Bidirectional RNNs

Sargur Srihari
srihari@buffalo.edu

Topics in Sequence Modeling

- Overview

1. Unfolding Computational Graphs
2. Recurrent Neural Networks
3. Bidirectional RNNs
4. Encoder-Decoder Sequence-to-Sequence Architectures
5. Deep Recurrent Networks
6. Recursive Neural Networks
7. The Challenge of Long-Term Dependencies
8. Echo-State Networks
9. Leaky Units and Other Strategies for Multiple Time Scales
10. LSTM and Other Gated RNNs
11. Optimization for Long-Term Dependencies
12. Explicit Memory

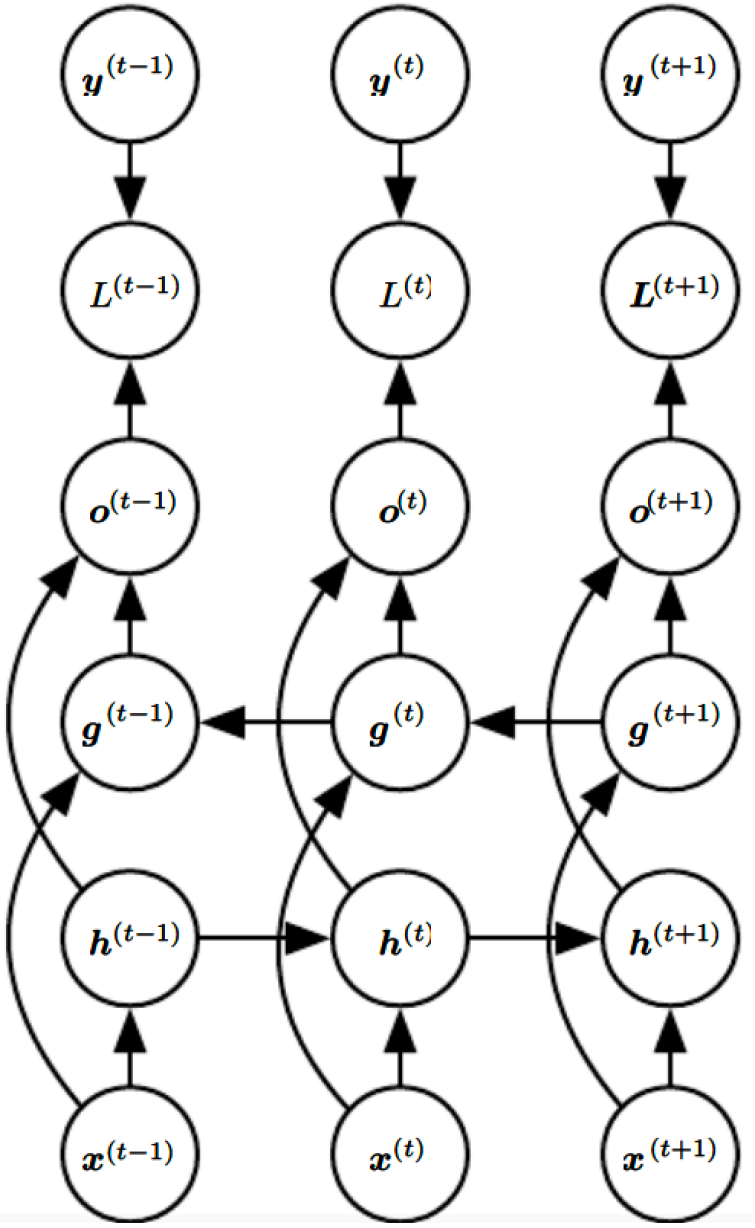
Need for bidirectionality

- In speech recognition, the correct interpretation of the current sound may depend on the next few phonemes because of coarticulation and the next few words because of linguistic dependencies
- Also true of handwriting recognition

A birectional RNN

- Combine an RNN that moves forward through time from the start of the sequence
- Another RNN that moves backward through time beginning from the end of the sequence

A typical bidirectional RNN



Maps input sequences $\boldsymbol{x}$ to target sequences $\boldsymbol{y}$ with loss $L^{(t)}$ at each step t

$\boldsymbol{h}$ recurrence propagates to the right
$\boldsymbol{g}$ recurrence propagates to the left.

This allows output units $\boldsymbol{o}^{(t)}$ to compute a representation that depends both the past and the future