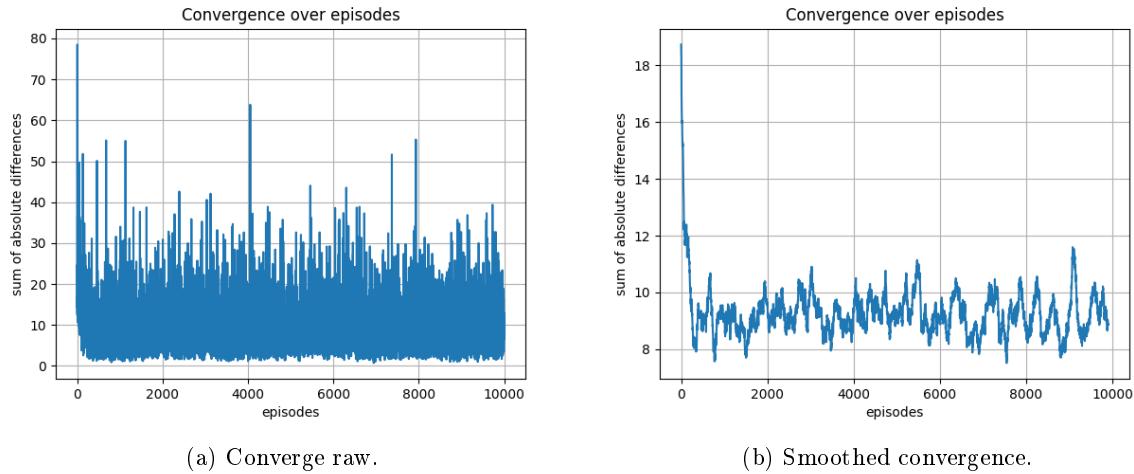


Figure 57: Convergence of value function.

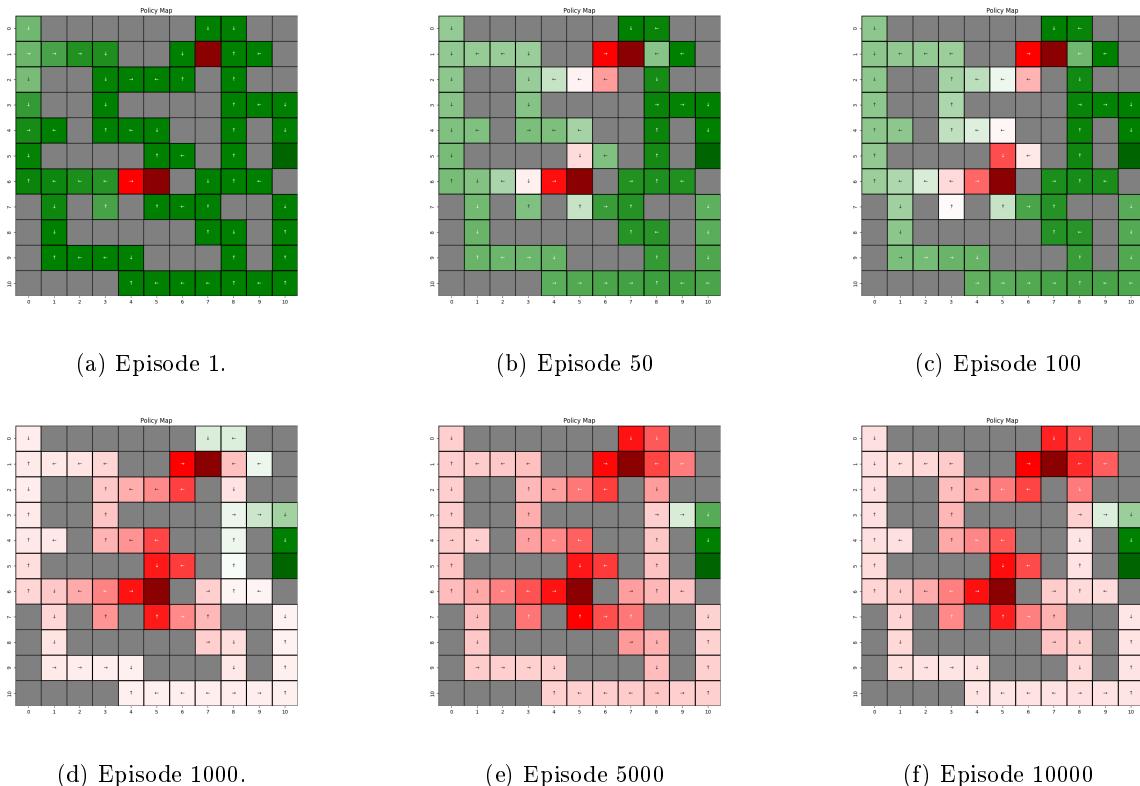


Figure 58: Evolution of policy maps throughout episodes.

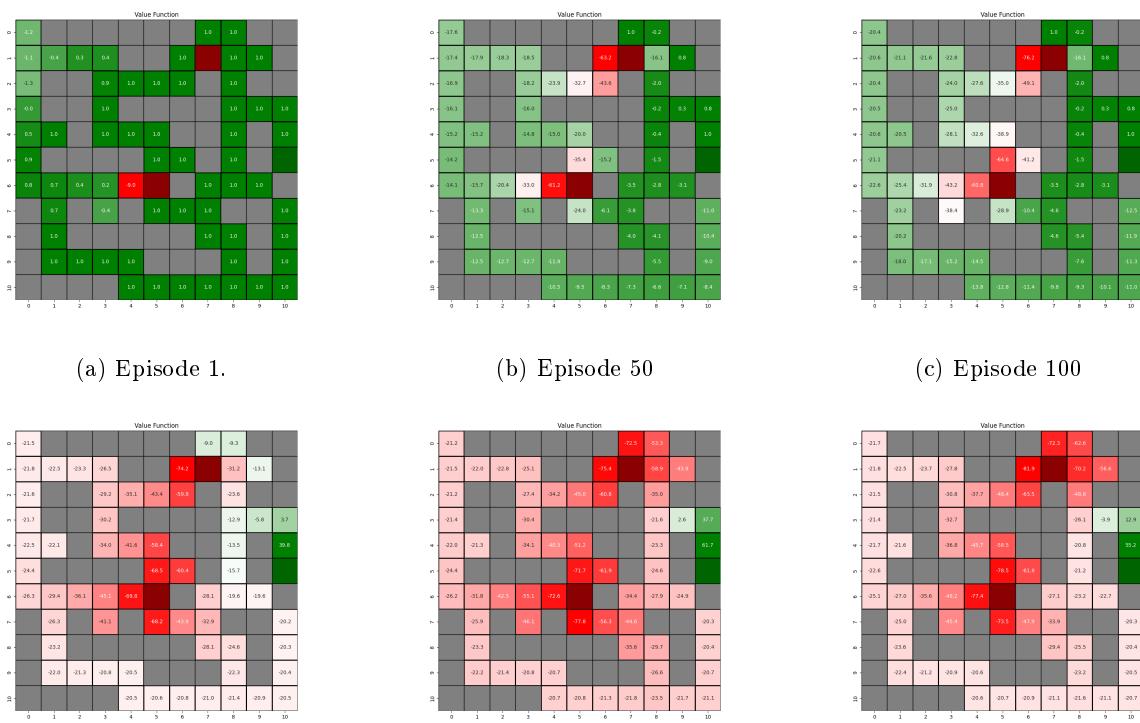


Figure 59: Evolution of value function throughout episodes.

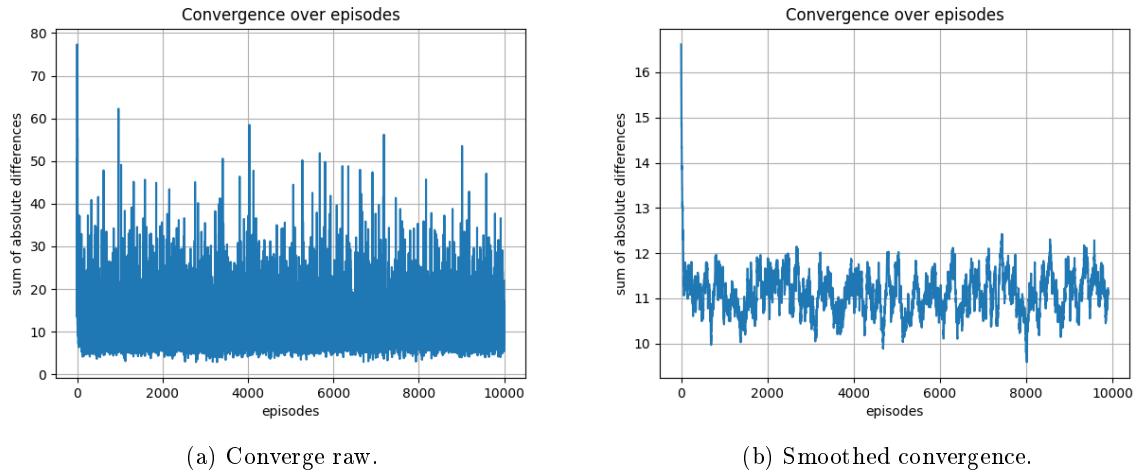


Figure 60: Converge of value function.

## 2.8 Effect of Epsilon in Q-Learning

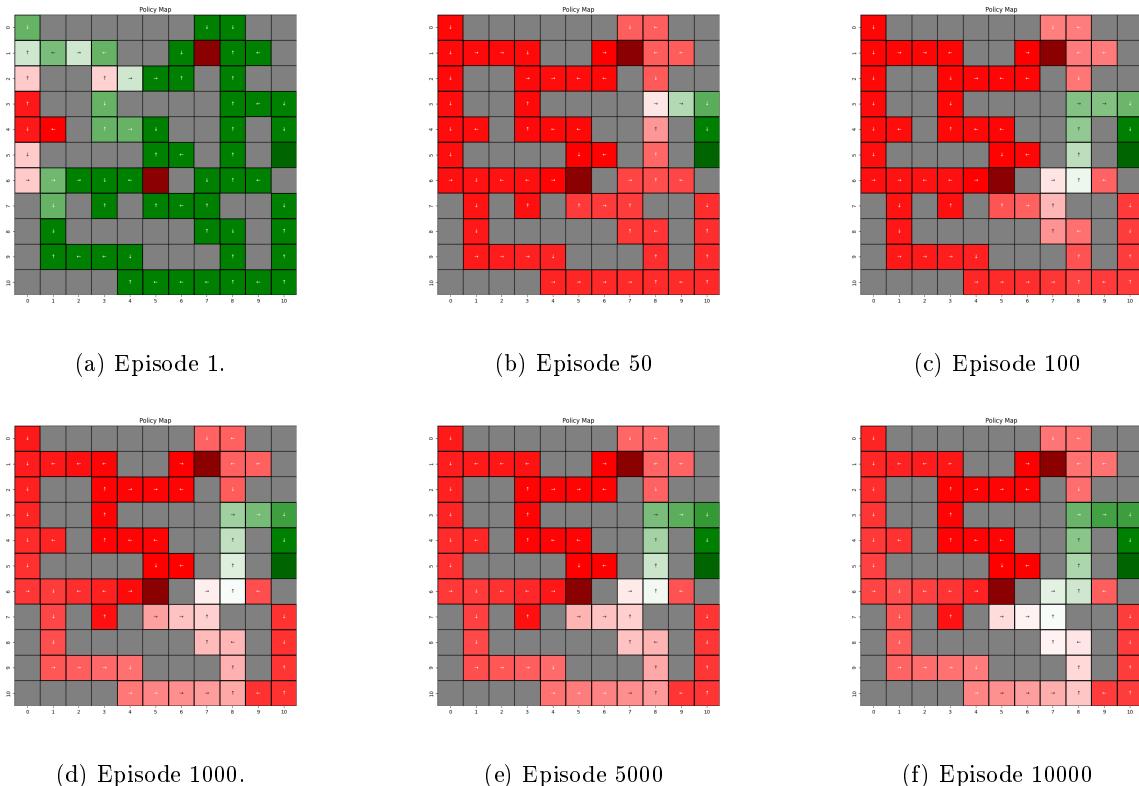


Figure 61: Evolution of policy maps throughout episodes.

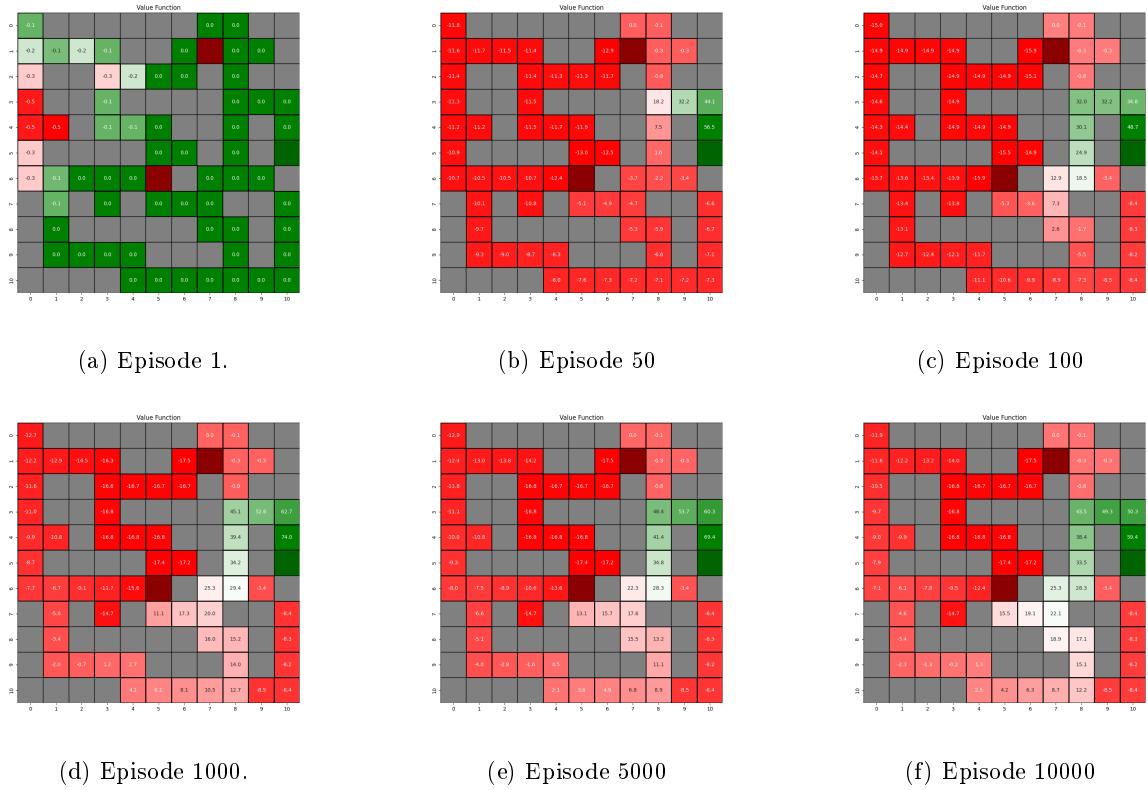


Figure 62: Evolution of value function throughout episodes.

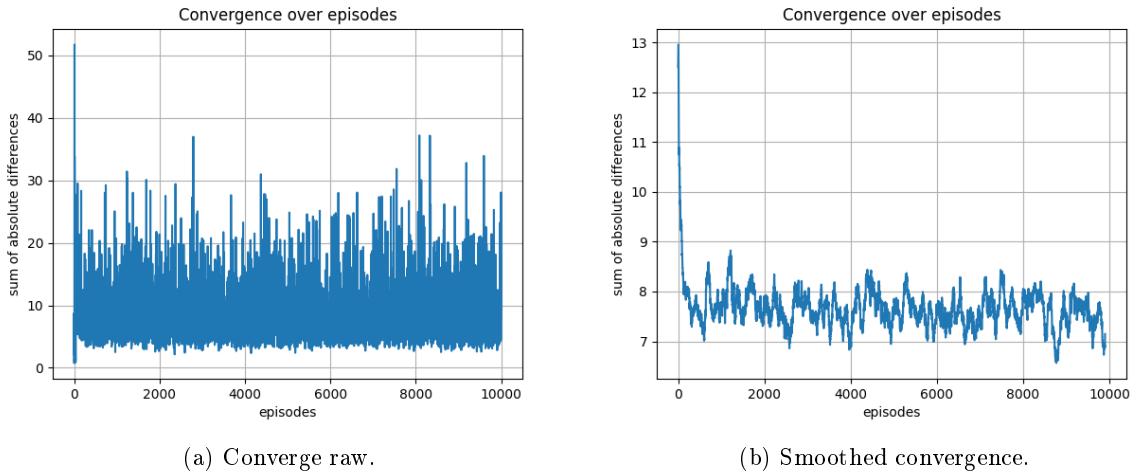


Figure 63: Convergence of value function.

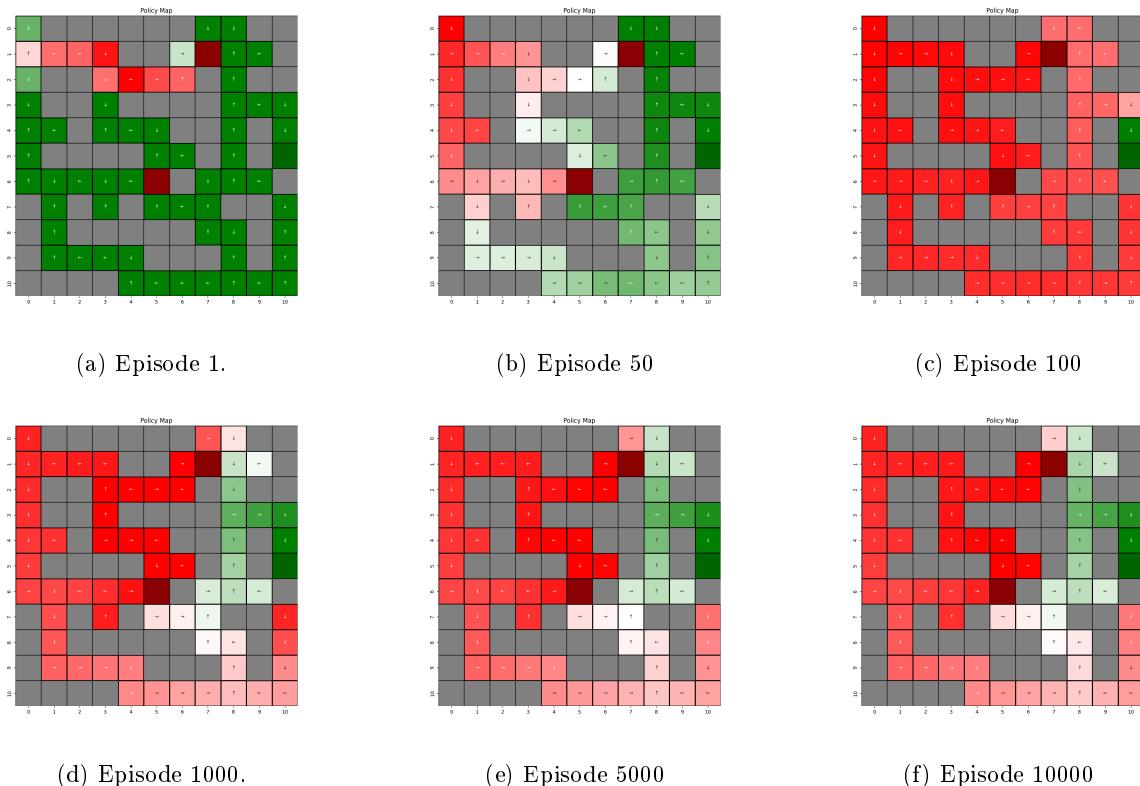


Figure 64: Evolution of policy maps throughout episodes.

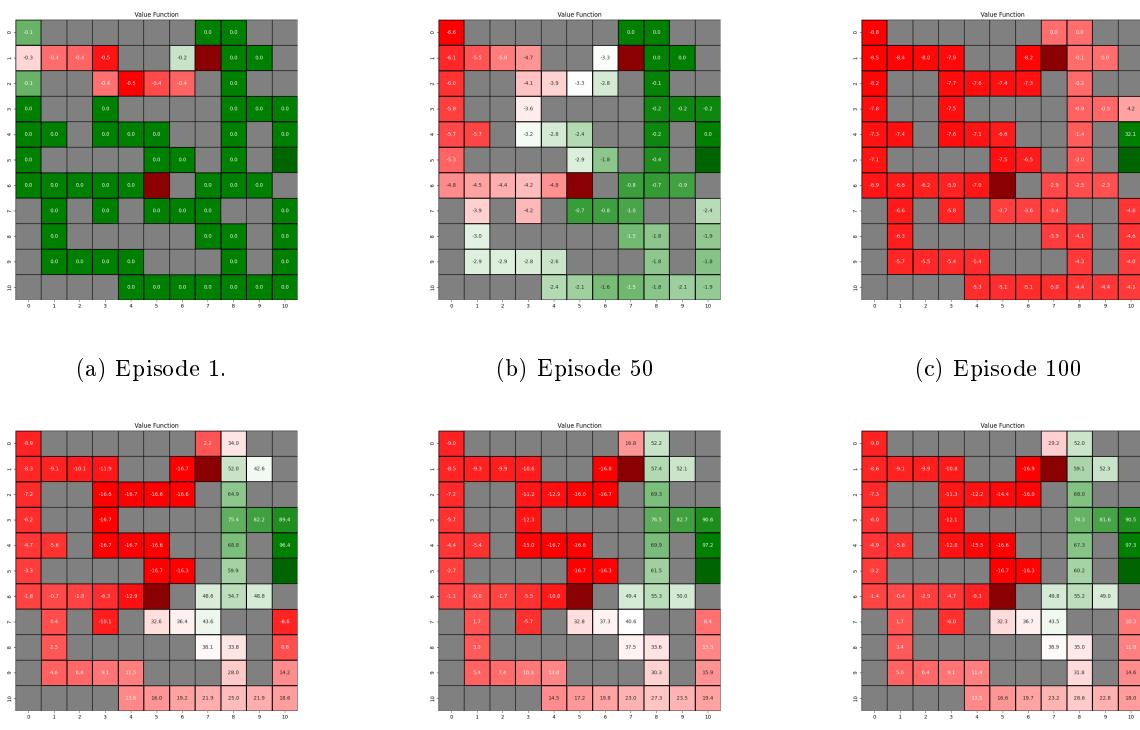


Figure 65: Evolution of value function throughout episodes.

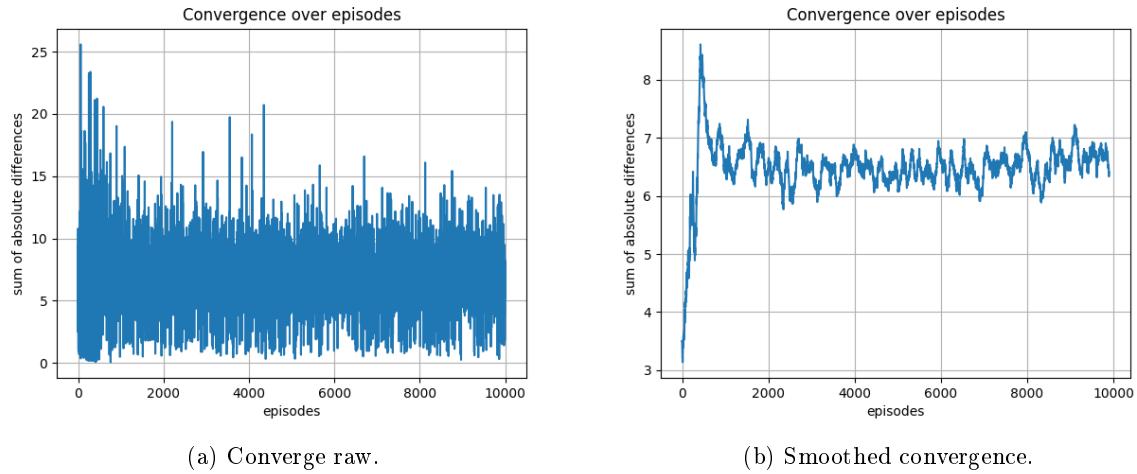


Figure 66: Converge of value function.

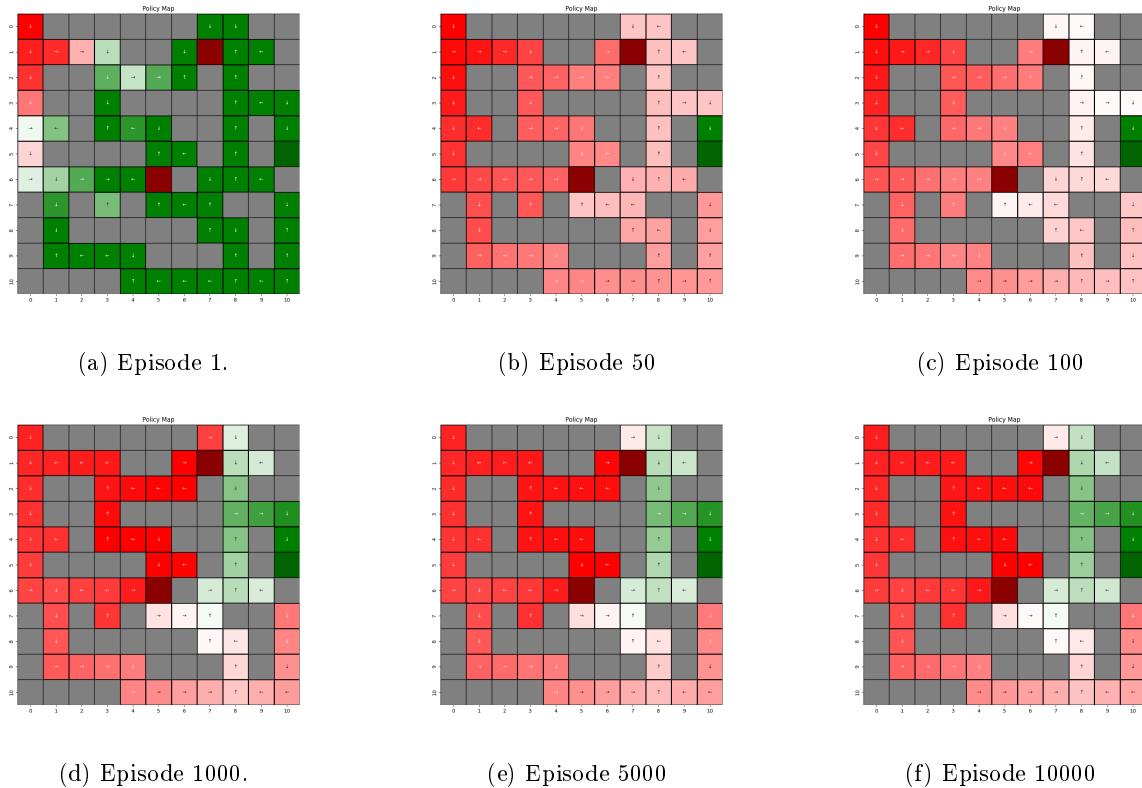


Figure 67: Evolution of policy maps throughout episodes.

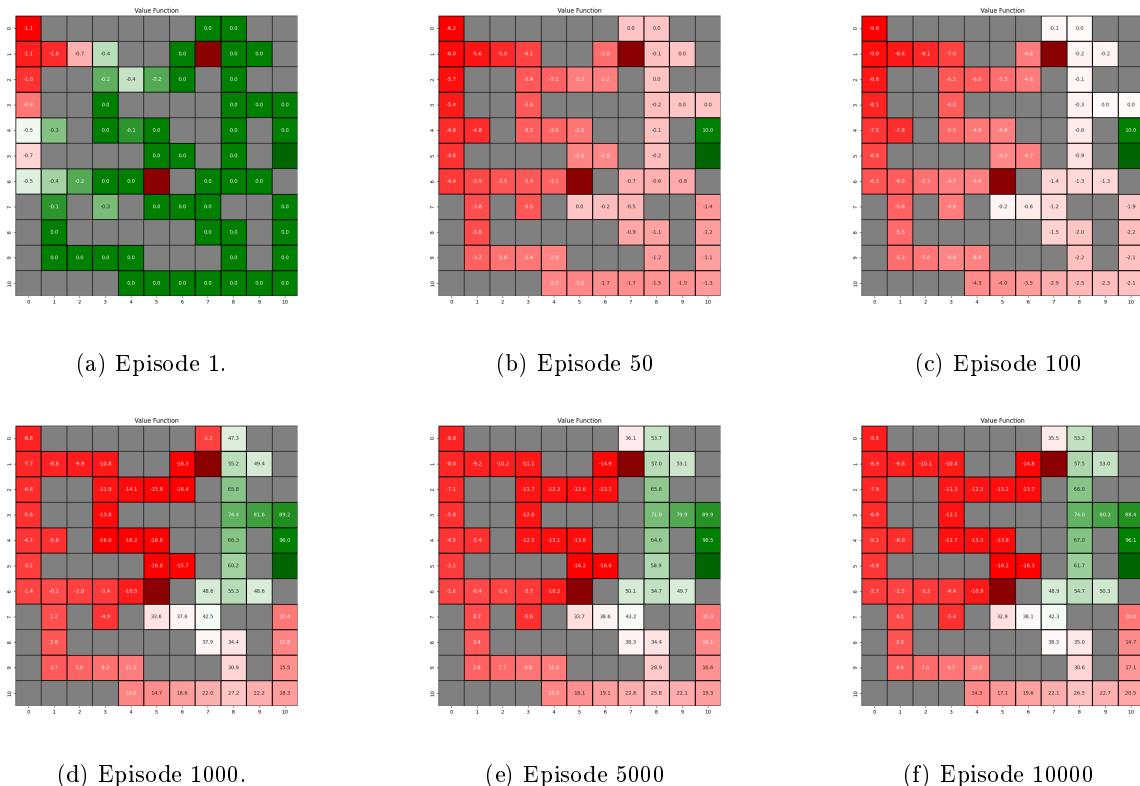


Figure 68: Evolution of value function throughout episodes.

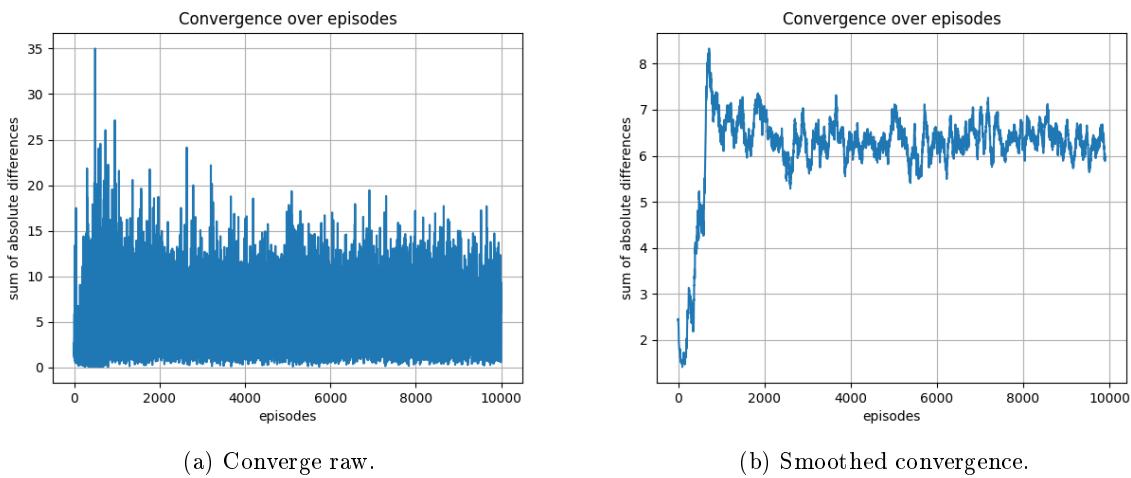


Figure 69: Convergence of value function.

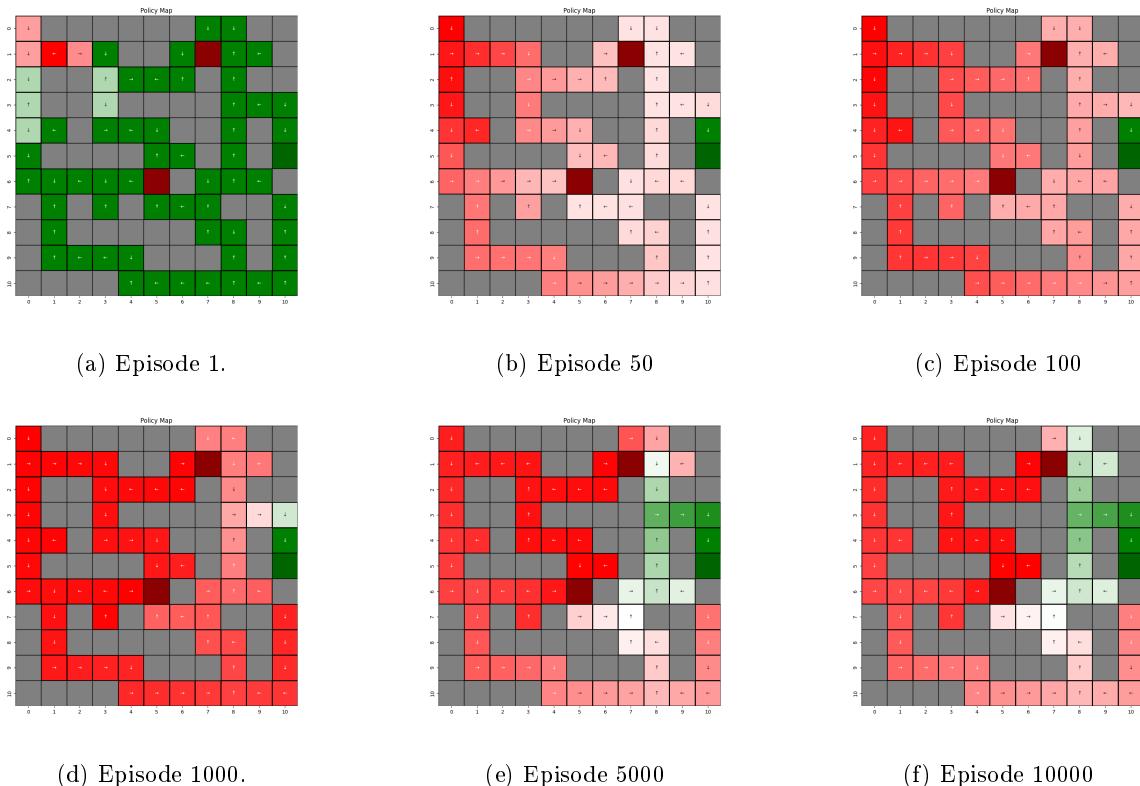


Figure 70: Evolution of policy maps throughout episodes.

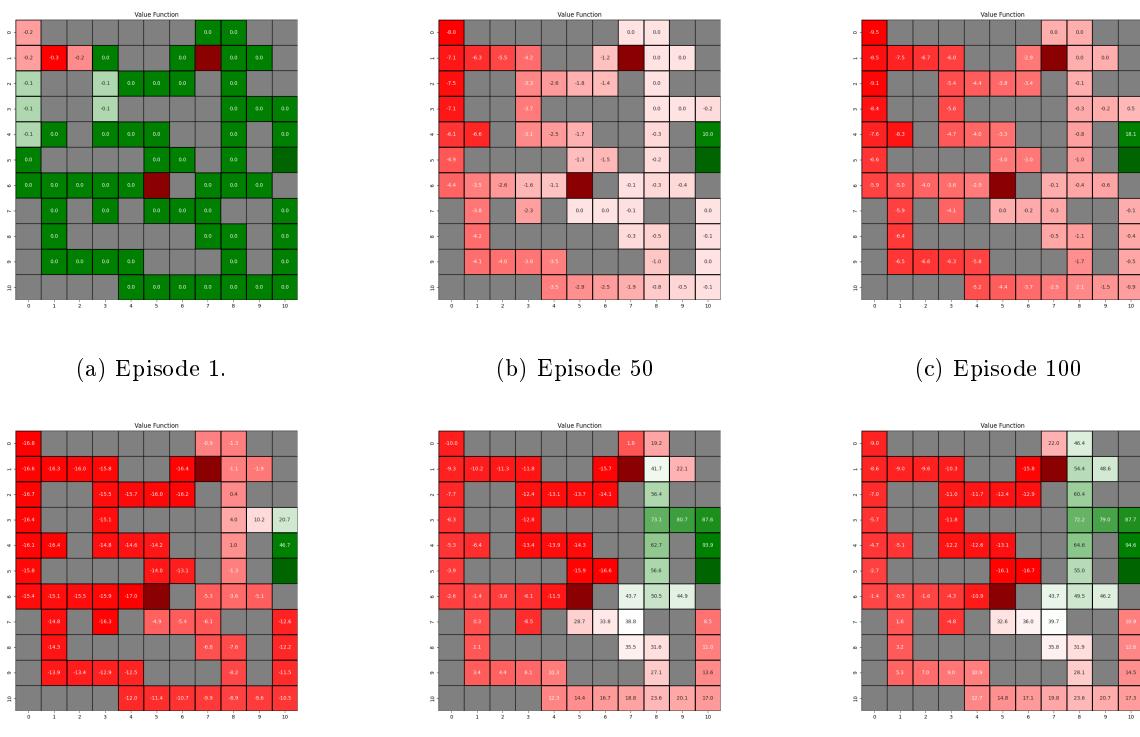


Figure 71: Evolution of value function throughout episodes.

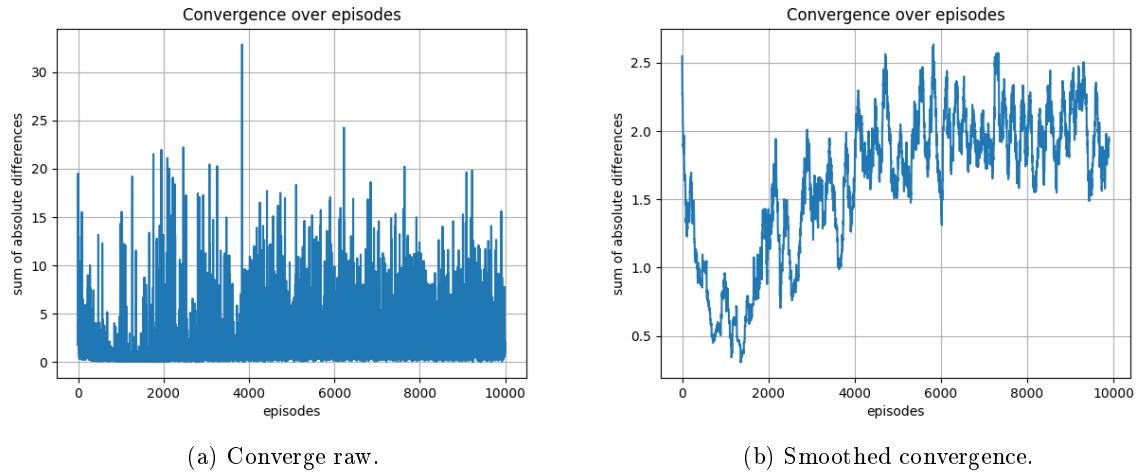


Figure 72: Converge of value function.

### 3 Discussions

#### 3.1 Q1

#### 3.2 Q2

#### 3.3 Q3

#### 3.4 Q4

#### 3.5 Q5

#### 3.6 Q6

#### 3.7 Q7

#### 3.8 Q8

#### 3.9 Q9

#### 3.10 Q10

### Appendix

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

Submitted by Ahmet Akman 2442366 on May 24, 2024.