









## Abstract

Deep Reinforcement Learning and policy gradient methods majorly contributed to the most recent advances in the field of Artificial Intelligence. These Methods enabled machines to surpass human performance for Atari console games (**mnih2015atari** ), boardgames like Chess, Shogi (**Shogi17** ) or Go (**Go2017** ) and most recently even complex team-based computer games (**OpenAI\_dota** ).

Simulating complex environments like the Atari console games can be

The goal of this thesis is the implementation and evaluation of the "Actor-Critic with Experience Replay"(ACER) algorithm proposed by **ACER** on the Atari console games.



## Contents

<b>1 Abstract</b>	<b>1</b>
<b>2 Introduction</b>	<b>1</b>
2.1 Motivation/Objectives . . . . .	1
<b>3 Reinforcement Learning Frameworks</b>	<b>2</b>
3.1 Elements of Reinforcement Learning . . . . .	2
3.2 Markov Decision Process . . . . .	3
<b>4 Actor-Critic Methods</b>	<b>4</b>
4.1 Actor-Only Methods . . . . .	4
4.2 Critic-Only Methods . . . . .	4
4.3 Actor-Critic Methods . . . . .	5
4.4 A3C : Asynchronous Advantage Actor-Critic . . . . .	5
<b>5 Off-Policy Learning</b>	<b>6</b>
5.1 Importance-Sampling . . . . .	6
5.2 Q-Trace . . . . .	6
<b>6 ACER : Actor-Critic with Experience Replay</b>	<b>7</b>
<b>7 Experimental Setup</b>	<b>8</b>
<b>8 Experiments</b>	<b>9</b>
8.1 Hyperparameter Robustness . . . . .	9
8.2 Adversarial Attacks . . . . .	9
8.3 Results . . . . .	10
<b>9 Conclusion</b>	<b>10</b>
<b>10 References</b>	<b>11</b>
<b>List of Figures</b>	<b>12</b>
<b>List of Tables</b>	<b>12</b>





## **1 Abstract**

## **2 Introduction**

### **2.1 Motivation/Objectives**

## **3 Reinforcement Learning Frameworks**

### **3.1 Elements of Reinforcement Learning**

**3.2 Markov Decision Process**

## **4 Actor-Critic Methods**

### **4.1 Actor-Only Methods**

### **4.2 Critic-Only Methods**

**4.3 Actor-Critic Methods**

**4.4 A3C : Asynchronous Advantage Actor-Critic**

## 5 Off-Policy Learning

### 5.1 Importance-Sampling

### 5.2 Q-Trace

## **6 ACER : Actor-Critic with Experience Replay**

## 7 Experimental Setup



## **8 Experiments**

### **8.1 Hyperparameter Robustness**

### **8.2 Adversial Attacks**

### 8.3 Results

## 9 Conclusion

## 10 References

**List of Figures****List of Tables**