



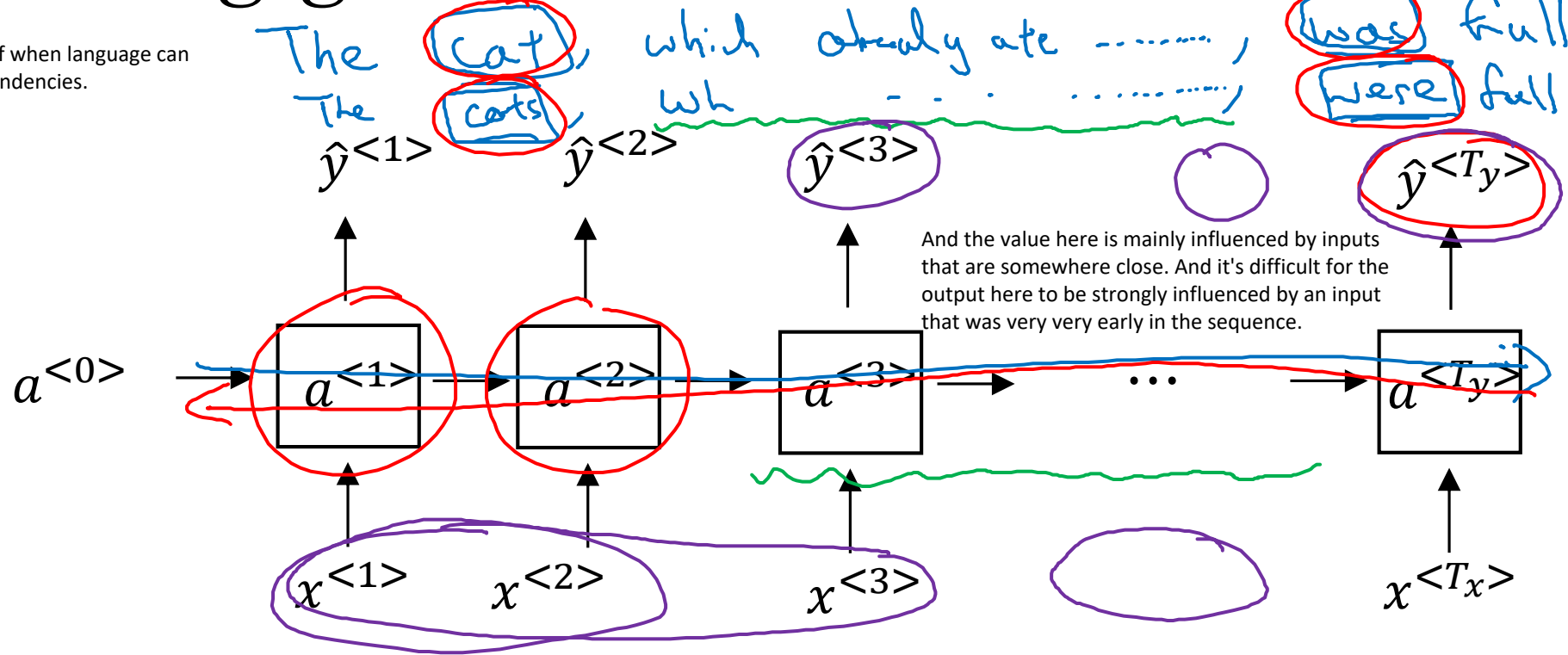
deeplearning.ai

Recurrent Neural Networks

Vanishing gradients with RNNs

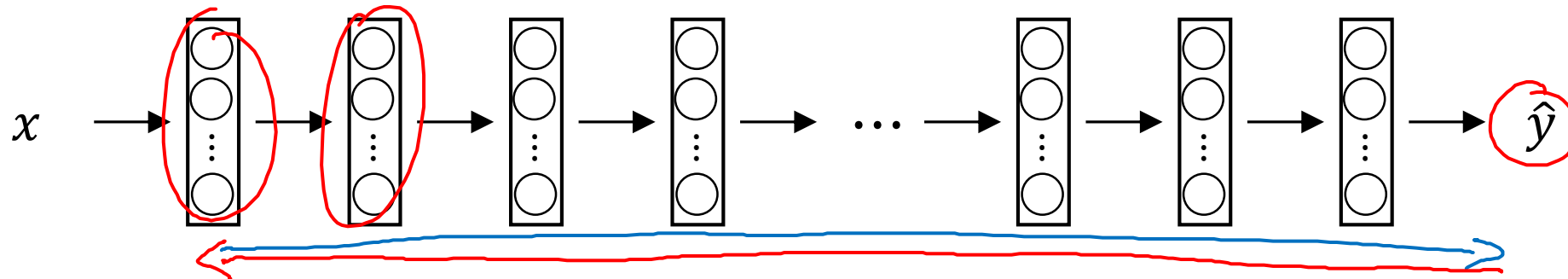
Vanishing gradients with RNNs

And this is one example of when language can have very long-term dependencies.



It's just very difficult for the area to back propagate all the way to the beginning of the sequence and therefore to modify how the neural network is doing computations earlier in the sequence.

So this is a weakness of the basic RNN algorithm.



Exploding gradients.

NaN

Gradient clipping