

DEPARTMENT OF COMPUTER SCIENCE
SERIES OF PUBLICATIONS A
REPORT A-2014-0

Example Thesis with a Long Title

John Smith

*To be presented in ... text of a long permission notice. Text of
a long permission notice. Text of a long permission notice. Text
of a long permission notice. Text of a long permission notice.
Text of a long permission notice.*

UNIVERSITY OF HELSINKI
FINLAND

Supervisor

John Smythe, University of Helsinki, Finland

Pre-examiners

Eva Johnson, University of Edinburgh, United Kingdom

Matti Suomalainen, University of Turku, Finland

Opponent

John Smithe, University of Amsterdam, The Netherlands

Custos

Petra Nielsen, University of Helsinki, Finland

Contact information

Department of Computer Science
P.O. Box 68 (Gustaf Hållströmin katu 2b)
FI-00014 University of Helsinki
Finland

Email address: info@cs.helsinki.fi

URL: <http://cs.helsinki.fi/>

Telephone: +358 2941 911, telefax: +358 9 876 4314

Copyright © 2014 John Smith

ISSN 1238-8645

ISBN 000-00-0000-0 (paperback)

ISBN 000-00-0000-0 (PDF)

Computing Reviews (1998) Classification: A.0, C.0.0

Helsinki 2014

Unigrafia

Example Thesis with a Long Title

John Smith

Department of Computer Science
P.O. Box 68, FI-00014 University of Helsinki, Finland
John.Smith@cs.helsinki.fi
<http://cs.helsinki.fi/John.Smith/>

PhD Thesis, Series of Publications A, Report A-2014-0
Helsinki, September 2014, 7 pages
ISSN 1238-8645
ISBN 000-00-0000-0 (paperback)
ISBN 000-00-0000-0 (PDF)

Abstract

This is a sample sentence that should look like normal text, and this is another. This is a sample sentence that should look like normal text, and this is another. This is a sample sentence that should look like normal text, and this is another.

This is a sample sentence that should look like normal text, and this is another. This is a sample sentence that should look like normal text, and this is another.

Computing Reviews (1998) Categories and Subject Descriptors:

A.0 Example Category
C.0.0 Another Example

General Terms:

thesis, example, another example, still more examples, more and more examples

Additional Key Words and Phrases:

example, an example phrase with many words

Acknowledgements

This is a sample sentence that should look like normal text, and this is another. This is a sample sentence that should look like normal text, and this is another. This is a sample sentence that should look like normal text, and this is another.

Contents

1	Reinforcement Learning	1
1.1	Markov Decision Process	1
1.2	Partially Observable Markov Decision Process	1
1.3	Dynamic Programming	1
1.4	Reinforcement Learning Methods	1
1.4.1	Temporal Difference Learning	1
1.4.2	Q-Learning	1
1.4.3	Adaptive Heuristic Critic	1
1.4.4	Prioritised Sweeping	1
1.4.5	Policy Gradient Methods	1
1.5	Classification of the Regarded RL Problems	1
1.5.1	High-Dimensionality	1
1.5.2	Partial-Observability	1
1.5.3	Continuous State and Action Spaces	1
1.5.4	Data-Efficiency	1
2	Recurrent Neural Networks	3
2.1	Feedforward Neural Networks	4
2.2	Recurrent Neural Networks	4
2.2.1	Finite Unfolding in Time	4
2.2.2	Overshooting	4
2.2.3	Dynamical Consistency	4
2.3	Universal Approximation	4
2.3.1	Approximation by FFNN	4
2.3.2	Approximation by RNN	4
2.4	Training of RNN	4
2.4.1	Shared Weight Extended Backpropagation	4
2.4.2	Learning Methods	4
2.4.3	Learning Long-Term Dependencies	4
2.5	Improved Model-Building with RNN	4

2.5.1	Handling Data Noise	4
2.5.2	Handling the Uncertainty of the Initial State	4
2.5.3	Optimal Weight Initialisation	4
3	Prior Arts of Combining RNN and RL	5
3.1	Neural Actor-Critic(idasi's group)	5
3.2	LSTM with POMDP objective function	5
3.3	PhD thesis, by Remi Coulom ?	5
3.4	DQN?	5
3.5	Hybrid Approch(RL with RNN)	5
3.6	Recurrent Models of Visual Attention?	5
3.7	stanley gecco021 2002?	5
4	Experiment	7
4.1	RNN(LSTM) Implementation	7
4.2	Cart-pole Balancing Simulator	7
4.3	Learning a task of stacking wooden blocks	7
	References	9

Chapter 1

Reinforcement Learning

1.1 Markov Decision Process

1.2 Partially Observable Markov Decision Process

1.3 Dynamic Programming

1.4 Reinforcement Learning Methods

1.4.1 Temporal Difference Learning

1.4.2 Q-Learning

1.4.3 Adaptive Heuristic Critic

1.4.4 Prioritised Sweeping

1.4.5 Policy Gradient Methods

1.5 Classification of the Regarded RL Problems

1.5.1 High-Dimensionality

1.5.2 Partial-Observability

1.5.3 Continuous State and Action Spaces

1.5.4 Data-Efficiency

Chapter 2

Recurrent Neural Networks

2.1 Feedforward Neural Networks

2.2 Recurrent Neural Networks

2.2.1 Finite Unfolding in Time

2.2.2 Overshooting

2.2.3 Dynamical Consistency

2.3 Universal Approximation

2.3.1 Approximation by FFNN

2.3.2 Approximation by RNN

2.4 Training of RNN

2.4.1 Shared Weight Extended Backpropagation

2.4.2 Learning Methods

2.4.3 Learning Long-Term Dependencies

2.5 Improved Model-Building with RNN

2.5.1 Handling Data Noise

2.5.2 Handling the Uncertainty of the Initial State

2.5.3 Optimal Weight Initialisation

Chapter 3

Prior Arts of Combining RNN and RL

- 3.1 Neural Actor-Critic(idasi's group)
- 3.2 LSTM with POMDP objective function
- 3.3 PhD thesis, by Remi Coulom ?
- 3.4 DQN?
- 3.5 Hybrid Approach(RL with RNN)
- 3.6 Recurrent Models of Visual Attention?
- 3.7 stanley gecco021 2002?

Chapter 4

Experiment

4.1 RNN(LSTM) Implementation

4.2 Cart-pole Balancing Simulator

4.3 Learning a task of stacking wooden blocks

This is a sample sentence that should look like normal text, and this is another. This is a sample sentence that should look like normal text, and this is another. This is a sample sentence that should look like normal text, and this is another.

Theorem 4.1 *This is a sample sentence that should look like normal text, and this is another:*

$$y = x + 3$$

Proof. This is a sample sentence. \square

References

- [1] Smith, John. Example Document.