Modern RNNs
TOTAL POINTS 4

1.Question 1
Choose correct statements about the exploding gradient problem:

(**X**) Exploding gradient problem is easy to detect.

() ReLU nonlinearity helps with the exploding gradient problem.

(**X**) The reason of the exploding gradient problem in the simple RNN is the recurrent weight matrix W. Nonlinearities sigmoid, tanh, and ReLU does not cause the problem.

() The threshold for gradient clipping should be as low as possible to make the training more efficient.

1 point

2.Question 2
Choose correct statements about the vanishing gradient problem:

() Vanishing gradient problem is easy to detect.

(**X**) Both nonlinearity and the recurrent weight matrix W cause the vanishing gradient problem.

(**X**) Orthogonal initialization of the recurrent weight matrix helps with the vanishing gradient problem.

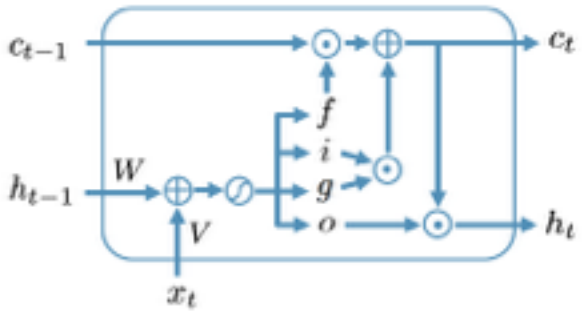
() Truncated BPTT helps with the vanishing gradient problem.

1 point

3. Consider the LSTM architecture: 1 point

$$\begin{pmatrix} g_t \\ i_t \\ o_t \\ f_t \end{pmatrix} = \begin{pmatrix} \tilde{f} \\ \sigma \\ \sigma \\ \sigma \end{pmatrix} (Vx_t + Wh_{t-1} + b)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot g_t, \quad h_t = o_t \cdot \tilde{f}(c_t)$$



Choose correct statements about this architecture:

Choose correct statements about this architecture:

(X) The LSTM needs four times more parameters than the simple RNN.

() Gradients do not vanish on the way through memory cells c in the LSTM with forget gate.

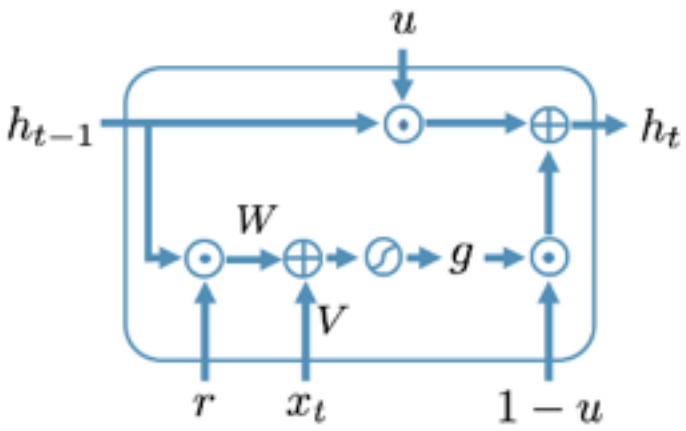
() There is a combination of the gates values which makes the LSTM completely equivalent to the simple RNN.

(X) The exploding gradient problem is still possible in LSTM on the way between h_{t-1} and h_{t}

4. Consider the GRU architecture:

$$g_t = \tilde{f}\left(V_g x_t + W_g(h_{t-1} \cdot r_t) + b_g\right)$$

$$h_t = (1 - u_t) \cdot g_t + u_t \cdot h_{t-1}$$



Which combination of the gate values makes this model equivalent to the simple RNN? Here value zero corresponds to a closed gate and value one corresponds to an open gate.

() Both reset and update gates are open.

() Both reset and update gates are closed.

(**X**) Reset gate is open and update gate is closed.

() Update gate is open and reset gate is closed.

1 point

I, Chun-Min Jen, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account. Learn more about Coursera's Honor Code