

Is attention really all you need?

Kyle Roth

Brigham Young University
kylrth@gmail.com

Contents

1 Introduction	1
2 Theoretical generalization	1
2.1 Self-attention	1
2.2 Attention between pairs	1
2.3 Attention on arbitrary graphs	1
2.4 Attention as a lookup	1
2.5 Active memory	1
3 Application	1
4 Interpretation	1

Abstract

This is a wonderful abstract.

1 Introduction

The attention mechanism is an exciting development in the artificial intelligence community. The concept is simple, and rooted in our understanding of attention in biological intelligence [Larochelle and Hinton, 2010; Hassabis *et al.*, 2017]. Attention mechanisms also lend themselves to more intuitive interpretation, an attractive feature as concerns about model interpretability gain traction.

Ironically, great attention has been given to attention mechanisms since their conception in 2015. Initially formulated as a word alignment mechanism for text translation [Bahdanau *et al.*, 2015],

2 Theoretical generalization

2.1 Self-attention

2.2 Attention between pairs

2.3 Attention on arbitrary graphs

2.4 Attention as a lookup

2.5 Active memory

Neural memory resources aren't the same thing [Graves *et al.*, 2014]?

3 Application

4 Interpretation

References

- [Bahdanau *et al.*, 2015] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *CoRR*, abs/1409.0473, 2015.
- [Graves *et al.*, 2014] Alex Graves, Greg Wayne, and Ivo Danihelka. Neural turing machines. *CoRR*, abs/1410.5401, 2014.
- [Hassabis *et al.*, 2017] Demis Hassabis, Dhharshan Kumaran, Christopher Summerfield, and Matthew Botvinick. Neuroscience-inspired artificial intelligence. *Neuron*, 95:245–258, 2017.
- [Larochelle and Hinton, 2010] Hugo Larochelle and Geoffrey Hinton. Learning to combine foveal glimpses with a third-order boltzmann machine. In *Proceedings of the 23rd International Conference on Neural Information Processing Systems - Volume 1*, NIPS'10, pages 1243–1251, USA, 2010. Curran Associates Inc.