Problem 6.1

Suppose you have a recurrent neural network with input $x[n]$, target $y[n]$, output $\hat{y}[n]$, and error metric

$$E = -\frac{1}{N} \sum_{n=0}^{N-1} (y[n] \ln \hat{y}[n] + (1 - y[n]) \ln(1 - \hat{y}[n]))$$

where

$$\hat{y}[n] = \sigma(e[n]),$$

$$e[n] = x[n] + \sum_{m=1}^{M-1} w[m] \hat{y}[n - m],$$

and where $\sigma(\cdot)$ is the logistic sigmoid. Write $dE/dw[3]$ in terms of the signals $y[n]$ and $\hat{y}[m]$. You can invent auxiliary signals such as $\dot{\sigma}[n]$, $\epsilon[n]$, or $\delta[n]$ if you wish, but you need to define them clearly. You may assume that $\hat{y}[n] = 0$ for $n < 0$.

Problem 6.2

Suppose that

$$h_0 = x^3$$

$$h_1 = \cos(x) + \sin(h_0)$$

$$\hat{y} = \frac{1}{2} (h_1^2 + h_0^2)$$

What is $d\hat{y}/dx$? Express your answer as a function of $x$ only, without the variables $h_0$ or $h_1$ in your answer.

Problem 6.3

Consider a one-gate recurrent neural net, defined as follows:

$$c[n] = c[n - 1] + w_c x[n] + u_c h[n - 1] + b_c$$

$$h[n] = o[n] c[n]$$

$$o[n] = \sigma(w_o x[n] + u_o h[n - 1] + b_o)$$
where \( \sigma(\cdot) \) is the logistic sigmoid, \( x[n] \) is the network input, \( c[n] \) is the cell, \( o[n] \) is the output gate, and \( h[n] \) is the output. Suppose that you’ve already completed synchronous back-prop, which has given you the following quantity:

\[
\epsilon[n] = \frac{\partial E}{\partial h[n]}
\]

Find asynchronous back-prop formulas, i.e., find formulas for the following quantities, in terms of one another, and/or in terms of the other quantities defined above:

\[
\delta_h[n] = \frac{dE}{dh[n]}
\]
\[
\delta_o[n] = \frac{dE}{do[n]}
\]
\[
\delta_c[n] = \frac{dE}{dc[n]}
\]

**Problem 6.4**

Using the CReLU nonlinearity for both \( \sigma_h \) and \( \sigma_g \) in an LSTM, choose weights and biases, \( \{b_c, u_c, w_c, b_f, u_f, w_f, b_i, u_i, w_i, b_o, u_o, w_o\} \), that will cause an LSTM to count the number of nonzero inputs, and output the tally only when the input is zero:

\[
h[n] = \begin{cases} 
\sum_{m=0}^{n} \mathbb{1}[x[m] \geq 1] & x[n] = 0 \\
0 & x[n] \geq 1 
\end{cases}
\]

where \( \mathbb{1}[\cdot] \) is the unit indicator function, and you may assume that \( x[n] \) is always a non-negative integer.