Memory based Model

A Overview

Lujun Zhao

2016/03/10

School of Mathematical Sciences
Fudan University

Memory Networks[Weston et al. 2015]

A MEMORY NETWORK consists of a memory $\mathbf{m}$ and four components I, G, O and R as follows:

- I: (input feature map) – converts the incoming input to the internal feature representation.
- G: (generalization) – updates old memories given the new input.
- O: (output feature map) – produces a new output, given the memories.
- R: (response) – converts the output into the response format desired.

- **Large - scale QA**

  Dataset: ReVerb ClueWeb09 Extractions (14M statements, stored as (subject, relation, object) triples)

Method	F1
(Fader et al., 2013)	0.54
(Bordes et al., 2014b)	0.73
MemNN (embedding only)	0.72
MemNN (with BoW features)	0.82

Figure: Results on the large-scale QA task

Experiments

- **Simulated world QA**

  Dataset: generated 7k statements and 3k questions from the simulator

- **Combining simulated data and large - scale QA**

Neural Turing Machines[Graves et al. 2014]

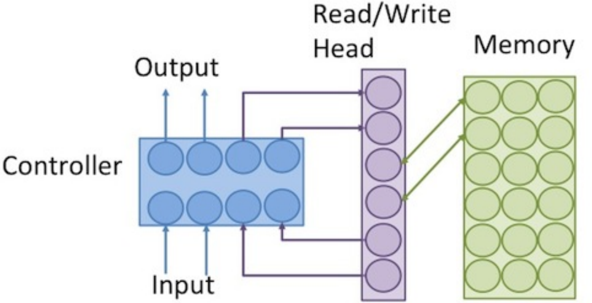

Figure: Neural Turing Machine Architecture

Experiments

- **Copy**
  Dataset: unknown (sequence of random binary vectors)
- **Repeat Copy**
- **Associative Recall**
- **Dynamic N-Grams**
- **Priority Sorts**

End-To-End Memory Networks[Sukhbaatar et al. 2015]

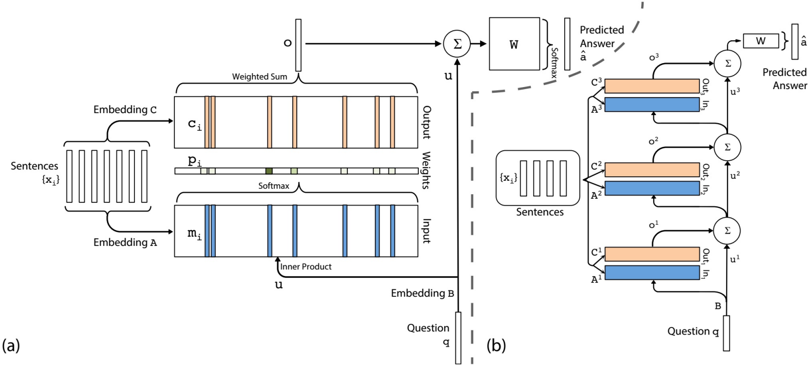

Figure: End-To-End Memory Networks

Experiments

- **Synthetic Question and Answering**
  Dataset: Facebook bAbI Dataset

- **Language Modeling**
  Dataset: Penn Tree Bank and Text8

Dynamic Memory Networks[Kumar et al. 2015]

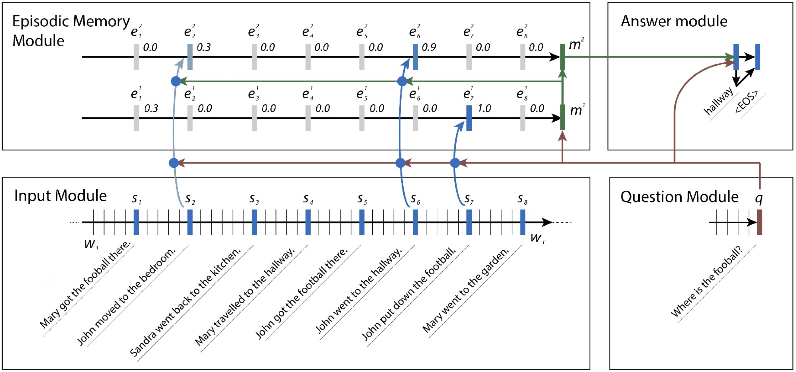

Figure: Dynamic Memory Networks

Experiments

- **Question Answering**
  Dataset: Facebook bAbI Dataset

- **Sentiment Analysis**
  Dataset: Stanford Sentiment Treebank

- **Part-of-Speech Tagging**
  Dataset: Wall Street Journal dataset

Stack RNN[Joulin et al. 2015]

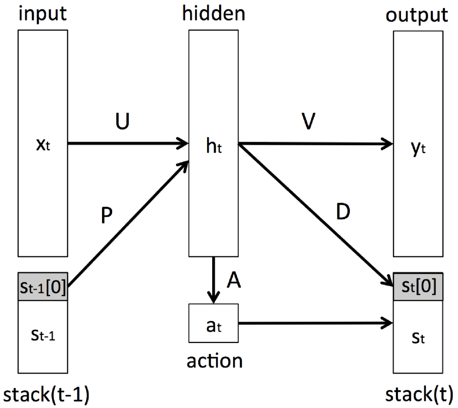

Figure: Neural network extended with push-down stack

Experiments

- **Learning simple algorithmic patterns**
  Dataset: generated by simple algorithms

- **Language modeling**
  Dataset: Penn Tree Bank

Model	Ngram	Ngram + Cache	RNN	LSTM	SRCN [24]	Stack RNN
Validation perplexity	-	-	137	**120**	**120**	124
Test perplexity	141	125	129	**115**	**115**	118

Figure: Comparison of RNN, LSTM, SRCN and Stack RNN on language modeling task

Neural Stack[Grefenstette et al. 2015]

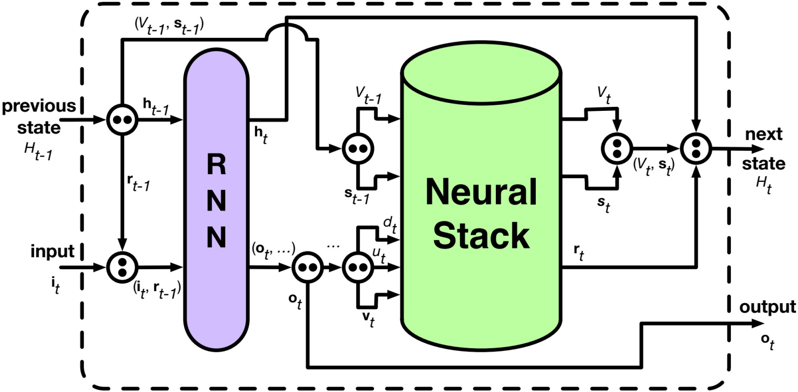

Figure: Illustrating a Neural Stack's Operations, Recurrent Structure, and Control

Experiments

- **Synthetic Transduction Tasks**
  Dataset: randomly generated from a vocabulary of 128 meaningless symbols
  - Sequence Copying
  - Sequence Reversal
  - Bigram flipping
- **ITG Transduction Tasks**
  Dataset: the source and target sequence are jointly generated by **I**nversion **T**ransduction **G**rammars

Structured-Memory NTMs[Zhang et al. 2015]

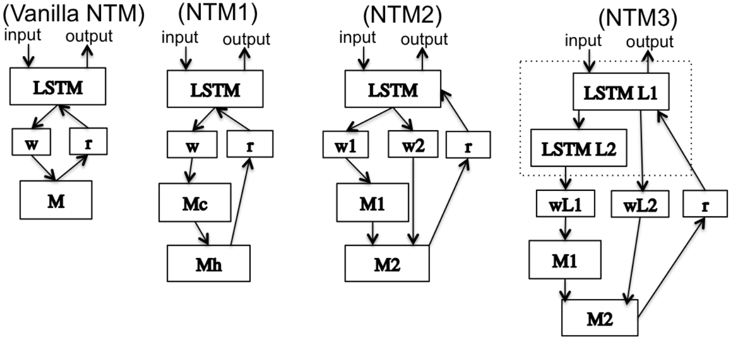

Figure: NTM and NTM variants that use LSTM as controllers

Experiments

The experiments is to show the convergence speed and quality of those three variants, compared to the NTM setting.

- **Copy Task**
  Dataset: randomly generated

- **Associative Recall Task**
  Dataset: randomly generated

Evolving Neural Turing Machines[Greve et al. 2015]

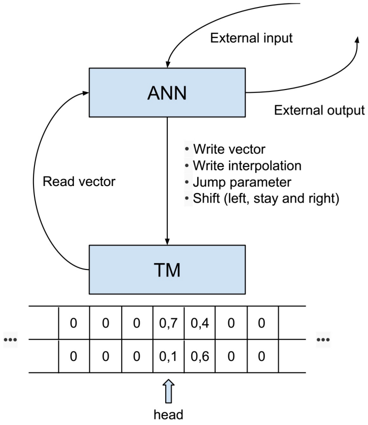

Figure: Evolvable Neural Turing Machine

Experiments

- **Copy Task**
  Dataset: unknown

- **T-Maze Task**
  Dataset: unknown

References I

A. Graves, G. Wayne, and I. Danihelka.
Neural turing machines.
arXiv preprint arXiv:1410.5401, 2014.

E. Grefenstette, K. M. Hermann, M. Suleyman, and P. Blunsom.
Learning to transduce with unbounded memory.
In *Advances in Neural Information Processing Systems*, pages 1819–1827, 2015.

R. B. Greve, E. J. Jacobsen, and S. Risi.
Evolving neural turing machines.

References II

A. Joulin and T. Mikolov.
Inferring algorithmic patterns with stack-augmented recurrent nets.
In *Advances in Neural Information Processing Systems*, pages 190–198, 2015.

A. Kumar, O. Irsoy, J. Su, J. Bradbury, R. English, B. Pierce, P. Ondruska, I. Gulrajani, and R. Socher.
Ask me anything: Dynamic memory networks for natural language processing.
arXiv preprint arXiv:1506.07285, 2015.

References III

- S. Sukhbaatar, J. Weston, R. Fergus, et al.
  End-to-end memory networks.
  In *Advances in Neural Information Processing Systems*, pages 2431–2439, 2015.
- J. Weston, S. Chopra, and A. Bordes.
  Memory networks.
  arXiv preprint arXiv:1410.3916, 2014.
- W. Zhang and Y. Yu.
  Structured memory for neural turing machines.
  arXiv preprint arXiv:1510.03931, 2015.