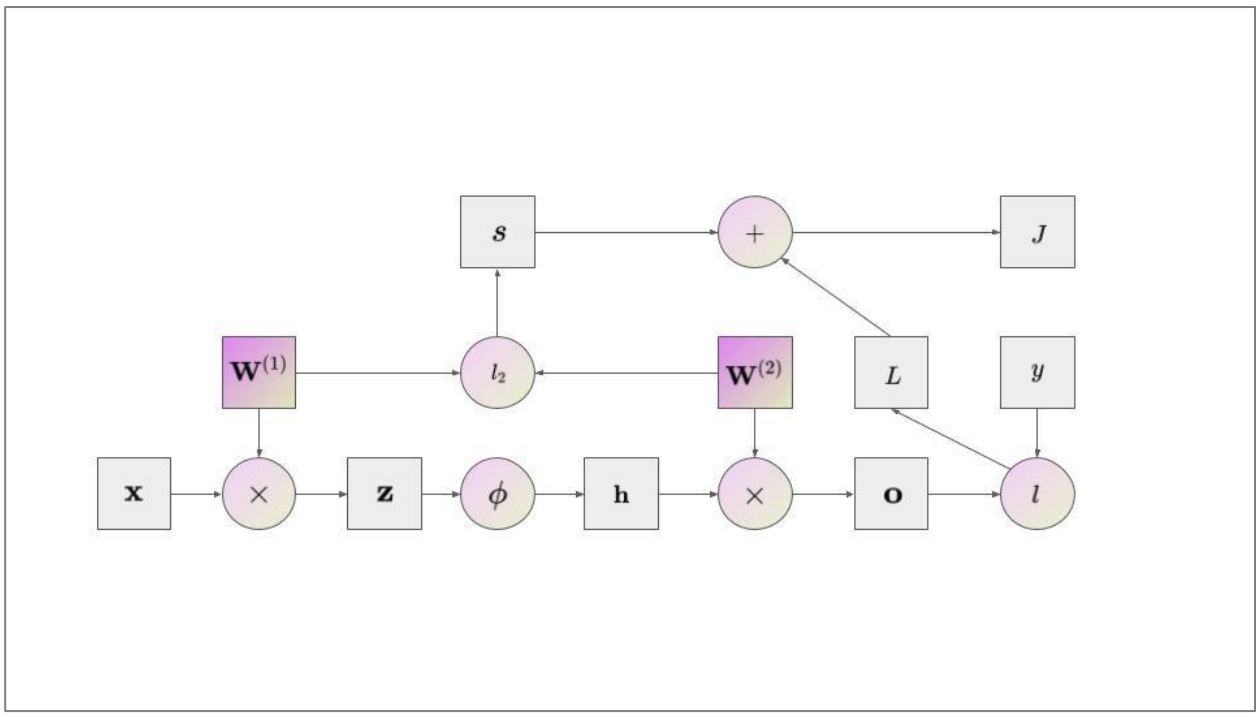
Assignment # 03

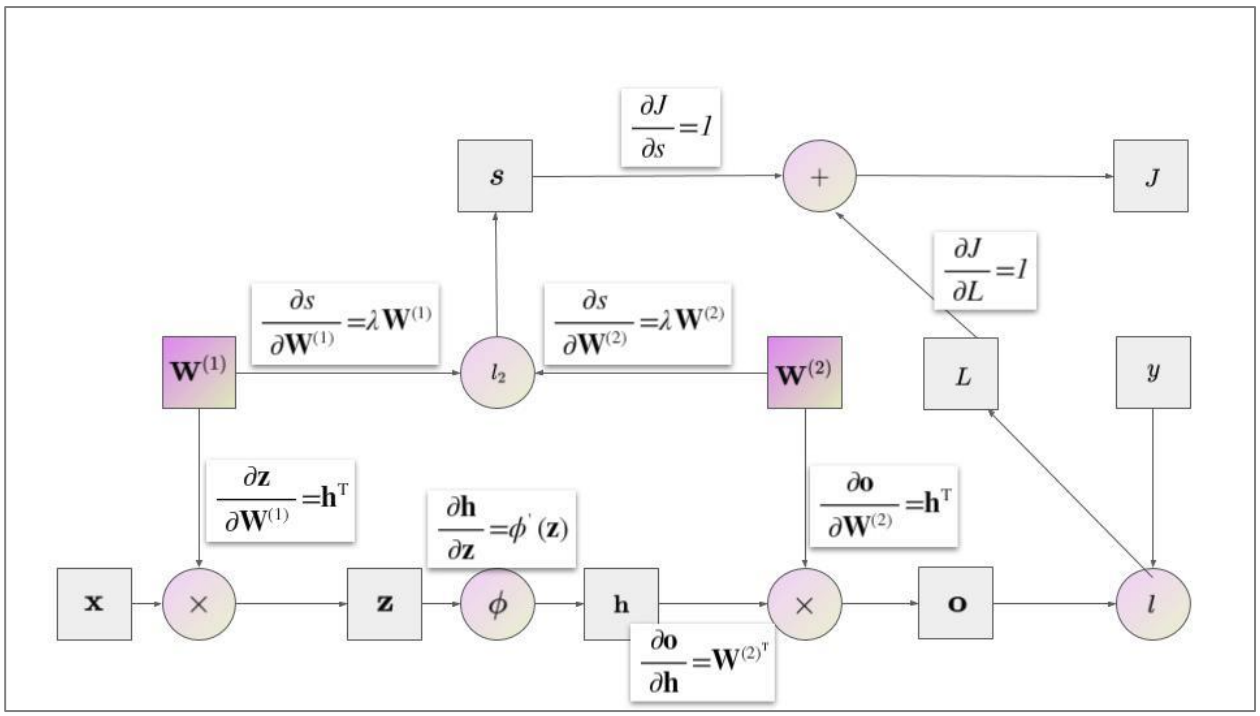
Part 1: Backpropagation in Computation Graph

Question 1: Provide the equations for the backward propagation step in the figure below:

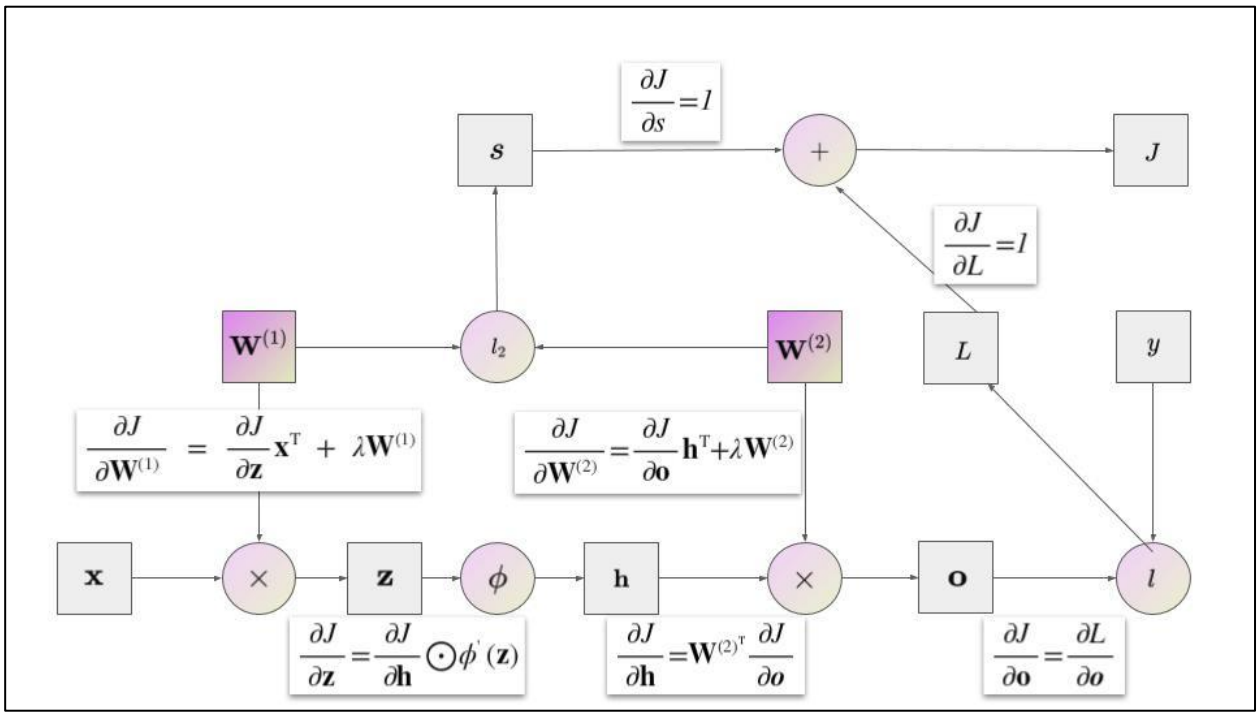


Answer: The above computation graph is for Forward Propagation. We now need to calculate the partial derivatives of 'J' with respect to the parameters '$W^{(1)}$' and '$W^{(2)}$' using backpropagation. For this, we need partial derivative of 'J' with other intermediate variables as well.

In order to show the backpropagation equations to denote the partial derivatives of output 'J' with respect to each of the variables in above graph, we first need to know the partial derivatives of each variable w.r.t. the variable that comes just before them in forward propagation. It is as shown below:



Then the next step is to use chain rule and sum and product rules to arrive at the final partial derivatives of 'J' with respect to all the intermediate connecting variables. It is as shown below:



In the above figure, we see how we calculate all the intermediate partial derivatives and then finally arrive at the required gradients: partial derivatives of 'J' with respect to the parameters '$W^{(1)}$' and '$W^{(2)}$'.

Part 2: Implementation of Xavier Initialization

Task: Create a new Google Colab notebook and show how to implement Xavier initialization from scratch. Compare it with the PyTorch built-in Xavier implementation. Then show how you would use it effectively to initialize your parameters and train a simple model as used in the notebook.

Solution: Below is a link to the GitHub page with colab notebook in which I have implemented the Xavier Initialization and compared it with the PyTorch built-in implementation:

https://github.com/quickSilverShanks/Dive-into-Deep-Learning/blob/master/Assignment_3___Xavier_Initialization.ipynb

Submitted by:

Name: Shashank Prakash

e-Mail: shashank708.2606@gmail.com

Date of Submission: 26th September 2020 (time - 23:45)