

**POLITECNICO DI MILANO**  
**Master's Degree in Computer Science and Engineering**  
**Dipartimento di Elettronica, Informazione e Bioingegneria**

**FAST REINFORCEMENT LEARNING  
USING DEEP STATE-ACTION  
FEATURES**

**AI & R Lab**  
**Laboratorio di Intelligenza Artificiale**  
**e Robotica del Politecnico di Milano**

**Supervisor: Prof. Marcello Restelli**  
**Co-supervisor: Prof. Andrea Bonarini**

**Master's Thesis by:**  
**Daniele Grattarola, matricola 853101**

**Academic Year 2016-2017**



*To nobody*



# Indice

|   |            |
|---|------------|
| <b>Summary</b>  | <b>III</b> |
| <b>Aknowledgments</b>                                 | <b>V</b>   |
| <b>1 Introduction</b>                                 | <b>1</b>   |
| 1.1 Inquadramento generale . . . . .                  | 1          |
| 1.2 Breve descrizione del lavoro . . . . .            | 2          |
| 1.3 Struttura della tesi . . . . .                    | 2          |
| <b>2 State Of The Art</b>                             | <b>3</b>   |
| 2.1 Deep Learning . . . . .                           | 3          |
| 2.2 Reinforcement Learning . . . . .                  | 3          |
| 2.3 Deep Reinforcement Learning . . . . .             | 3          |
| <b>3 Impostazione del problema di ricerca</b>         | <b>5</b>   |
| <b>4 Progetto logico della soluzione del problema</b> | <b>7</b>   |
| <b>5 Architettura del sistema</b>                     | <b>9</b>   |
| <b>6 Realizzazioni sperimentali e valutazione</b>     | <b>11</b>  |
| <b>7 Direzioni future di ricerca e conclusioni</b>    | <b>13</b>  |
| <b>Bibliografia</b>                                   | <b>15</b>  |
| <b>A Documentazione del progetto logico</b>           | <b>17</b>  |
| <b>B Documentazione della programmazione</b>          | <b>19</b>  |
| <b>C Listato</b>                                      | <b>21</b>  |
| <b>D Il manuale utente</b>                            | <b>23</b>  |

|          |                           |           |
|----------|---------------------------|-----------|
| <b>E</b> | <b>Esempio di impiego</b> | <b>25</b> |
| <b>F</b> | <b>Datasheet</b>          | <b>27</b> |



# Summary

Il sommario deve contenere 3 o 4 frasi tratte dall'introduzione di cui la prima inquadra l'area dove si svolge il lavoro (eventualmente la seconda inquadra la sottoarea più specifica del lavoro), la seconda o la terza frase dovrebbe iniziare con le parole "Lo scopo della tesi è ..." e infine la terza o quarta frase riassume brevemente l'attività svolta, i risultati ottenuti ed eventuali valutazioni di questi.

NB: se il relatore effettivo è interno al Politecnico di Milano nel frontesimo si scrive Relatore, se vi è la collaborazione di un altro studioso lo si riporta come Correlatore come sopra. Nel caso il relatore effettivo sia esterno si scrive Relatore esterno e poi bisogna inserire anche il Relatore interno. Nel caso il relatore sia un ricercatore allora il suo Nome COGNOME dovrà essere preceduto da Ing. oppure Dott., a seconda dei casi.





# Acknowledgments

I thank...



# Capitolo 1

## Introduction

*“Terence: Rotta a nord con circospezione*

*Bud: Ehi, gli ordini li do io qui!*

*Terence: Ok, comante*

*Bud: Rotta a nord*

*Terence: Soltanto?*

*Bud: Con circospezione!”*

Chi Trova un Amico Trova un Tesoro

L'introduzione deve essere atomica, quindi non deve contenere nè sottosezioni nè paragrafi nè altro. Il titolo, il sommario e l'introduzione devono sembrare delle scatole cinesi, nel senso che lette in quest'ordine devono progressivamente svelare informazioni sul contenuto per incatenare l'attenzione del lettore e indurlo a leggere l'opera fino in fondo. L'introduzione deve essere tripartita, non graficamente ma logicamente:

### 1.1 Inquadramento generale

La prima parte contiene una frase che spiega l'area generale dove si svolge il lavoro; una che spiega la sottoarea più specifica dove si svolge il lavoro e la terza, che dovrebbe cominciare con le seguenti parole “lo scopo della tesi è ...”, illustra l'obiettivo del lavoro. Poi vi devono essere una o due frasi che contengano una breve spiegazione di cosa e come è stato fatto, delle attività sperimentali, dei risultati ottenuti con una valutazione e degli sviluppi futuri. La prima parte deve essere circa una facciata e mezza o due

## 1.2 Breve descrizione del lavoro

La seconda parte deve essere una esplosione della prima e deve quindi mostrare in maniera più esplicita l'area dove si svolge il lavoro, le fonti bibliografiche più importanti su cui si fonda il lavoro in maniera sintetica (una pagina) evidenziando i lavori in letteratura che presentano attinenza con il lavoro affrontato in modo da mostrare da dove e perché è sorta la tematica di studio. Poi si mostrano esplicitamente le realizzazioni, le direttive future di ricerca, quali sono i problemi aperti e quali quelli affrontati e si ripete lo scopo della tesi. Questa parte deve essere piena (ma non grondante come la sezione due) di citazioni bibliografiche e deve essere lunga circa 4 facciate.

## 1.3 Struttura della tesi

La terza parte contiene la descrizione della struttura della tesi ed è organizzata nel modo seguente. “La tesi è strutturata nel modo seguente.

Nella sezione due si mostra ...

Nella sez. tre si illustra ...

Nella sez. quattro si descrive ...

Nelle conclusioni si riassumono gli scopi, le valutazioni di questi e le prospettive future ...

Nell'appendice A si riporta ... (Dopo ogni sezione o appendice ci vuole un punto).”

I titoli delle sezioni da 2 a M-1 sono indicativi, ma bisogna cercare di mantenere un significato equipollente nel caso si vogliano cambiare. Queste sezioni possono contenere eventuali sottosezioni.

## Capitolo 2

# State Of The Art

This section will be dedicated to the analysis of the two main areas of machine learning around which this thesis revolves: *deep learning* and *reinforcement learning*. After an introduction of both areas, we will describe how the two classes of algorithms can be used together in what is called *deep reinforcement learning*.

### 2.1 Deep Learning

Deep Learning is a branch of machine learning which exploits *Artificial Neural Networks* (ANN) with more than one hidden layer to learn an abstract representation of the input space [1].

Deep learning techniques can be applied to the three main classes of problems of machine learning (supervised, semi-supervised, and unsupervised), and have been used to achieve state-of-the-art results in a variety of learning tasks.

### 2.2 Reinforcement Learning

### 2.3 Deep Reinforcement Learning



## Capitolo 3

# Impostazione del problema di ricerca

*“Bud: Apri!”*

*Cattivo: Perché, altrimenti vi arrabbiate?*

*Bud e Terence: Siamo già arrabbiati!”*

Altrimenti ci arrabbiamo

In questa sezione si deve descrivere l’obiettivo della ricerca, le problematiche affrontate ed eventuali definizioni preliminari nel caso la tesi sia di carattere teorico.





## Capitolo 4

# Progetto logico della soluzione del problema

*“Bud: No, calma, calma, stiamo calmi, noi siamo su un’isola deserta, e per il momento non t’ammazzo perché mi potresti servire come cibo ...”*

Chi trova un amico trova un tesoro

In questa sezione si spiega come è stato affrontato il problema concettualmente, la soluzione logica che ne è seguita senza la documentazione.



## Capitolo 5

# Architettura del sistema

*“Terence: Ma scusa di che ti preoccupi, i piedipiatti hanno altro a cui pensare, in questo momento stanno cercando due cadaveri scomparsi  
Bud: Se non spegni quella sirena uno di quei due cadaveri scomparsi lo trovano di sicuro!”*

Nati con la camicia

Si mostra il progetto dell'architettura del sistema con i vari moduli.



## Capitolo 6

# Realizzazioni sperimentali e valutazione

*“Bambino: Questo è l’ultimo avviso per voi e i vostri rubagalline  
Il pistolero si alza: Che avete detto?  
Bambino: RUBAGALLINE  
Il pistolero si risiede: Aaah.”*

Lo chiamavano Trinità ...

Si mostra il progetto dal punto di vista sperimentale, le cose materialmente realizzate. In questa sezione si mostrano le attività sperimentali svolte, si illustra il funzionamento del sistema (a grandi linee) e si spiegano i risultati ottenuti con la loro valutazione critica. Bisogna introdurre dati sulla complessità degli algoritmi e valutare l’efficienza del sistema.



## Capitolo 7

# Direzioni future di ricerca e conclusioni

*“Terence: Mi fai un gelato anche a me? Lo vorrei di pistacchio.*

*Bud: Non ce l’ho il pistacchio. C’ho la vaniglia, cioccolato, fragola, limone e caffè.*

*Terence: Ah bene. Allora fammi un cono di vaniglia e di pistacchio.*

*Bud: No, non ce l’ho il pistacchio. C’ho la vaniglia, cioccolato, fragola, limone e caffè.*

*Terence: Ah, va bene. Allora vediamo un po’, fammelo al cioccolato, tutto coperto di pistacchio.*

*Bud: Ehi, macché sei sordo? Ti ho detto che il pistacchio non ce l’ho!*

*Terence: Ok ok, non c’è bisogno che t’arrabbi, no? Insomma, di che ce l’hai?*

*Bud: Ce l’ho di vaniglia, cioccolato, fragola, limone e caffè!*

*Terence: Ah, ho capito. Allora fammene uno misto: mettimi la fragola, il cioccolato, la vaniglia, il limone e il caffè. Charlie, mi raccomando il pistacchio, eh.”*

Pari e dispari

Si mostrano le prospettive future di ricerca nell’area dove si è svolto il lavoro. Talvolta questa sezione può essere l’ultima sottosezione della precedente. Nelle conclusioni si deve richiamare l’area, lo scopo della tesi, cosa è stato fatto, come si valuta quello che si è fatto e si enfatizzano le prospettive future per mostrare come andare avanti nell’area di studio.





# Bibliografia

- [1] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.



## Appendice A

# Documentazione del progetto logico

Documentazione del progetto logico dove si documenta il progetto logico del sistema e se è il caso si mostra la progettazione in grande del SW e dell'HW. Quest'appendice mostra l'architettura logica implementativa (nella Sezione 4 c'era la descrizione, qui ci vanno gli schemi a blocchi e i diagrammi).



## Appendice B

# Documentazione della programmazione

Documentazione della programmazione in piccolo dove si mostra la struttura ed eventualmente l'albero di Jackson.



## Appendice C

### Listato

Il listato (o solo parti rilevanti di questo, se risulta particolarmente esteso)  
con l'autodocumentazione relativa.





## Appendice D

# Il manuale utente

Manuale utente per l'utilizzo del sistema



## Appendice E

# Esempio di impiego

Un esempio di impiego del sistema realizzato.



## Appendice F

# Datasheet

Eventuali Datasheet di riferimento.