

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

Summary	III
Aknowledgments	V
1 Introduction	1
1.1 Inquadramento generale	1
1.2 Breve descrizione del lavoro	1
1.3 Struttura della tesi	2
2 State Of The Art	3
2.1 Deep Learning	3
2.1.1 Artificial Neural Networks	3
2.1.2 Convolutional Neural Networks	3
2.2 Reinforcement Learning	4
2.2.1 Markov Decision Processes	4
2.2.2 Sequential Decision Making	4
2.2.3 Value-based optimization	4
2.2.4 Policy-based optimization	4
2.3 Deep Reinforcement Learning	4
2.3.1 Deep Q-Learning	5
2.3.2 Asynchronous Advantage Actor Critic	5
3 Effects of Feature Selection on Agent Performance	7
3.1 Preliminary Definitions	7
3.2 Tree-based Recursive Feature Selection	7
3.3 Impact of RFS on Agent Performance	7
3.4 RFS on Deep Features	7
4 Technical Details and Implementation	9
4.1 Atari Games	9
4.2 Deep Neural Network for Feature Extraction	9

4.3	Recursive Feature Selection	9
4.4	Code Implementation	9
5	Experimental Results	11
6	Conclusions and Future Developments	13
6.1	Future Developments	13
	References	15

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 frontesizio 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

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.

2.1.1 Artificial Neural Networks

Feed-forward Artificial Neural Networks (ANN) are function approximators inspired by the connected structure of neurons and synapses in the animal brain.

2.1.2 Convolutional Neural Networks

Convolutional Neural Networks (CNN) are a type of ANN inspired by the visual cortex in animal brains. CNNs exploit spatially-local correlations in the neurons of adjacent layers through the use of a *receptive field*, a set of

weights which is used to transform a local subset of the input neurons of a layer.

Autoencoders

Autoencoders are a type of ANN which are used to learn a sparse and compressed representation of the input space, by sequentially compressing and reconstructing the inputs under some sparsity constraint (typically by minimizing the $L1$ norm of the activations in the innermost hidden layers).

2.2 Reinforcement Learning

2.2.1 Markov Decision Processes

Markov Decision Processes are discrete-time, stochastic control processes defined as 5-tuples (S, A, P, R, γ) , where:

- S
- A
- P
- R
- γ

2.2.2 Sequential Decision Making

2.2.3 Value-based optimization

Q-learning

Fitted Q-Iteration

2.2.4 Policy-based optimization

Actor-Critic

2.3 Deep Reinforcement Learning

Deep Reinforcement Learning is the application of RL's sequential decision making techniques and DL's abstract representation capabilities to control problems, usually on MDPs where the state representation is too complex to be tractable with classic dynamic programming techniques.

2.3.1 Deep Q-Learning

2.3.2 Asynchronous Advantage Actor Critic

Capitolo 3

Effects of Feature Selection on Agent Performance

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.

3.1 Preliminary Definitions

3.2 Tree-based Recursive Feature Selection

3.3 Impact of RFS on Agent Performance

3.4 RFS on Deep Features

Capitolo 4

Technical Details and Implementation

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

4.1 Atari Games

4.2 Deep Neural Network for Feature Extraction

4.3 Recursive Feature Selection

4.4 Code Implementation

Capitolo 5

Experimental Results

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 6

Conclusions and Future Developments

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.

6.1 Future Developments

Bibliografia

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