



deeplearning.ai

LSTM and GRU are used to capture long range dependencies.

# Recurrent Neural Networks

---

**LSTM** (long short  
term memory) unit

# GRU and LSTM

## GRU

$$\tilde{c}^{<t>} = \tanh(W_c[\Gamma_r * \underline{c}^{<t-1>}, x^{<t>}] + b_c)$$

$$\Gamma_u = \sigma(W_u[\underline{c}^{<t-1>}, x^{<t>}] + b_u)$$

$$\Gamma_r = \sigma(W_r[\underline{c}^{<t-1>}, x^{<t>}] + b_r)$$

$$\underline{c}^{<t>} = \Gamma_u * \tilde{c}^{<t>} + (1 - \Gamma_u) * \underline{c}^{<t-1>}$$

$\underbrace{a^{<t>} = c^{<t>}}$

$\uparrow$   
 $\Gamma_f$

## LSTM

$$\tilde{c}^{<t>} = \tanh(W_c[a^{<t-1>}, x^{<t>}] + b_c)$$

(update)  $\Gamma_u = \sigma(W_u[a^{<t-1>}, x^{<t>}] + b_u)$

(forget)  $\Gamma_f = \sigma(W_f[a^{<t-1>}, x^{<t>}] + b_f)$

(output)  $\Gamma_o = \sigma(W_o[a^{<t-1>}, x^{<t>}] + b_o)$

$$\underline{c}^{<t>} = \Gamma_u * \tilde{c}^{<t>} + \Gamma_f * \underline{c}^{<t-1>}$$

$$\underline{a}^{<t>} = \Gamma_o * \underline{c}^{<t>}$$

# LSTM in pictures

$$\tilde{c}^{<t>} = \tanh(W_c[a^{<t-1>}, x^{<t>}] + b_c)$$

$$\Gamma_u = \sigma(W_u[a^{<t-1>}, x^{<t>}] + b_u)$$

$$\Gamma_f = \sigma(W_f[a^{<t-1>}, x^{<t>}] + b_f)$$

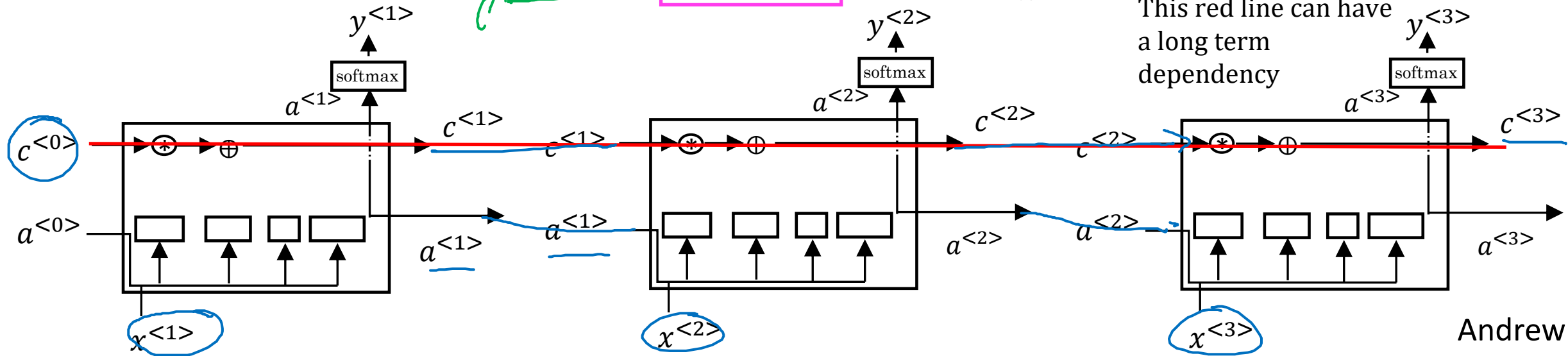
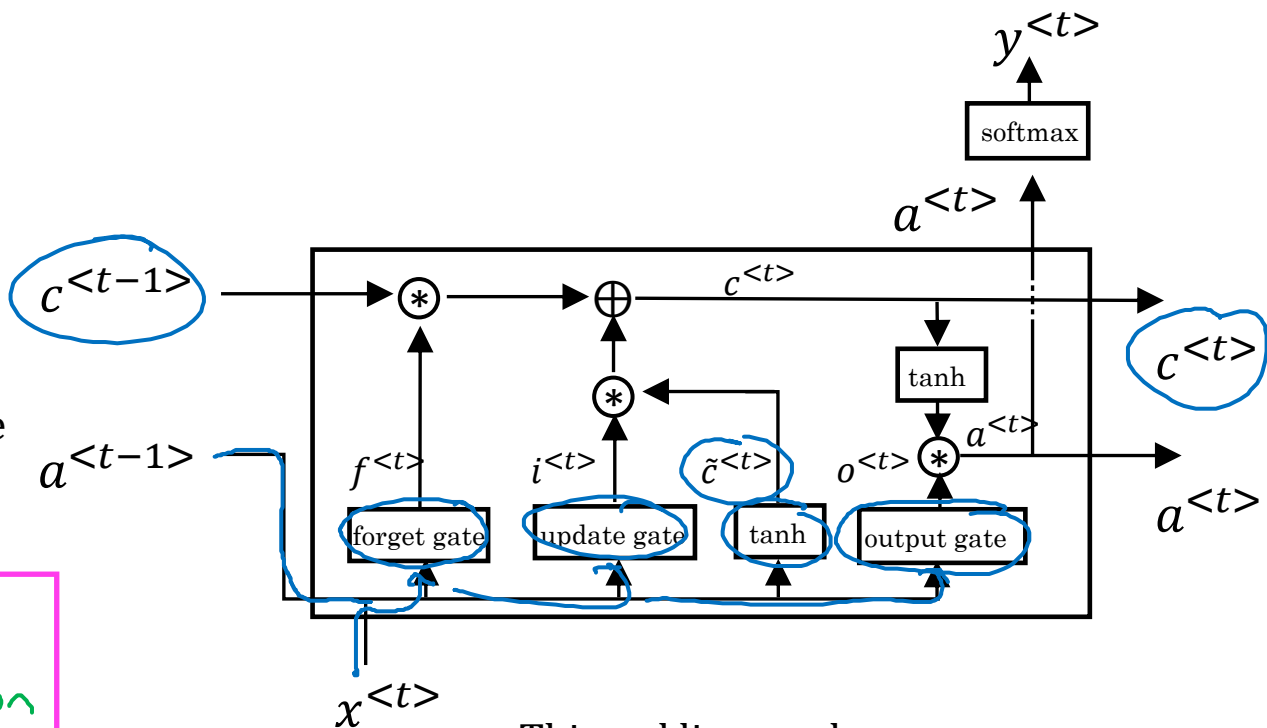
$$\Gamma_o = \sigma(W_o[a^{<t-1>}, x^{<t>}] + b_o)$$

$$c^{<t>} = \Gamma_u * \tilde{c}^{<t>} + \Gamma_f * c^{<t-1>}$$

$$a^{<t>} = \Gamma_o * c^{<t>}$$

Alternative version of LSTM

peephole connection



This red line can have a long term dependency