deeplearning.ai

Recurrent Neural Networks

Vanishing gradients with RNNs

Vanishing gradients with RNNs

The (cat), which already ate ...., (was) full
The (cats), wh ...., (were) full

$\hat{y}^{<1>}$ $\hat{y}^{<2>}$ $\hat{y}^{<3>}$ $\hat{y}^{<T_y>}$

$a^{<0>}$ → $a^{<1>}$ → $a^{<2>}$ → $a^{<3>}$ → ... → $a^{<T_y>}$

$x^{<1>}$ $x^{<2>}$ $x^{<3>}$ $x^{<T_x>}$

x → ... → $\hat{y}$

100
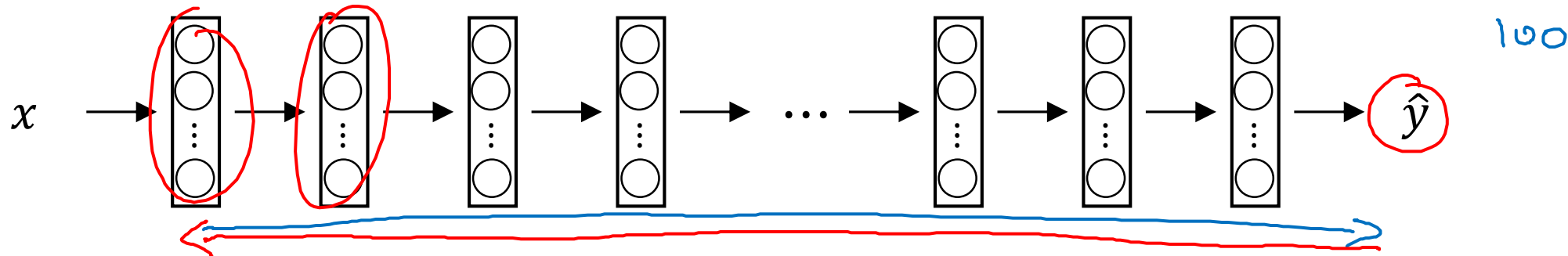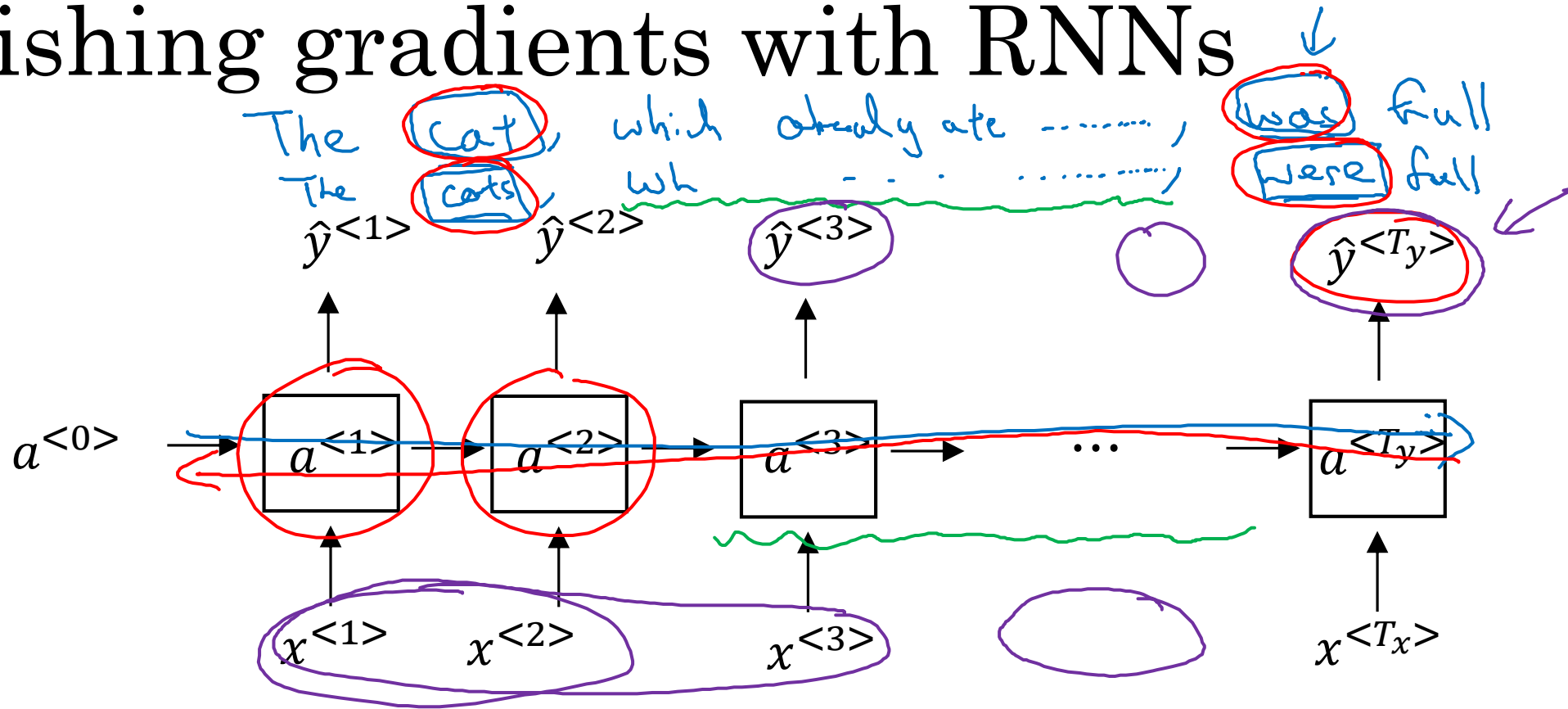
Exploding gradients.   NaN   Gradient clipping

Andrew Ng