NLP
Deep Learning

Recurrent Neural Networks
the cat
Recurrent Neural Networks

\[ h_t = \sigma(W_h h_{t-1} + W_x x_t) \]
RNN

\[ h_t = \sigma(W_h h_{t-1} + W_x x_t) \]
\[ y_t = \text{softmax}(W_y h_t) \]
RNN

The cat sat

\[ h_0 \xrightarrow{w_h} \sigma \xrightarrow{w_x} x_1 \]
\[ h_1 \xrightarrow{w_h} \sigma \xrightarrow{w_x} x_2 \]
\[ h_2 \xrightarrow{w_h} \sigma \xrightarrow{w_x} x_3 \]

\[ y_3 \xrightarrow{softmax} \]

\[ w_y \]
Updating Parameters of an RNN

Backpropagation through time

The cat sat

Cost
Notes

• RNNs are used to keep “memory”, just like finite-state automata
• They can be used as generators, acceptors, transducers, etc.
Example Application: Machine Translation

Sequence-to-sequence learning (Sutskever, Vinyals, and Le, 2014)
Applications

• Language Modeling (Mikolov 2012)
• Character-level RNNs for text generation
• Semantic Role Labeling:
  – http://www.aclweb.org/anthology/P15–1109
• Dependency parsing:
  (“producing state-of-art dependency parsing results while requiring minimal feature engineering”)

NLP