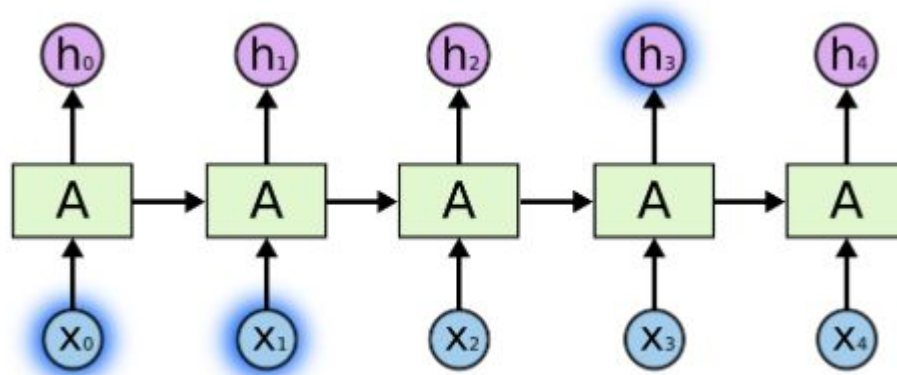
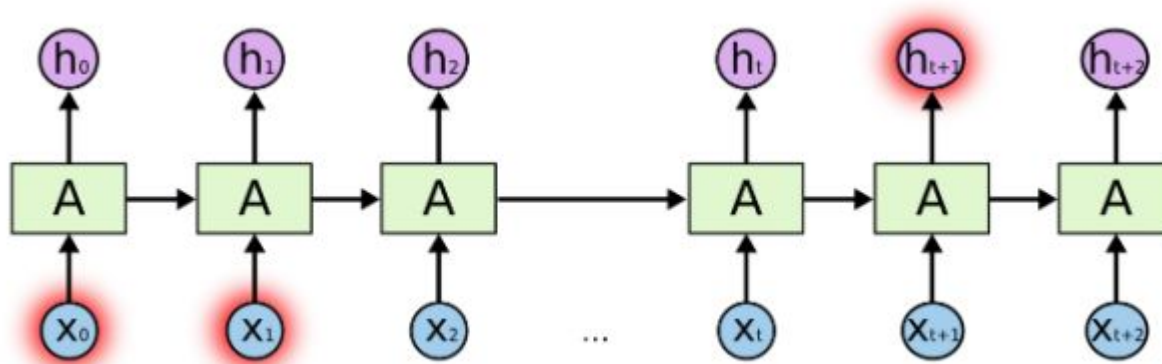


Recurrent Neural Networks



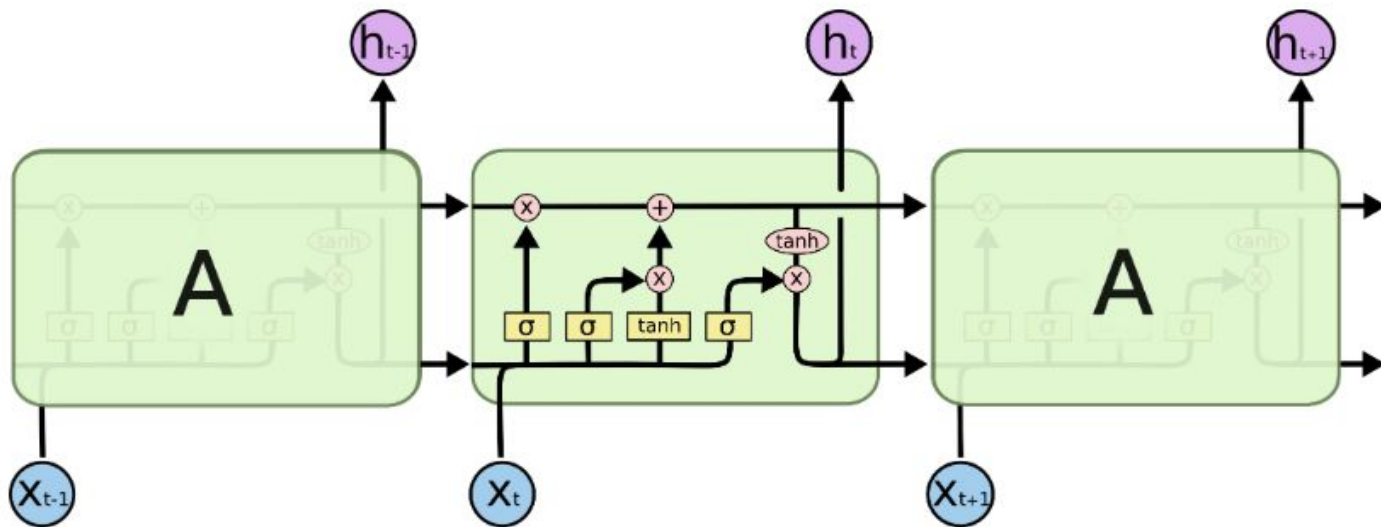
<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>

Vanishing Gradient



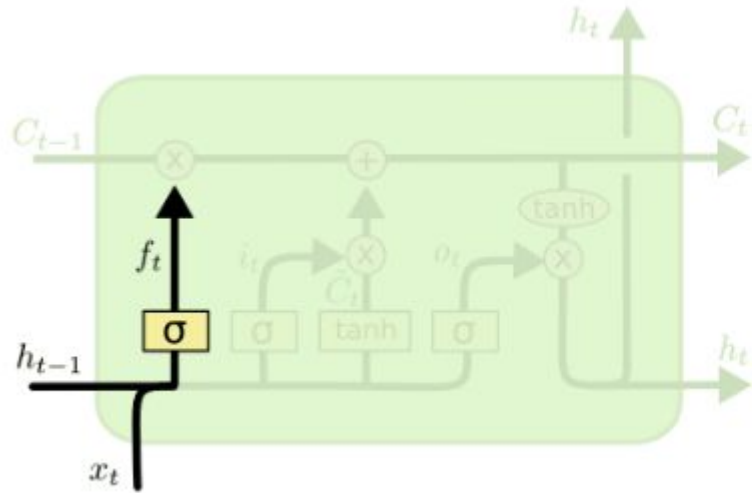
<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>

LSTMs avoid the long-term dependency problem



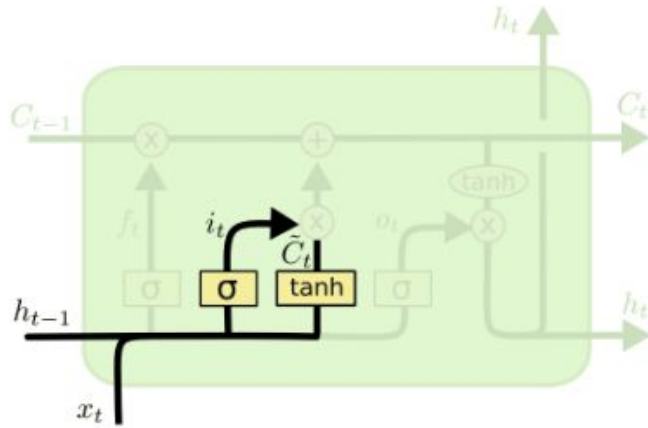
<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>

Forget Gate - Decide what to forget



$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f)$$

Update Gate - Decide what to update

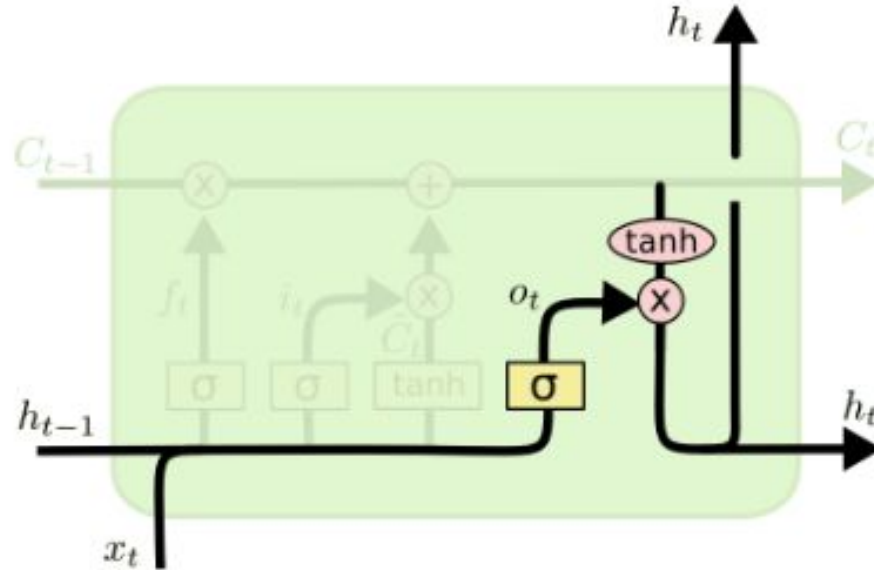


$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Output Gate - Decide what to output

Hochreiter & Schmidhuber



<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>