



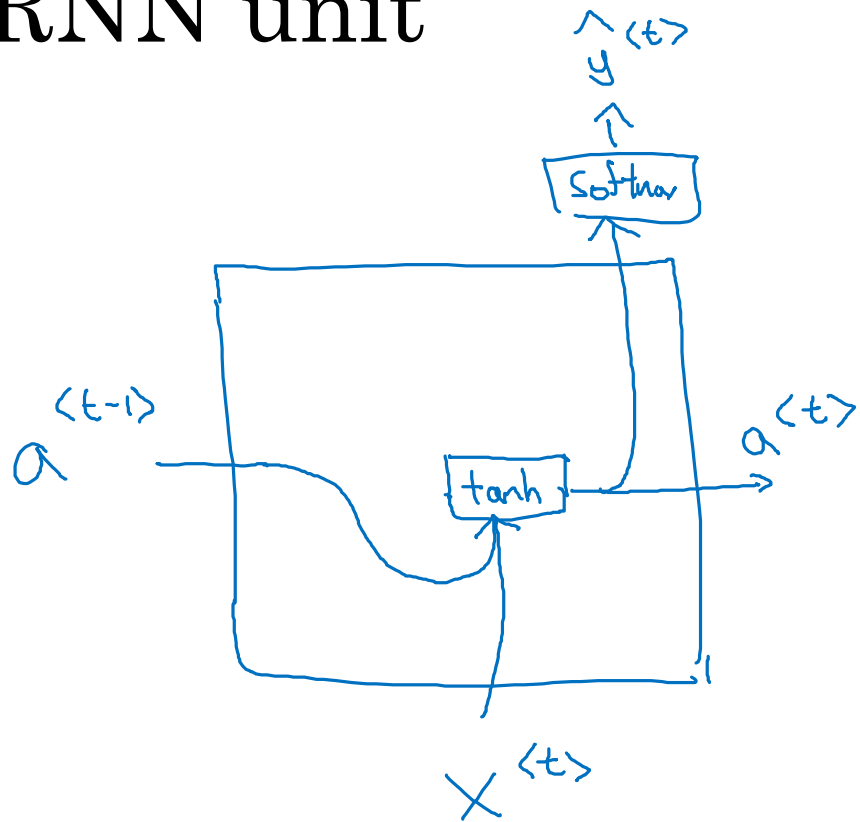
deeplearning.ai

# Recurrent Neural Networks

---

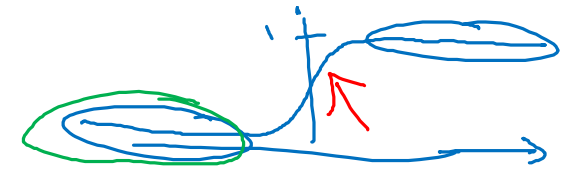
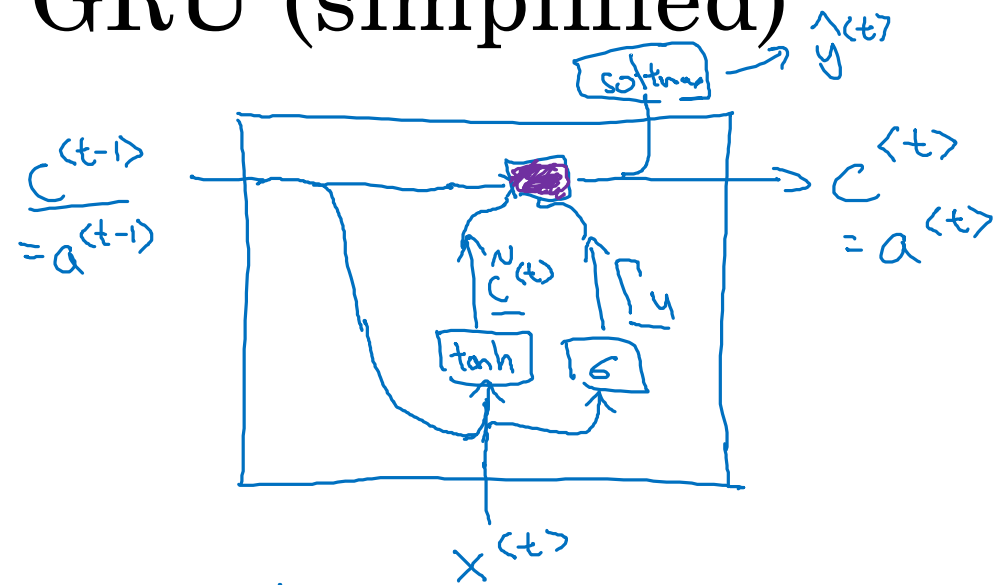
## Gated Recurrent Unit (GRU)

# RNN unit



$$\underline{a^{<t>}} = \overset{\substack{\text{tanh} \\ \downarrow}}{g}(\underbrace{W_a[a^{<t-1>}, x^{<t>}]}_{\uparrow} + \underline{b_a})$$

# GRU (simplified)



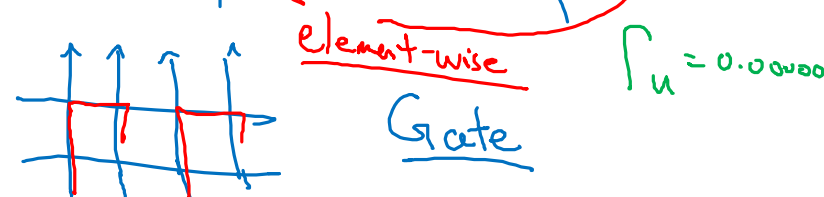
$C$  = memory cell

$$\rightarrow \boxed{C^{(t)}} = \underline{a}^{(t)}$$

$$\rightarrow \boxed{\tilde{C}^{(t)}} = \tanh(W_c [c^{(t-1)}, x^{(t)}] + b_c)$$

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

$$\boxed{C^{(t)}} = \underbrace{\Gamma_u}_{\leftarrow \text{"update"}} * \tilde{C}^{(t)} + (1 - \Gamma_u) * \boxed{C^{(t-1)}}$$



$\Gamma_u = 1$   
 $\tilde{C}^{(t)} = 1$   
 $\Gamma_u = 0 \quad \Gamma_u = 0 \quad \Gamma_u = 0 \quad \dots$   
 $\rightarrow$  The cat, which already ate ..., was full.

[Cho et al., 2014. On the properties of neural machine translation: Encoder-decoder approaches]

[Chung et al., 2014. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling]

Andrew Ng

# Full GRU

$$\tilde{h} \quad \tilde{c}^{<t>} = \tanh(W_c[\underbrace{c^{<t-1>} * \Gamma_r}_{\text{LSTM}}, \underbrace{x^{<t>}}_{\text{LSTM}}] + b_c)$$

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

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

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

LSTM

The cat, which ate already, was full.