Instructor: Vishnu Boddeti
CSE 891-001: Deep Learning
Homework 3
November, 17

**1 Binary Addition (2.5pts):** There are multiple possible solutions, here is one.

$$\begin{aligned} \boldsymbol{W}_{xh} &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \\ \boldsymbol{W}_{hh} &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\ \boldsymbol{W}_{hy} &= \begin{bmatrix} 1 & -1 & 1 \end{bmatrix} \\ \boldsymbol{b}_h &= \begin{bmatrix} -0.5 \\ -1.5 \\ -2.5 \end{bmatrix} \\ \boldsymbol{b}_y &= 0 \end{aligned} \tag{1}$$

**2 LSTM Gradient (3.5pts):**

$$\begin{aligned} \overline{h^{(t)}} &= \frac{\partial L^{(t)}}{\partial h^{(t)}} + \overline{g^{(t+1)}}(1-g^{(t+1)})w_{hg} + \overline{o^{(t+1)}}o^{(t+1)}(1-o^{(t+1)})w_{ho} + \overline{f^{(t+1)}}f^{(t+1)}(1-f^{(t+1)})w_{hf} \\ &\quad + \overline{i^{(t+1)}}i^{(t+1)}(1-i^{(t+1)})w_{hi} \\ \overline{c^{(t)}} &= \overline{h^{(t)}}o^{(t)}(1-tanh^2(c^{(t)})) + \overline{c^{(t+1)}}f^{(t+1)} \\ \overline{g^{(t)}} &= \overline{c^{(t)}}i^{(t)} \\ \overline{o^{(t)}} &= \overline{h^{(t)}}tanh(c^{(t)}) \\ \overline{f^{(t)}} &= \overline{c^{(t)}}c^{(t-1)} \\ \overline{i^{(t)}} &= \overline{c^{(t)}}g^{(t)} \end{aligned}$$

$$\overline{w_{xi}} = \sum_{t=1}^{T} \overline{i^{(t)}}i^{(t)}(1-i^{(t)})x^{(t)}$$

**3 Convolutional Neural Networks (2pts):**

1. (**1pt**) Number of parameters = 207,507,816.
2. (**1pt**) The one on the left.

**4 Autoregressive Generative Models (2pt):**

1. (a) $\mathcal{O}(dHWk^2)$
   (b) $\mathcal{O}(d)$
2. (a) $\mathcal{O}(dHWk^2)$
   (b) $\mathcal{O}(d\sqrt{HW})$