Memory Networks

Advanced Machine Learning for NLP
Jordan Boyd-Graber
INTRODUCTION

Slides adapted from Christopher Olah
What’s wrong with RNNs?
What’s wrong with RNNs?
Learning a regular expression

- Suppose we wanted to learn $a^n b^n$
- Need to encode arbitrarily long dependency in state of RNN
- More importantly, don’t know *when* that information is useful
Role of LSTMs

- Now standard baseline method for sequency labeling and classification tasks
- Like RNNs, hidden layer is critically important
- Can be applied to downstream tasks
- Also can help with vanishing gradient problem