Learning to Reason

PhD Thesis Proposal

Wojciech Zaremba WOJ.ZAREMBA@GMAIL.COM

Advisors:
Rob Fergus FERGUS@CS.NYU.EDU
Yann LeCun YANN@CS.NYU.EDU
New York University

Abstract

Neural networks proven to be a very powerful models for object recognition (Krizhevsky et al., 2012), natural language processing (Mikolov, 2012), speech recognition (Graves et al., 2013), and many others (Sutskever et al., 2014). However, there is still a huge gap between them, and an intelligent systems. I identify several potential unaddressed skills, which intelligent systems should possess: (1) reasoning abilities, (2) capability to integrate with external interfaces and (3) small sample complexity. My research focuses on tackling this problems.

Input:
```
f=(8794 if 8887<9713 else (3*8334))
print((f+574))
```
Target: 9368.
Model prediction: 9368.

Figure 1. aaa

1. Introduction

It's clear that by improving performance of current statistical learning systems, we won't be able to make them intelligent. Even if our object recognition system would yield 0% of prediction error, they wouldn't be intelligent. Same applies to speech recognition systems, machine translation and others. This work asks what skills are necessary for statistical learning system to become "intelligent". Moreover, it attempts to address this remaining unaddressed skills.

I believe, that crucial, poorly addressed skills that intelligent system has to poses are (1) reasoning abilities, (2) capability to integrate with external interfaces and (3) small sample complexity. I would like to address all this problems within seamless system. The same system should be used across different tasks, and should be able to emulate simpler models.

I have partially addressed some of proposed problems. I am referring to my work, and I am defining future goals.

2. Reasoning abilities

Reasoning - "the process of forming conclusions, judgments, or inferences from facts or premises" *.

There exists many rule based reasoning systems. However, intelligent system reasoning cannot be based only on predefined rules. It has to be based on pattern matching, and application of learnt heuristic algorithms. I would like to build a statistical reasoning system.

There are many domains where it is crucial. Domains of my interest include learning object relations in computer vision, proving mathematical theorems, and learning about computer programs.

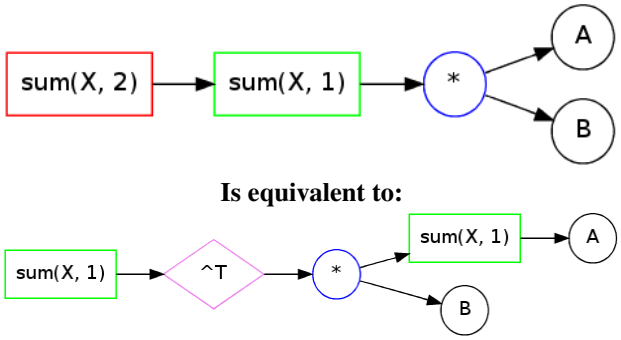

Figure 2. (Zaremba et al., 2014)

Figure 3. Jenny loves to play soccer but she is worried that Mike will kick the ball too hard.

2.1. Reasoning in computer programs

2.2. Reasoning in mathematics

2.3. Reasoning in computer vision

3. Interface learning

Contemporary machine learning systems are closed in a box, and cannot interact with external interfaces.

Initial advances in machine learning, where lead by engineering features, e.g histogram of colors, SIFT (Scale-invariant feature transform) features. This approach has its limitations, and its fragile. Moreover, it requires a human expert to train the system for the every new domain. Ideal system should have access to external interfaces that might give access to information, or simplify it processing. Interfaces that I have in mind are following (1) grid that helps to process images, (2) search engine, (4) Linux syscalls, or (5) execution environment like python interpreter. External interfaces are not-differentiable, and their state space is massive. This obstacles could potentially be addressed by learning a differentiable model that describes them (e.g. neural network). Neural network would simulate such external interface in a model-based reinforcement learning approach.

4. Toward all tasks one-shot learning

Current deep learning systems suffer of large sample complexity. Such high sample complexity hinders potential use of the systems in online learning systems (e.g. robots). Its expected that during the initial phase of learning any system without prior knowledge would need to consume a large number of samples. We hope that over the time of training, sample complexity should drop. However, this is not observed in current systems. I would like to build a meta-optimizer that could overcome this limitation. Such optimizer would consume gradients of a neural network, and would decide on the next update step. Optimizer itself could be parameterized with a neural network. Proper weights could simulate any first order, gradient-based, learning algorithm like SGD, momentum, LBFGs etc. This implies that meta-optimizer subsumes all first order, gradient-based optimization techniques. Trained meta-optimizer could update the network in a much more clever way, and a single sample could provide enough knowledge.

5. Discussion

Tackling aforementioned problems would take us much closer to the real intelligent systems, and defines for me

*Definition from http://dictionary.reference.com/browse/reasoning

the three main pillars of artificial intelligence. However, there are many other problems, which would need to be solved / integrated within such system to make it fully intelligent, e.g. navigation, learning by imitation, cooperation, and many others.

References

Graves, Alex, Mohamed, Abdel-rahman, and Hinton, Geoffrey. Speech recognition with deep recurrent neural networks. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pp. 6645–6649. IEEE, 2013.

Krizhevsky, Alex, Sutskever, Ilya, and Hinton, Geoffrey E. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pp. 1097–1105, 2012.

Mikolov, Tomáš. *Statistical language models based on neural networks*. PhD thesis, Ph. D. thesis, Brno University of Technology, 2012.

Sutskever, Ilya, Vinyals, Oriol, and Le, Quoc V. Sequence to sequence learning with neural networks. *arXiv preprint arXiv:1409.3215*, 2014.

Zaremba, Wojciech, Kurach, Karol, and Fergus, Rob. Learning to discover efficient mathematical identities. *arXiv preprint arXiv:1406.1584*, 2014.