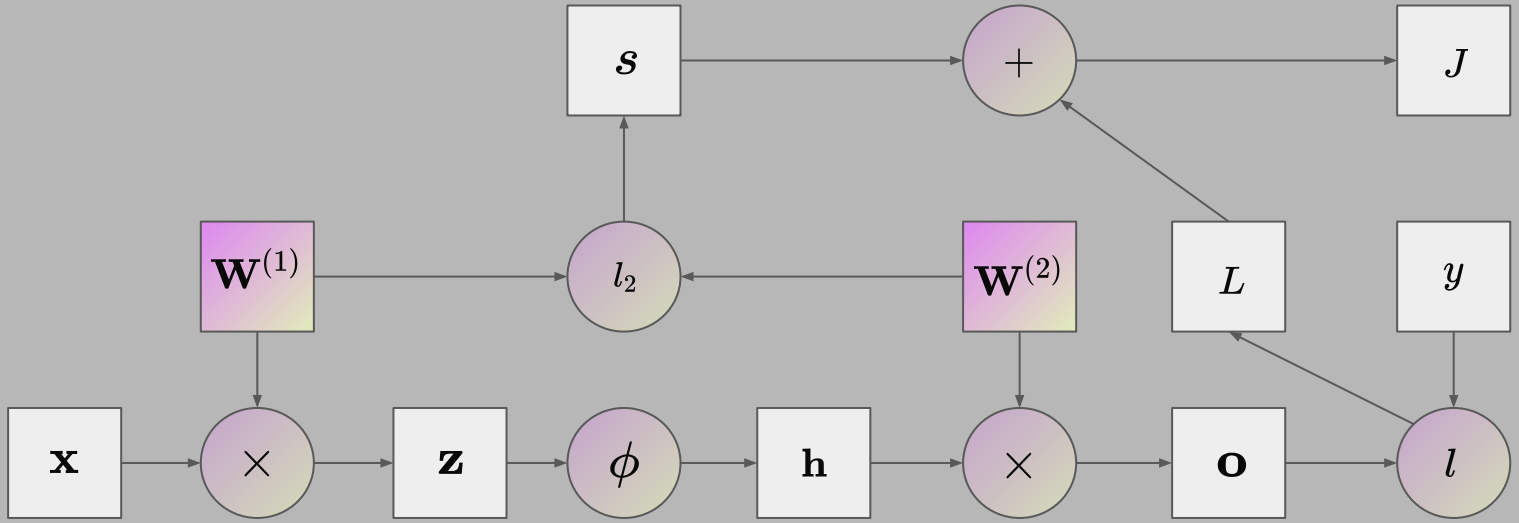
Forward Propagation

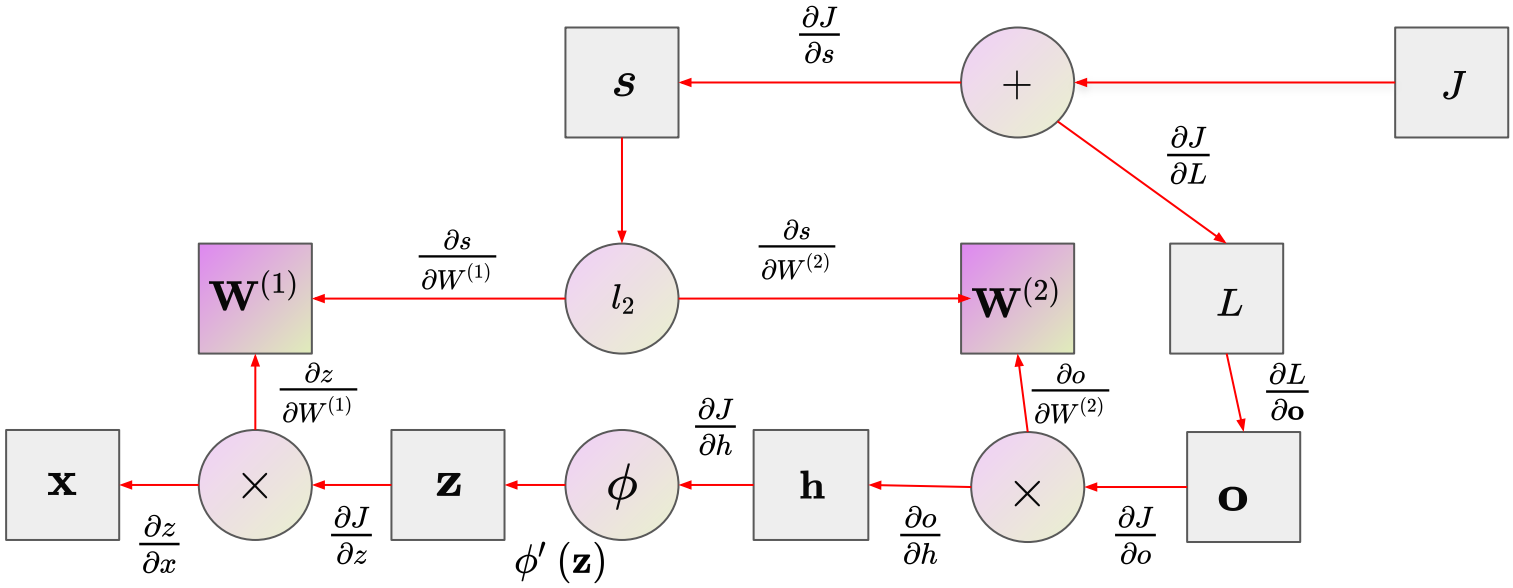


Backpropagation Equations

$$\frac{\partial J}{\partial W^{(1)}} = \frac{\partial J}{\partial z} x^T + \lambda W^{(1)}$$

$$\frac{\partial J}{\partial W^{(2)}} = \frac{\partial J}{\partial o} h^T + \lambda W^{(2)}$$

Backpropagation ---> (Path and direction



$$\frac{\partial o}{\partial W^{(2)}} = h^\top$$

$$\frac{\partial z}{\partial W^{(1)}} = x^\top$$

$$\frac{\partial J}{\partial \mathbf{z}} = \text{prod}\left(\frac{\partial J}{\partial \mathbf{h}}, \frac{\partial \mathbf{h}}{\partial \mathbf{z}}\right) = \frac{\partial J}{\partial \mathbf{h}} \odot \phi'(\mathbf{z})$$

$$\frac{\partial J}{\partial \mathbf{o}} = \text{prod}\left(\frac{\partial J}{\partial L}, \frac{\partial L}{\partial \mathbf{o}}\right) = \frac{\partial L}{\partial \mathbf{o}}$$

$$\frac{\partial J}{\partial \mathbf{h}} = \text{prod}\left(\frac{\partial J}{\partial \mathbf{o}}, \frac{\partial \mathbf{o}}{\partial \mathbf{h}}\right) = \mathbf{W}^{(2)\top} \frac{\partial J}{\partial \mathbf{o}}$$

Link for xavier implementation

https://colab.research.google.com/drive/1yAZFByZdlmamlVzxcD1H667ADxMbqUHE?usp=sharing