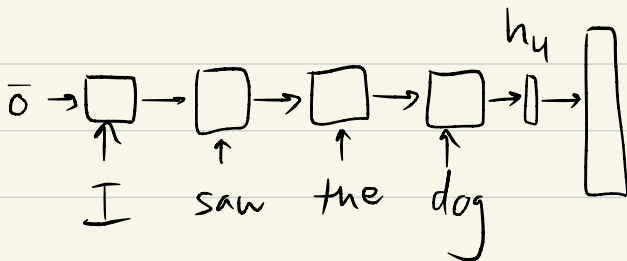
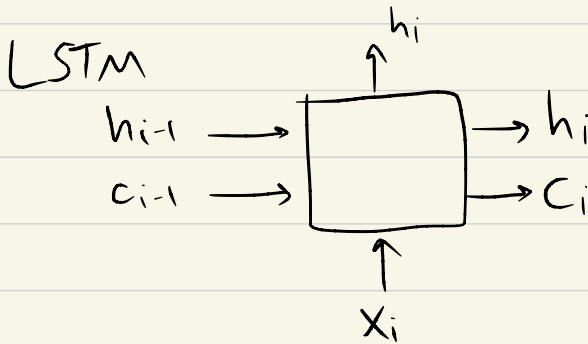


CS 378 Lecture 17

Announcements

- FP proposals back
- A4 due in 12 days

Recap RNNs: models including Elman networks, LSTMs, others like GRU



$$P(w_i | w_1, \dots, w_{i-1}) = \text{softmax}(W h_i)$$

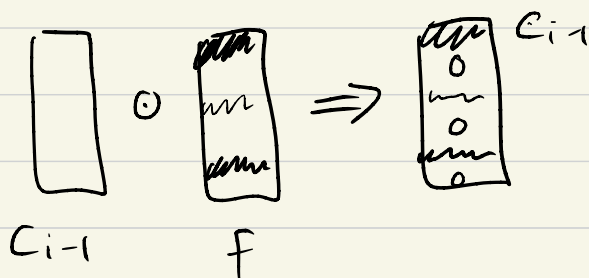
learned $\uparrow$

$|V| \times d$

LSTM: gated connections

$$C_i = C_{i-1} \odot f + \text{func}(x_i) \odot i$$

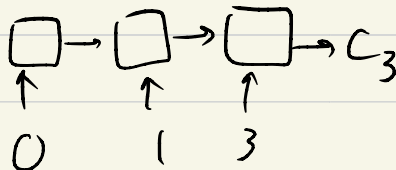
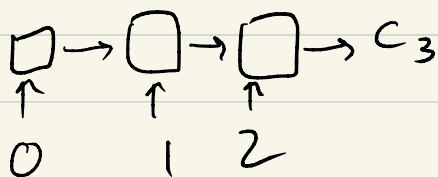
f, i "gate" vectors $\in [0, 1]^d$



d -dim $f = \sigma(W^1 x_i + W^2 h_{i-1})$

Today LSTM code ex
Implementation of RNNs,
+ training
Examples / visualization

Code LSTM w/random params



$$C_i = f \cdot C_{i-1} + \overline{W} X_i$$

Imagine $f \approx [0.5]^d \Rightarrow$

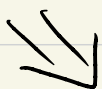
$$f = [1]^d$$

$$C_i = X_i + \frac{1}{2}X_{i-1} + \frac{1}{4}X_{i-2}$$



random case

high bias
case



$$C_i = X_i + X_{i-1} + X_{i-2}$$