Differentiable Neural Computers and their abilities

Report by:
Denys Kovalenko
Gagandeep Singh

University of Tartu

Introduction

Machine learning have seen many breakthroughs in recent years due to reduced cost of computing power and enhanced algorithms. But in most cases foundation of modern machine learning approaches was developed at least decade ago. This project aims to cast light to new concept in machine learning - enhance possibilities of neural networks with external memory. Maybe this has roots in fact that people have wanted to be able to connect external devices to brain and learn new languages, skills instantly by "downloading" data into their brain. This is still far from reality, but DeepMind shows that they are working on similar problem in artificial neural networks world - "let's connect external memory that will work like RAM to ANN and see what happens". In this combination neural net performs function of algorithm, which can be learned with gradient descent, and it can write to external memory and read from it, just like CPUs in computers work with RAM.

Goal of this project is to find out for what Differential Neural Computers can be used for, provide understanding how they can be trained and how they perform tasks. Also, we are briefly explaining idea that lays behind DNC.

Finally, we will demonstrate how DNC performs on real data, and describe characteristics related to functioning of it's memory, and describe limitations of DNCs.

Theory behind Differential Neural Computers

To get better understanding of DNCs we should provide brief overview of key components that forms foundation of what DNC is.

Turing machine

Turing machine is abstract machine, represented by moving head that is connected to infinite tape, and can perform read/write operation, and move along this tape. Set of those operations is defined in finita table with instructions.
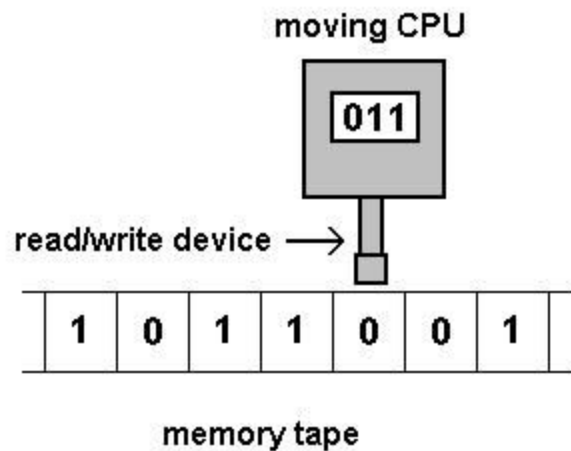

Image 1. representation of turing machine (http://ideonexus.com)

Given any computer algorithm, such turing machine can be constructed that it will implement this algorithm.
//todo - provide references and reasoning. Add more information that can be related later to DNC
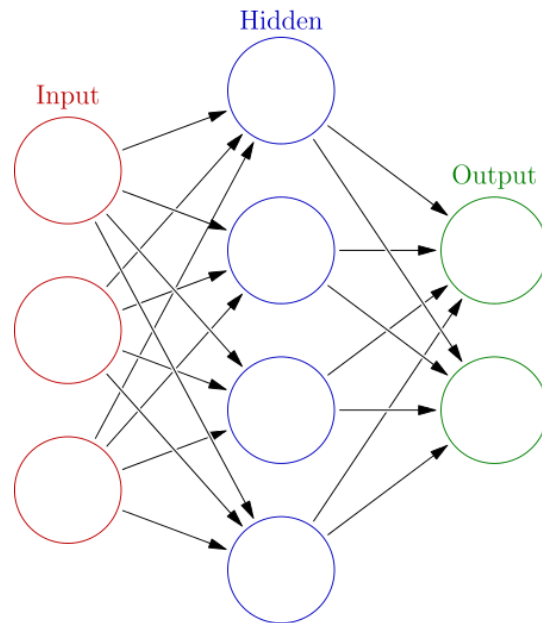
Artificial Neural network

Image 2. Feed-forward ANN

Artificial neural network is a computational model that mimics how neurons work in brain of humans and animals. It's heavily used in machine learning for image processing, speech recognition and others tasks for which no strict algorithm can be defined so it should be learned from datasets. "Knowledge" in ANN is represented in weights of connections between neurons.

DNC - Differential Neural Computer

DNC is a machine learning model, that have combined to approaches and took best from 2 different worlds - it can learn from data, thus it can be trained via supervised learning to solve different task, and it has ability that all computers have - store information in variables and form complex structures, and store data over long time periods without interference.

Structure of Differentiable Neural Computer
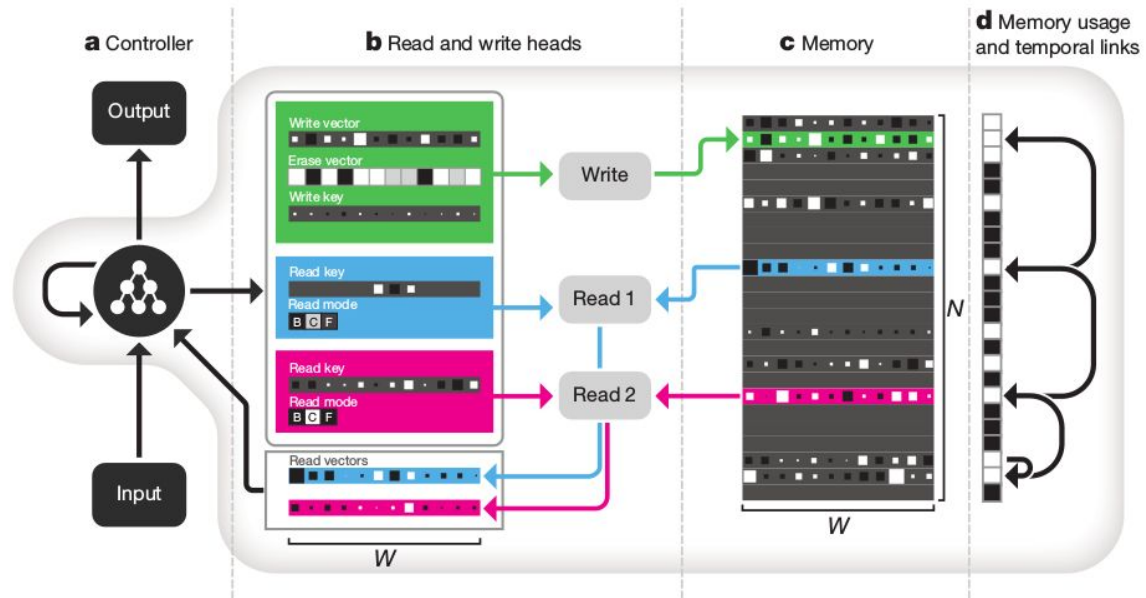
Picture 3. Structure of differential neural machine

Explanation of how DNC works on "copy" example

/*in development*/

Training and testing on real problem

/*Dataset is being chosed. Right now choosing between basic algorithms on graphs and trying to find NLP related tasks that we can solve*/

Conclusions