Project Proposal
Progressive Neural Networks

Dennis Verheijden s4455770 Joost Besseling s4796799

January 9, 2018

Project Description

One of the many unsolved problems of Artificial Intelligence is catastrophic forgetting, the knowledge of task A is forgotten when trained on task B. This is what we aim to address in this project.

We will do so by implementing a Neural Network architecture proposed by Google Deepmind in 2016. Here they proposes a novel way to tackle this problem by using a Neural Network architecture that they called a Progressive Neural Network (PNN) [2]. For our project we will be implementing this type of network and we will train it to play atari games, using the openAi gym [1].

We plan on implementing it in Chainer [3]. With a network similar in the original paper, where they use two convolutional layers and one fully connect layer [1]. The latter being tuned toward the task domain. The input to the neural networks are the raw pixel values of the rendered openAI scene.

We aim to replicate some of the papers findings by:

1. Implement a Progressive Neural Network
2. Train it on a source domain, which will be a Atari game (Pong probably)
3. Then train it on Pong variants and find out whether we can replicate the papers findings:
    - Does the Progressive Neural Network converge faster than a regular Neural Network (i.e. is there transfer of knowledge)
    - Try to compute the Average Pertubation Sensitivity to replicate the transfer matrix.

References

[1] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym, 2016. URL `https://github.com/openai/gym`.

[2] Andrei A Rusu, Neil C Rabinowitz, Guillaume Desjardins, Hubert Soyer, James Kirkpatrick, Koray Kavukcuoglu, Razvan Pascanu, and Raia Hadsell. Progressive neural networks. *arXiv preprint arXiv:1606.04671*, 2016. URL `https://arxiv.org/pdf/1606.04671`.

[3] Seiya Tokui, Kenta Oono, Shohei Hido, and Justin Clayton. Chainer: a next-generation open source framework for deep learning. In *Proceedings of Workshop on Machine Learning Systems (LearningSys) in The Twenty-ninth Annual Conference on Neural Information Processing Systems (NIPS)*, 2015. URL `http://learningsys.org/papers/LearningSys_2015_paper_33.pdf`.