Imperial College of Science, Technology and Medicine
Department of Computing

On the Feasibility of Using Convolutional Variational Autoencoders to Advance Deep Symbolic Reinforcement Systems

D. G. Sherburn

Submitted in part fulfilment of the requirements for the degree of
Master of Engineering in Joint Mathematics and Computing
of Imperial College, June 2017

Acknowledgements

I would like to express (whatever feelings I have) to:

- My supervisor
- My second supervisor
- Other researchers
- My family and friends

Dedication

Dedication here.

‘Quote text here.’

Guy Quoted

Contents

Acknowledgements **i**

1 Introduction **1**

1.1 Motivation and Objectives 1

1.2 Contributions 3

2 Background Theory **4**

2.1 Introduction 4

3 Conclusion **5**

3.1 Summary of Thesis Achievements 5

3.2 Applications 5

3.3 Future Work 5

Bibliography **5**

List of Tables

List of Figures

1.1 May 1997: Gary Kasparov makes his first move against IBM's Deep Blue. Deep Blue would later emerge the victor in the best of six games; the first defeat of a reigning world chess champion by computer. [5] 2

1.2 March 2016: Lee Sedol plays the first move of game three against AlphaGo. AlphaGo won four of five games. [4] 2

1.3 Overview of deep symbolic reinforcement learning system architecture. **A**: The neural back end maps high-dimensional raw input data to a low-dimensional compositionally structured symbolic representation. **B**: The low-dimensional compositionally structured symbolic representation. **C**: Reinforcement learning of mapping from symbolic representation to action with maximum expected reward over time. *Source: Garnelo et al.* [1]. 2

Chapter 1

Introduction

1.1 Motivation and Objectives

A long term goal of artificial intelligence (AI) is the development of artificial general intelligence (AGI). A number of theoretical frameworks have been presented as formalizations of what it means to achieve AGI, notably by Hutter [2].

Reinforcement learning is fundamental in Hutter's framework, and, as of March 2016, deep reinforcement learning (DRL) systems have mastered a wide range of tasks, including wide range of Atari 2600 games and Go [3, 6].

Though DRL systems have been remarkably successful, they still have a number of drawbacks [1]:

1. **Slow to learn**. Deep neural networks require large data sets and are therefore slow to learn.
2. **Fail to transfer past experience**. They often fail to perform well on tasks very similar to those they have mastered.
3. **Inability to reason abstractly**. They fail to exploit statistical regularities in the training data by using high-level processes like planning or causal reasoning.

4. **Hard to reason about**. It's often difficult to understand why the DRL system chose the action it did.

Figure 1.1: May 1997: Gary Kasparov makes his first move against IBM's Deep Blue. Deep Blue would later emerge the victor in the best of six games; the first defeat of a reigning world chess champion by computer. [5]

Figure 1.2: March 2016: Lee Sedol plays the first move of game three against AlphaGo. AlphaGo won four of five games. [4]

Deep symbolic reinforcement learning (DSRL) is a recent advance which overcomes all of these drawbacks without the drawbacks from classical symbolic AI, namely the symbol grounding problem [1].

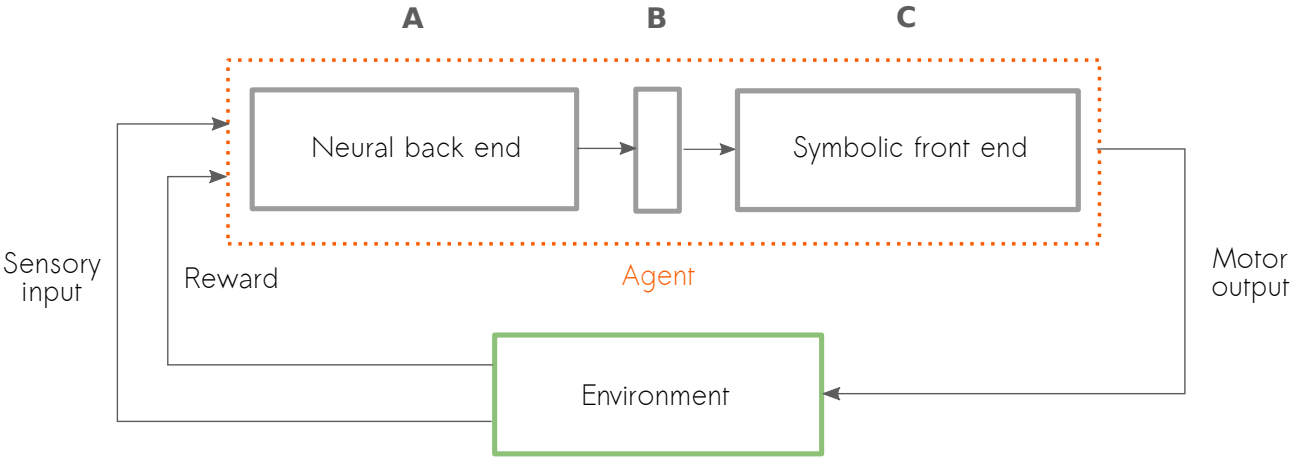

Figure 1.3: Overview of deep symbolic reinforcement learning system architecture. **A**: The neural back end maps high-dimensional raw input data to a low-dimensional compositionally structured symbolic representation. **B**: The low-dimensional compositionally structured symbolic representation. **C**: Reinforcement learning of mapping from symbolic representation to action with maximum expected reward over time. *Source: Garnelo et al.* [1].

The architecture of the DSRL is shown in Figure 1.3. The raw high-dimensional input, say an

image, is mapped to a low-dimensional symbolic representation. This is done by passing the high-dimensional input through a series of convolutional layers and extracting the activation spectra in the latent space. These spectra are then used to classify the

relies on the unsupervised extraction of disentangled features, allowing for transfer learning and high-level cognitive processes. However, the unsupervised extraction of features from a wide range of scenes is still a challenge in AI research [?]. Fortunately methods are getting better, and the first unsupervised scalable model β-VAE was developed recently.

1.2 Contributions

Contributions here.

Chapter 2

Background Theory

2.1 Introduction

Text of the Background.

Chapter 3

Conclusion

3.1 Summary of Thesis Achievements

Summary.

3.2 Applications

Applications.

3.3 Future Work

Future Work.

Bibliography

[1] M. Garnelo, K. Arulkumaran, and M. Shanahan. Towards Deep Symbolic Reinforcement Learning. *arXiv Preprint*, pages 1–13, 2016.

[2] M. Hutter. *Universal Artificial Intelligence*, volume 1. 2005.

[3] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.

[4] D. Ormerod. Lee Sedol plays the first move of game three against AlphaGo., 2016.

[5] R. Rosen. IBM's Deep Blue vs Gary Kasparov, 2012.

[6] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.