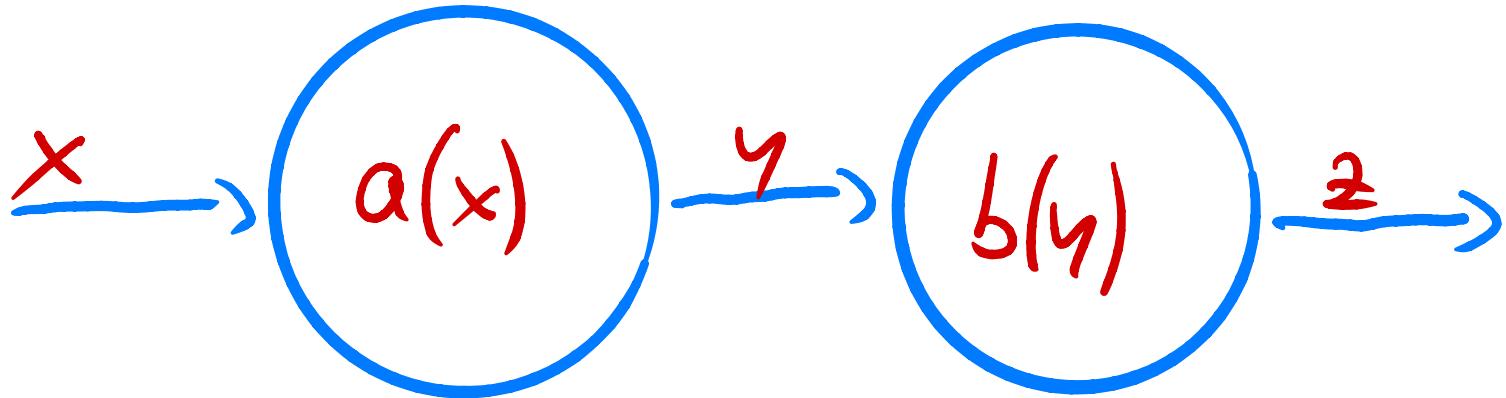
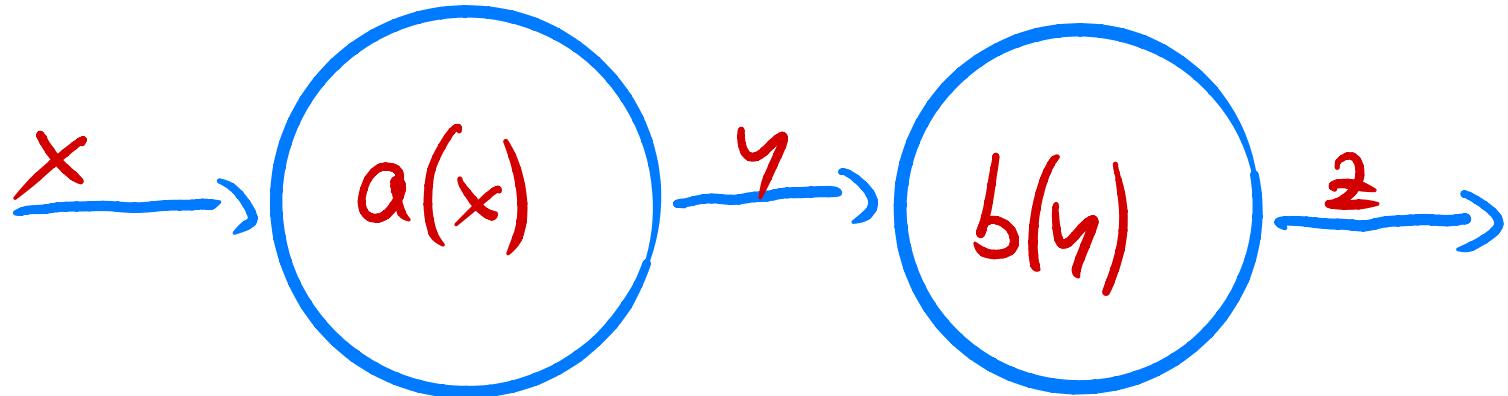
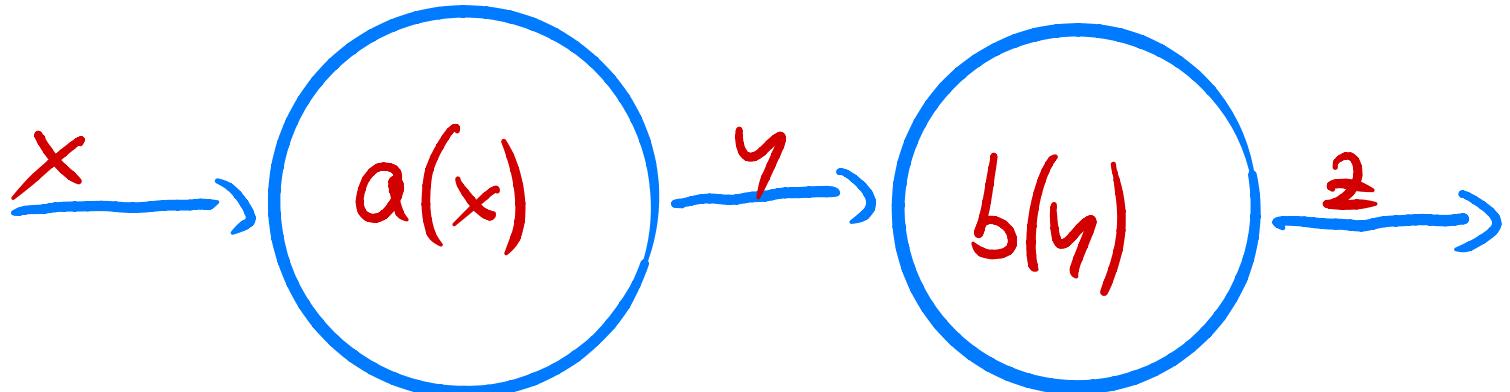


Backpropagation



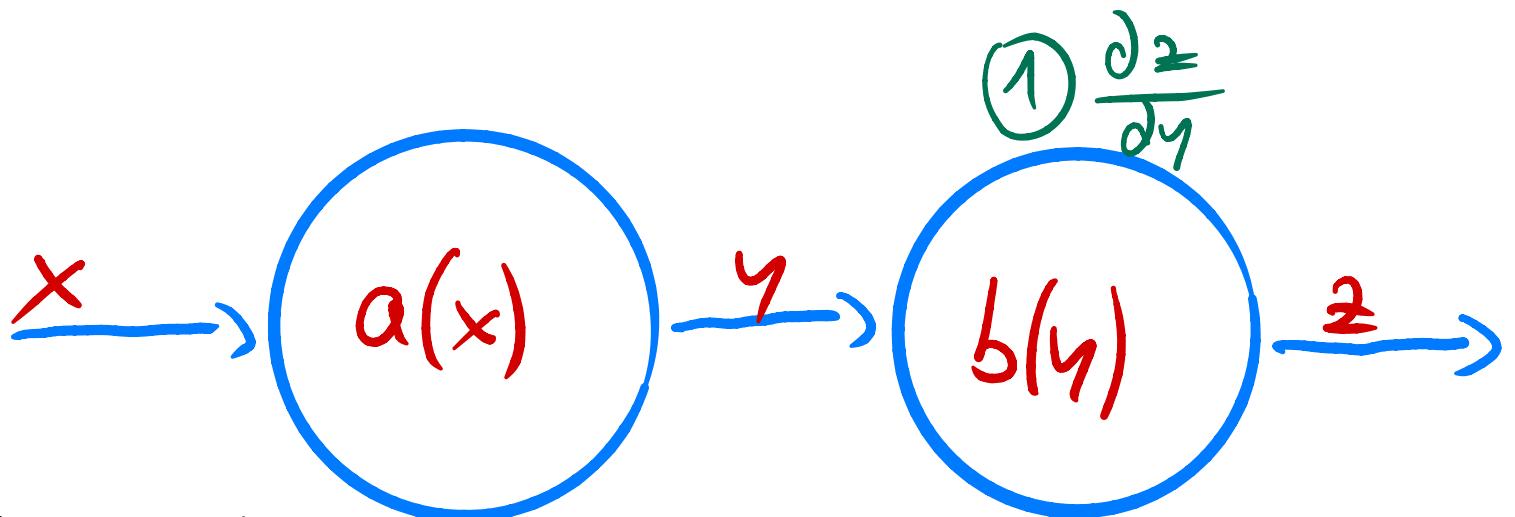


$$\frac{\partial z}{\partial x} \quad ?$$



Chain Rule:

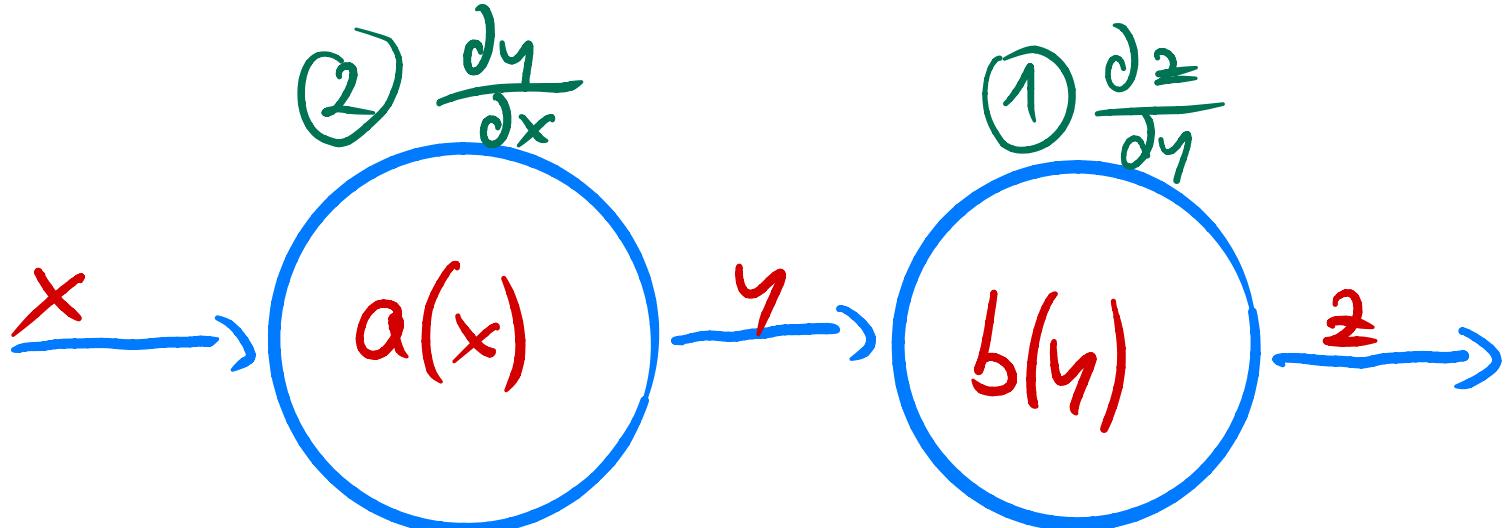
$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial x}$$



Chain Rule:

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial x}$$

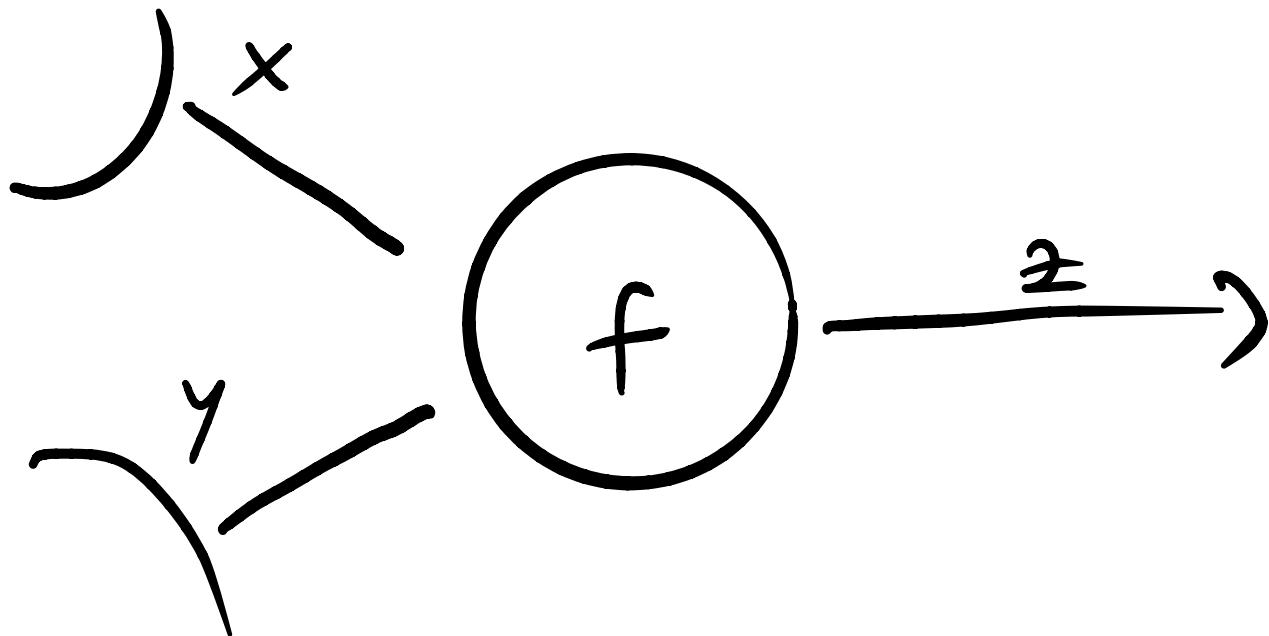
(1)



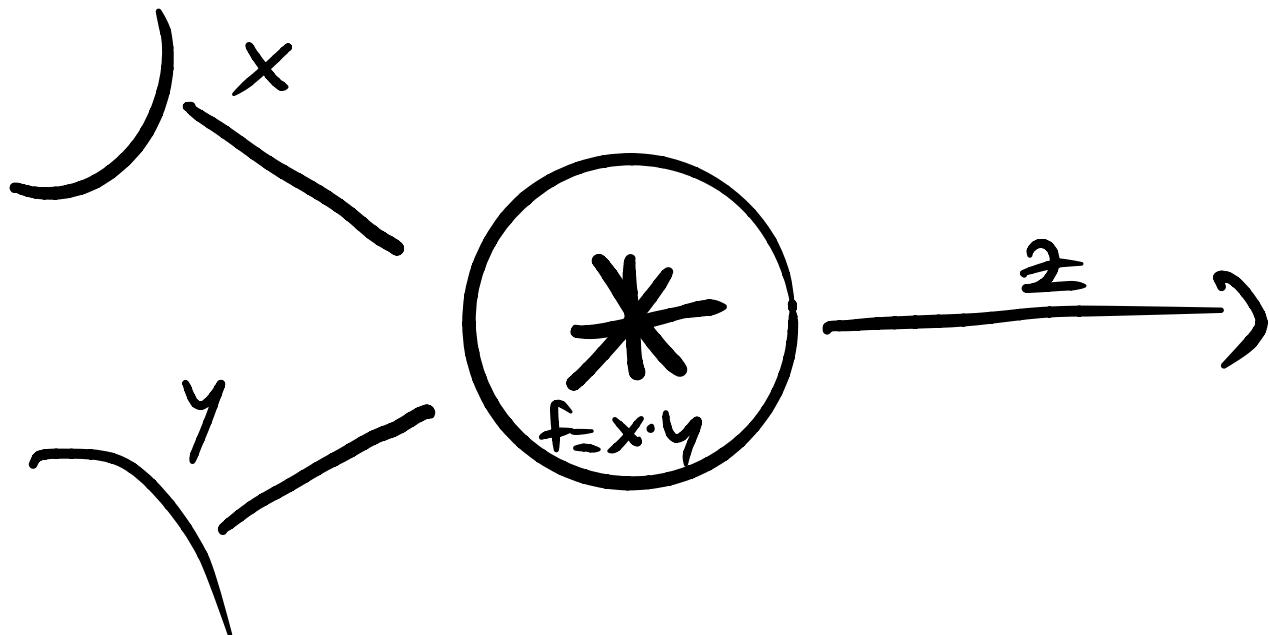
Chain Rule :

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial x}$$

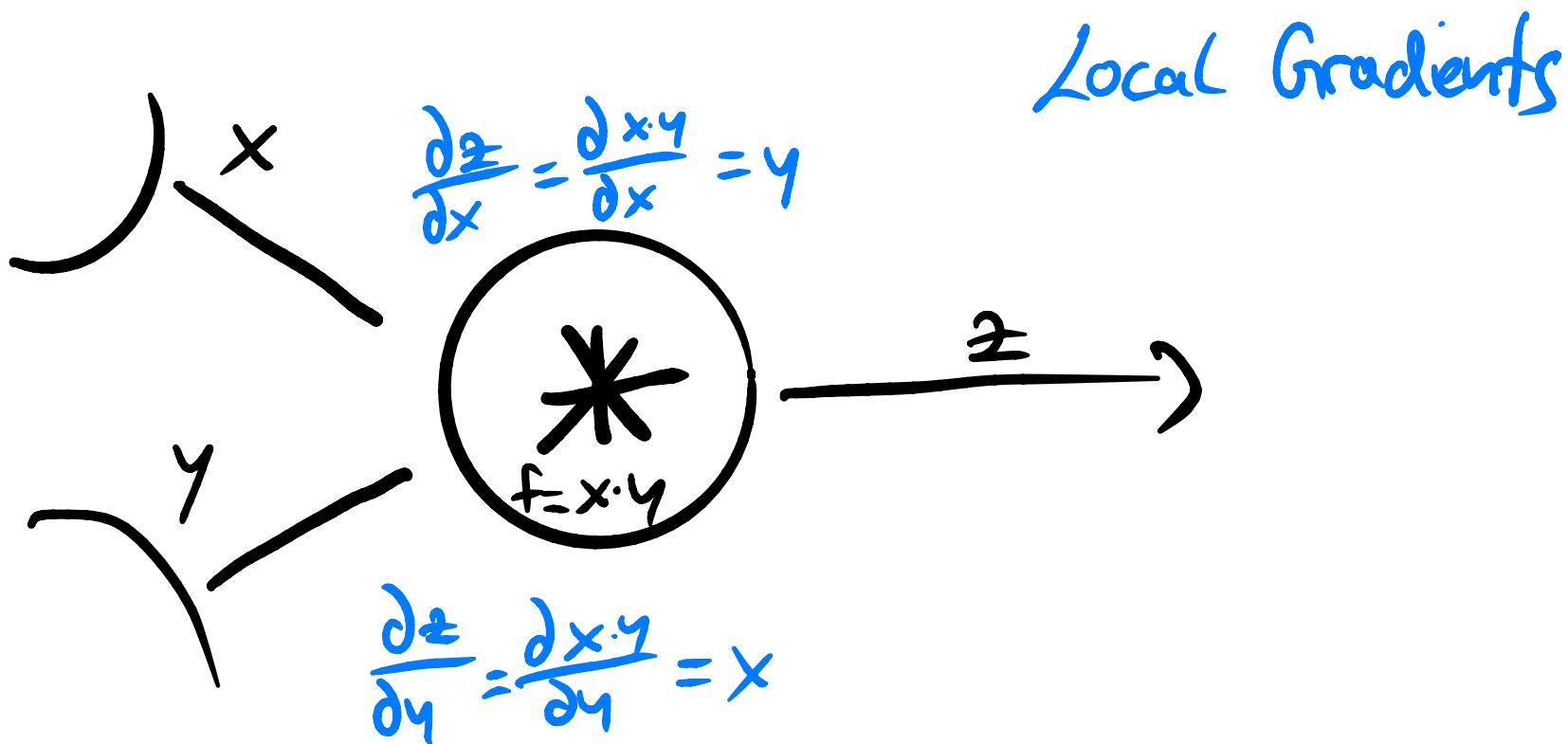
Computational Graph



Computational Graph



Computational Graph



Computational Graph

$$x \quad \frac{\partial z}{\partial x} = \frac{\partial x \cdot y}{\partial x} = y$$

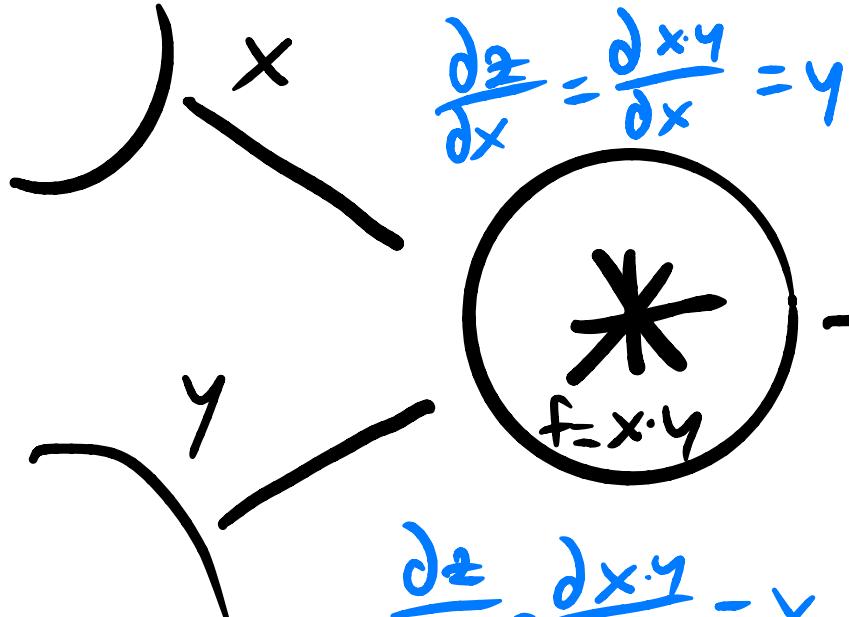
$$y \quad f = x \cdot y \quad z \rightarrow \text{Loss}$$

$$\frac{\partial z}{\partial y} = \frac{\partial x \cdot y}{\partial y} = x$$

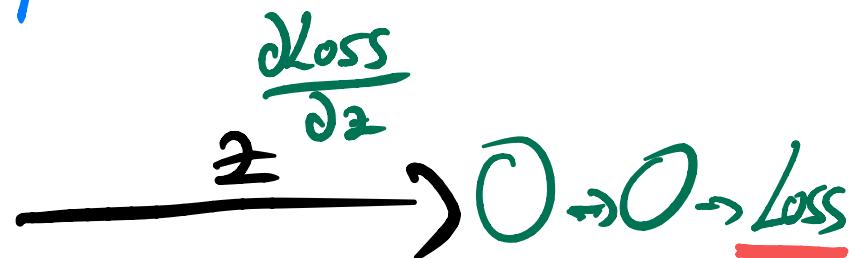
Local Gradients

$$\frac{\partial \text{Loss}}{\partial x} ?$$

Computational Graph

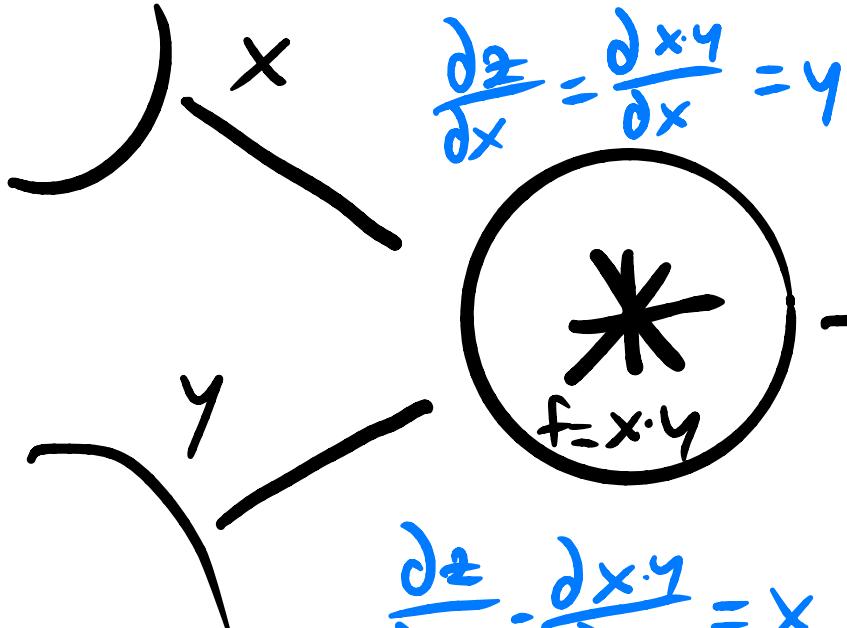


Local Gradients



$$\frac{\partial \text{loss}}{\partial x} ?$$

Computational Graph



$$\frac{\partial z}{\partial x} = \frac{\partial x \cdot y}{\partial x} = y$$

$$\frac{\partial z}{\partial y} = \frac{\partial x \cdot y}{\partial y} = x$$

Local Gradients

$$\frac{\text{loss}}{\partial z} \rightarrow 0 \rightarrow 0 \rightarrow \underline{\text{loss}}$$

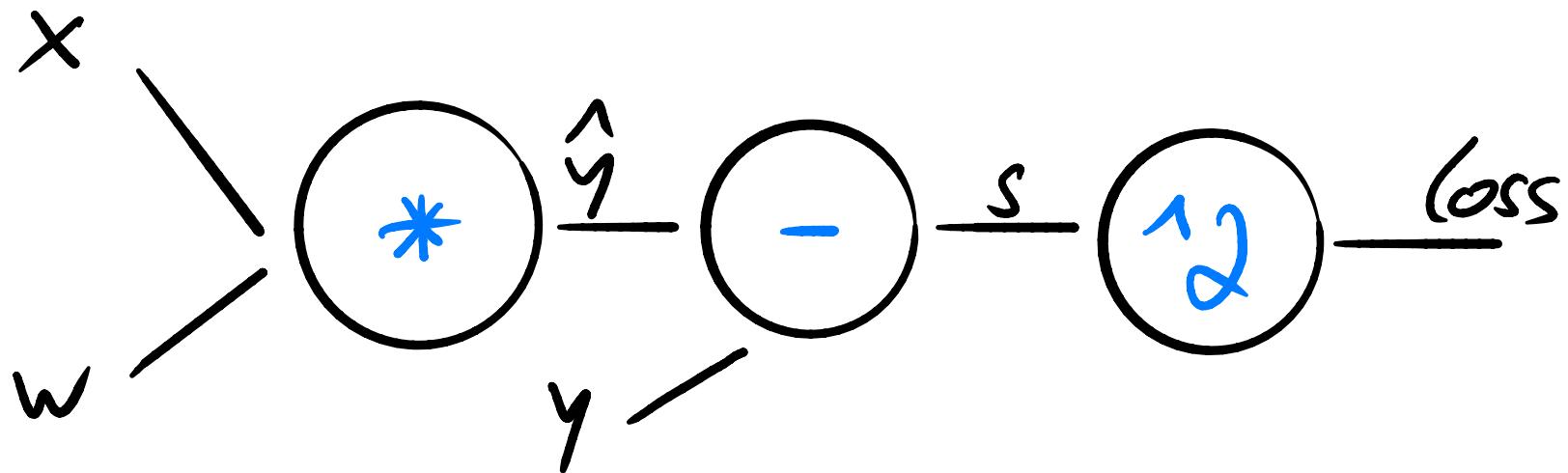
Chain Rule:

$$\frac{\partial \text{loss}}{\partial x} = \frac{\partial \text{loss}}{\partial z} \cdot \frac{\partial z}{\partial x}$$

- 1) Forward pass: Compute Loss
- 2) Compute local gradients
- 3) Backward pass: Compute
 $d\text{Loss} / d\text{Weights}$ using the
Chain Rule

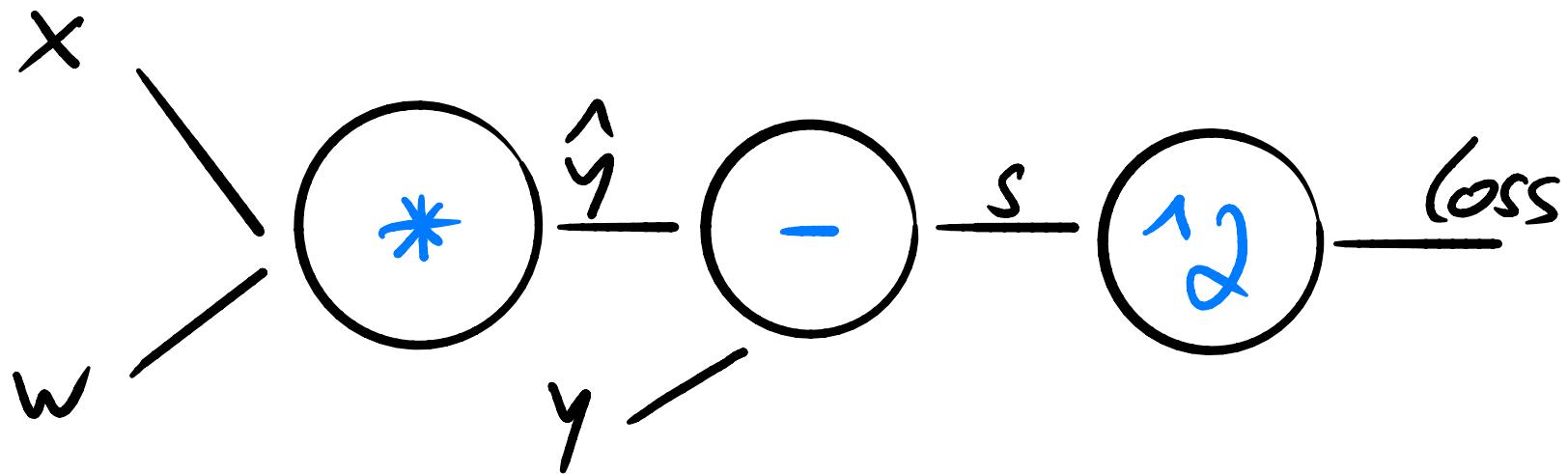
$$\hat{y} = w \cdot x$$

$$\text{loss} = (\hat{y} - y)^2 = (wx - y)^2$$



$$\hat{y} = w \cdot x$$

$$\text{loss} = (\hat{y} - y)^2 = (wx - y)^2$$



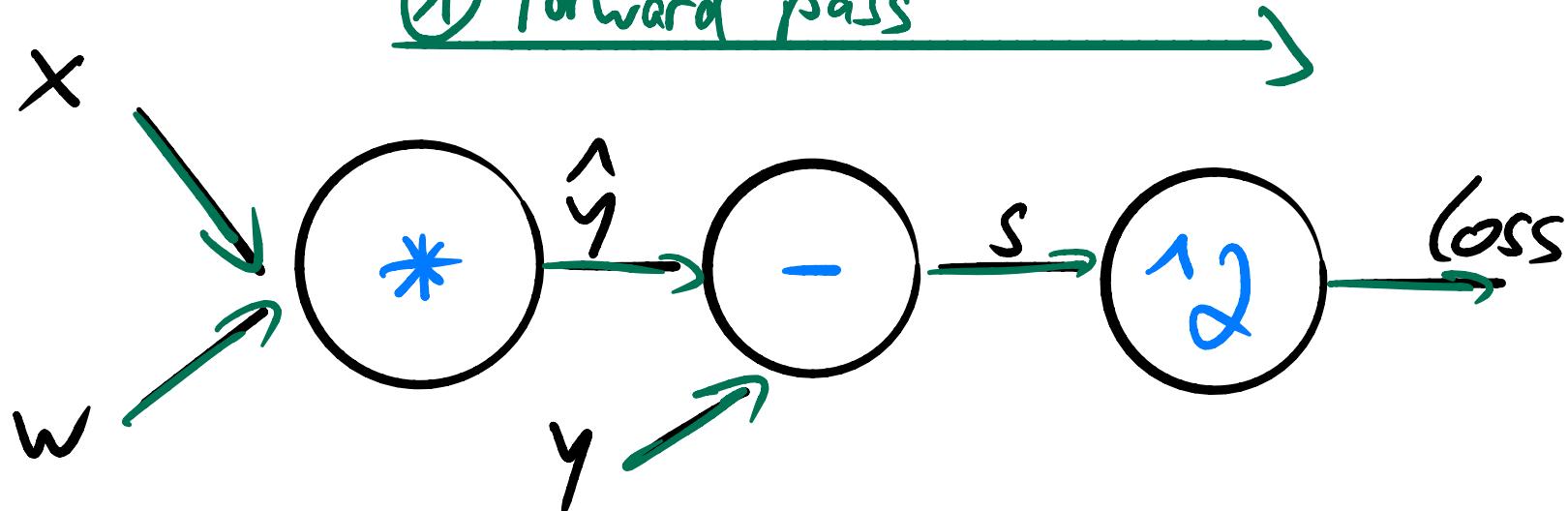
Minimize
Loss

$$\rightarrow \frac{\partial \text{loss}}{\partial w} ?$$

$$\hat{y} = w \cdot x$$

$$\text{loss} = (\hat{y} - y)^2 = (wx - y)^2$$

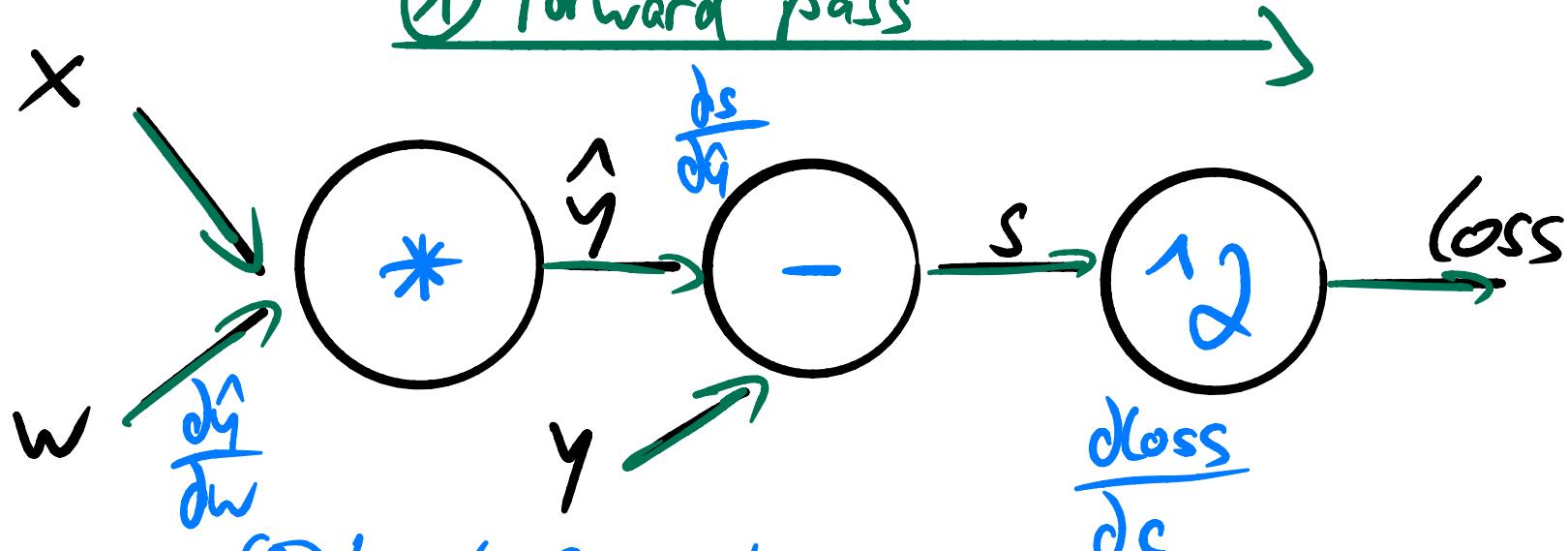
① Forward pass



$$\hat{y} = w \cdot x$$

$$\text{loss} = (\hat{y} - y)^2 = (wx - y)^2$$

① Forward pass

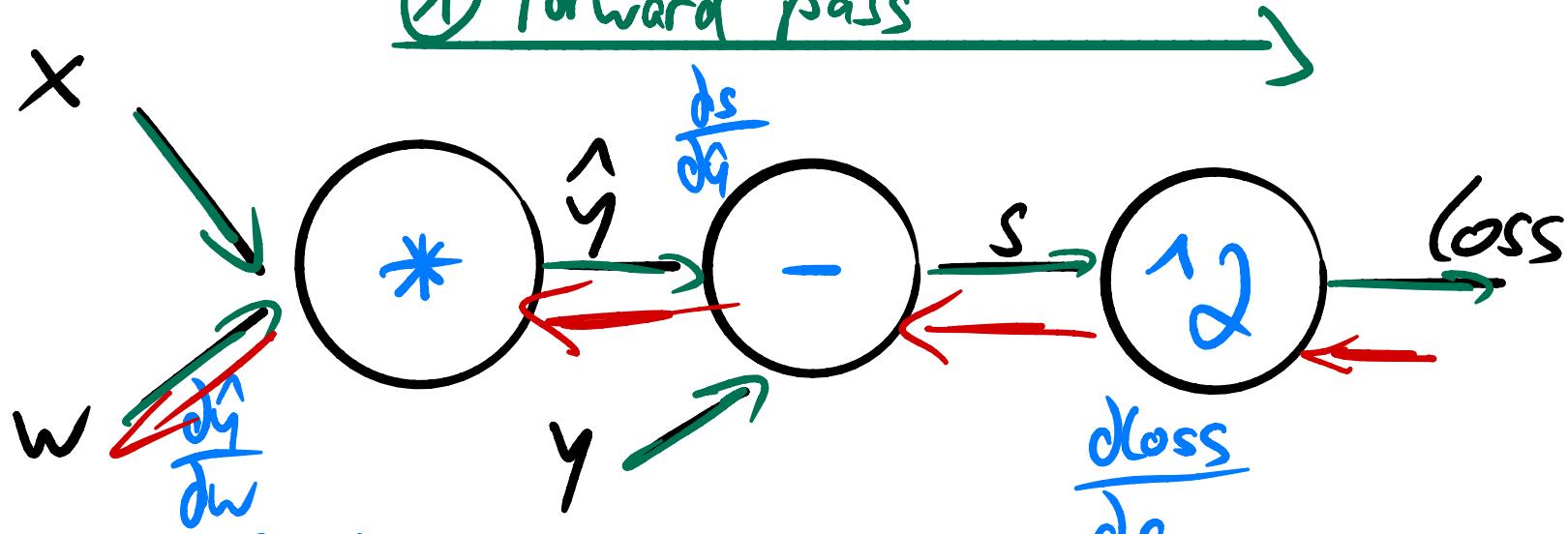


② Local Gradients

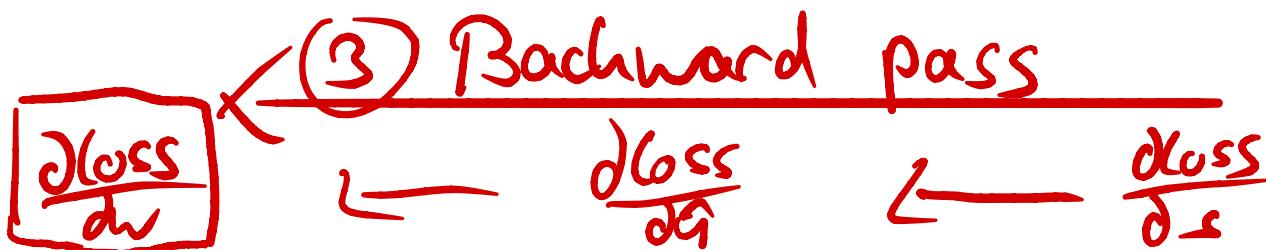
$$\hat{y} = w \cdot x$$

$$\text{loss} = (\hat{y} - y)^2 = (wx - y)^2$$

① Forward pass

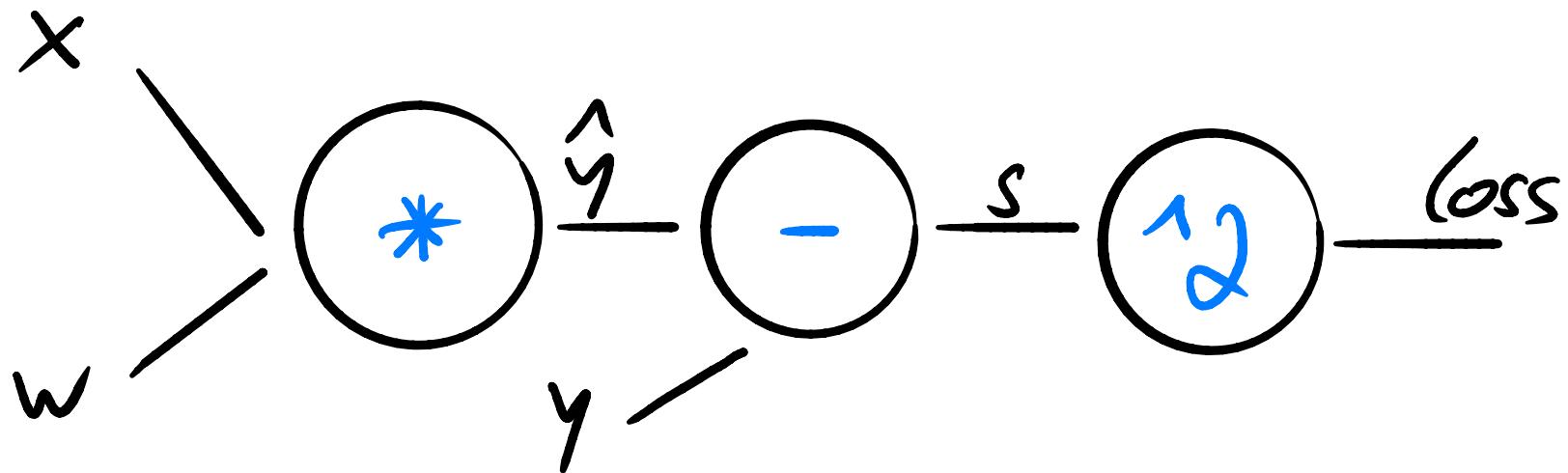


② Local Gradients

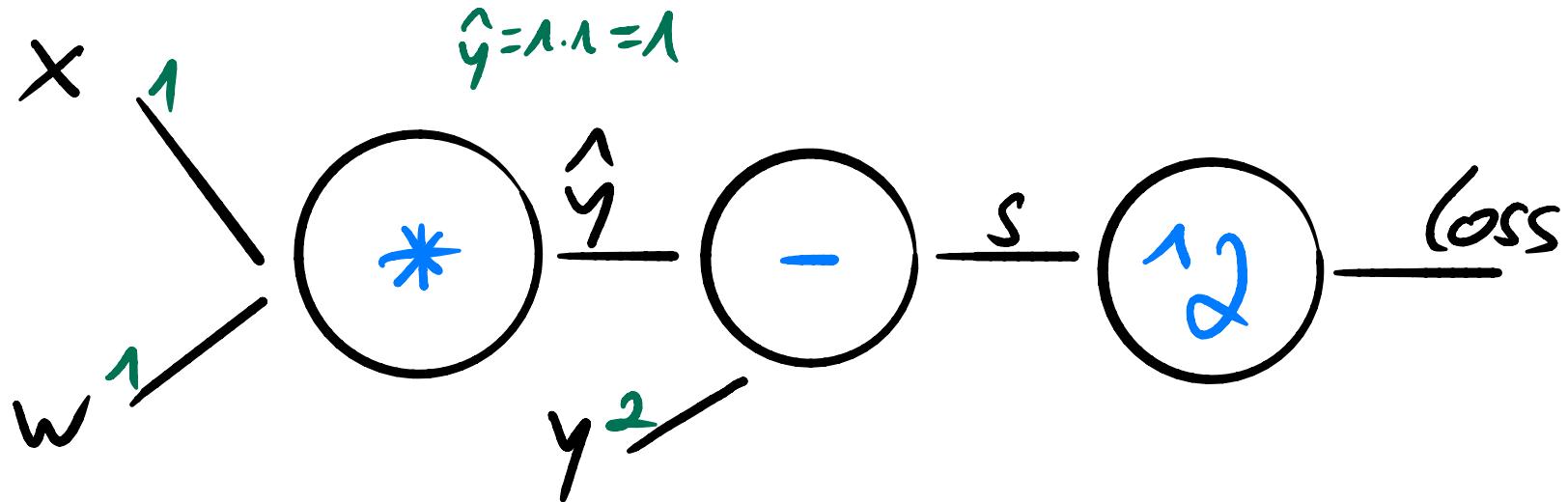


Use
Chain
Rule!

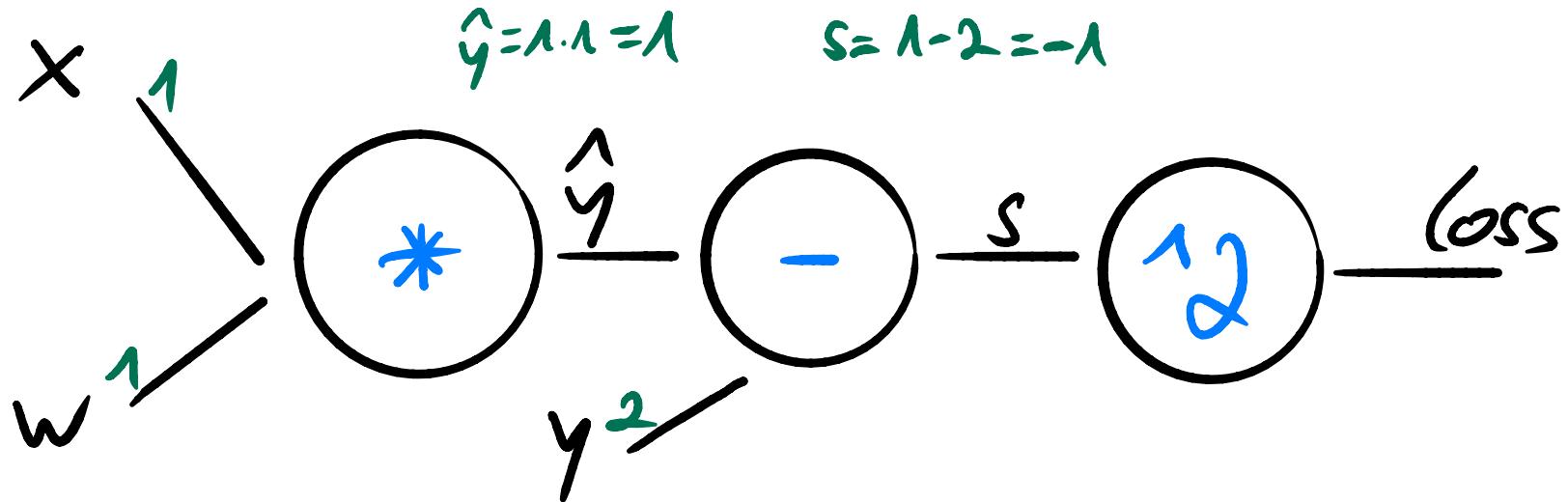
Forward pass $x=1$ $y=2$ $w=1$ \rightarrow



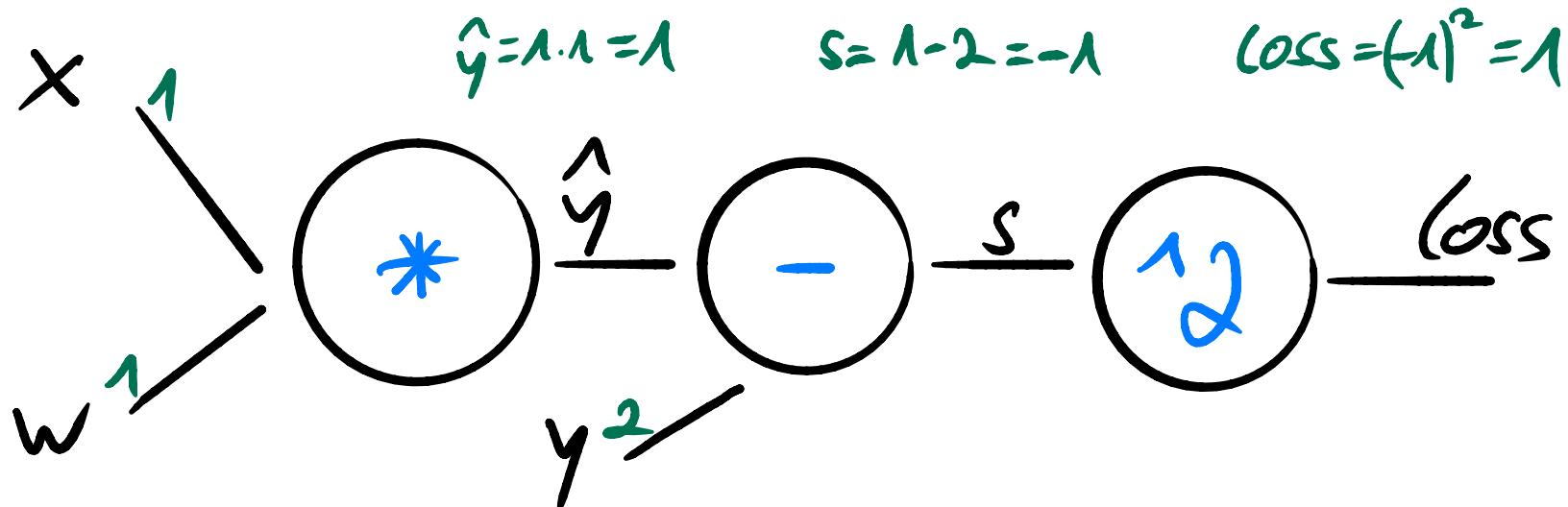
Forward pass $x=1$ $y=2$ $w=1$ \rightarrow



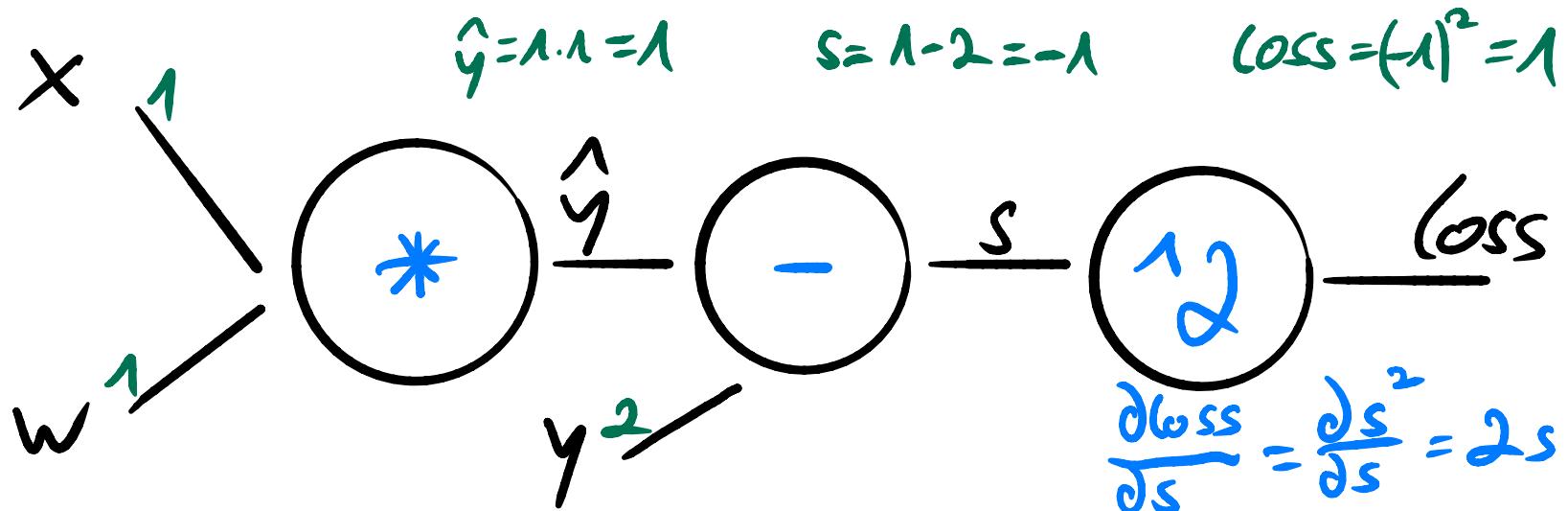
Forward pass $x=1$ $y=2$ $w=1$



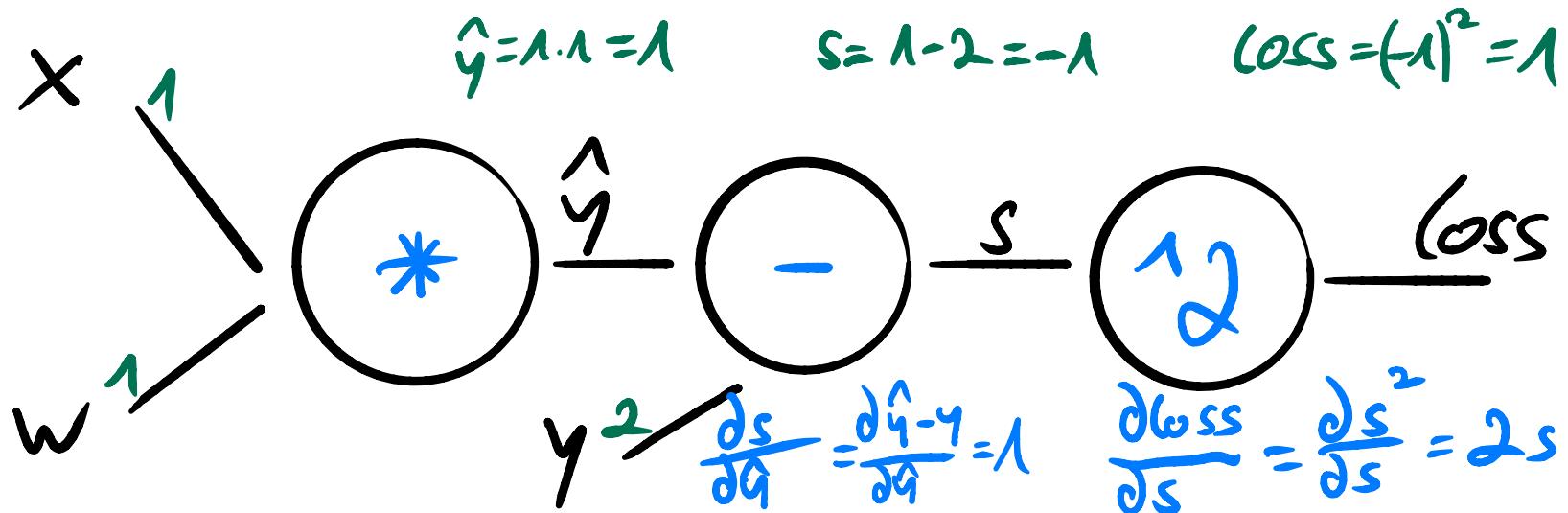
Forward pass $x=1$ $y=2$ $w=1$



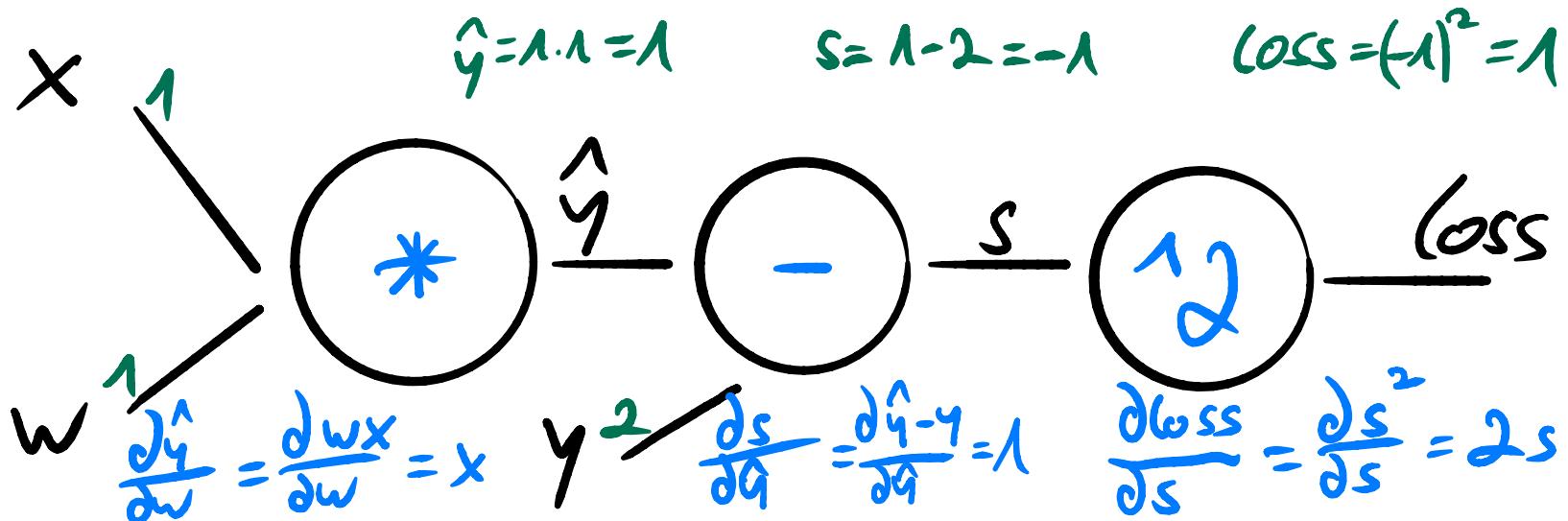
Forward pass $x=1$ $y=2$ $w=1$



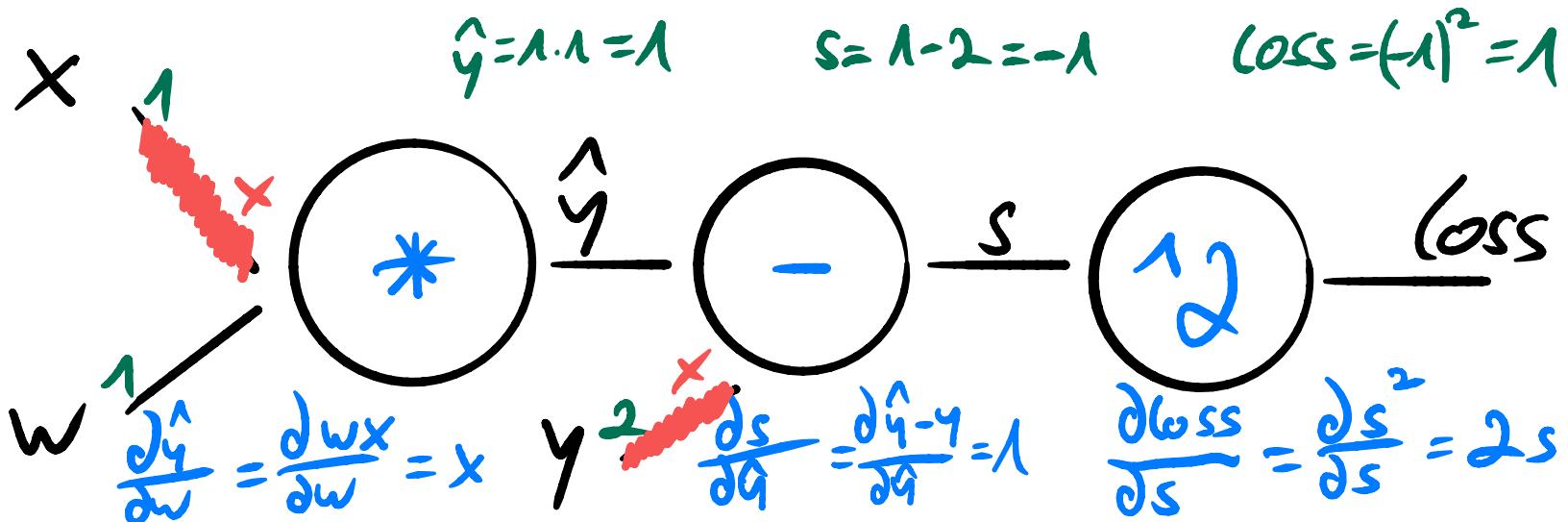
Forward pass $x=1$ $y=2$ $w=1$



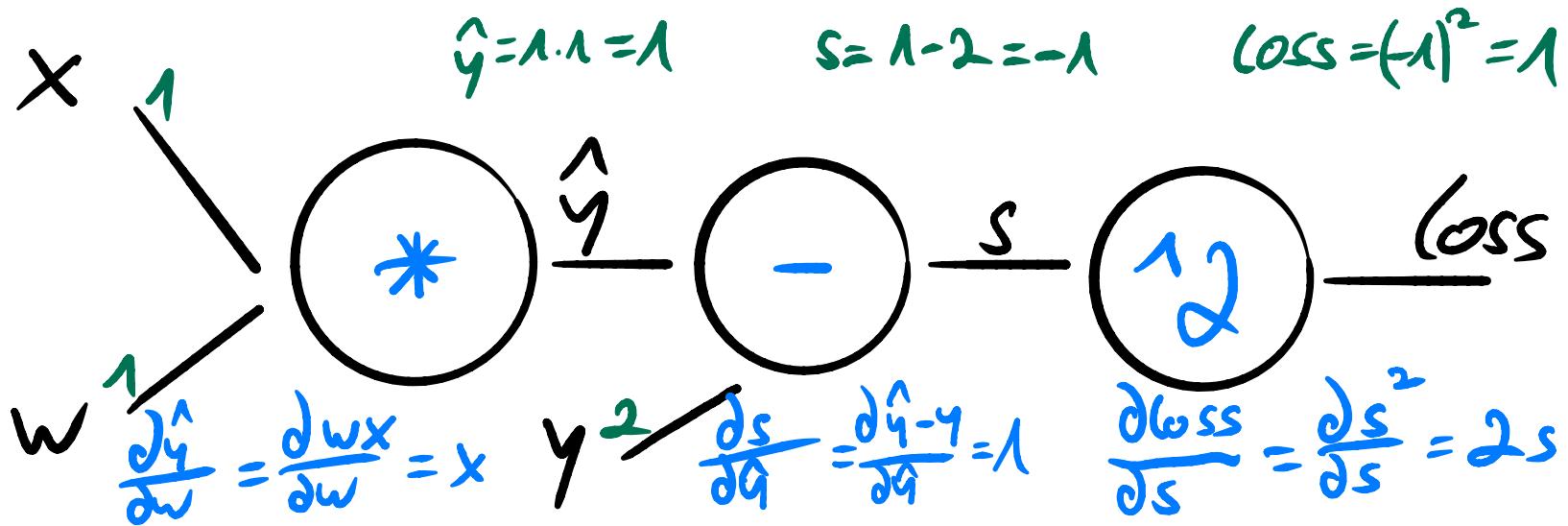
Forward pass $x=1$ $y=2$ $w=1$



Forward pass $x=1$ $y=2$ $w=1$

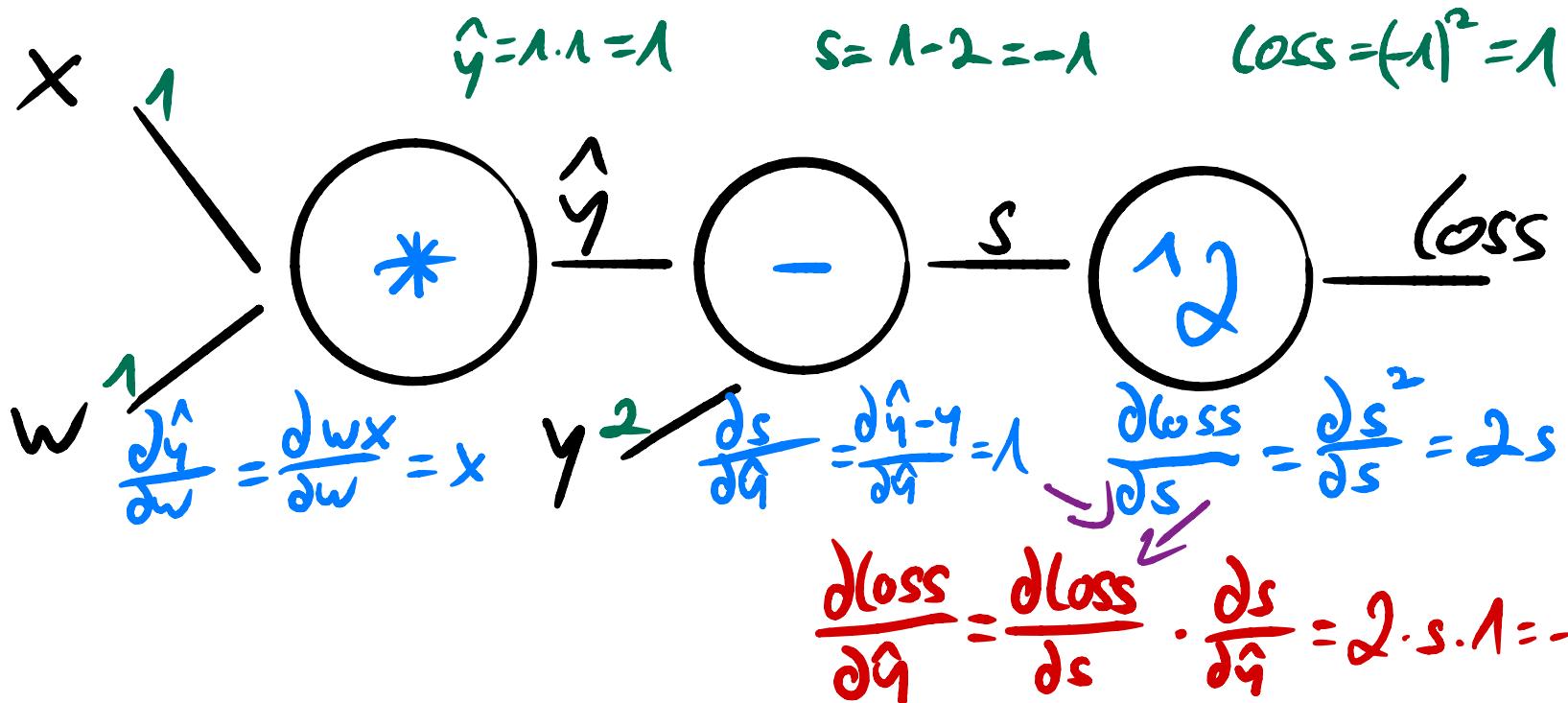


Forward pass $x=1$ $y=2$ $w=1$



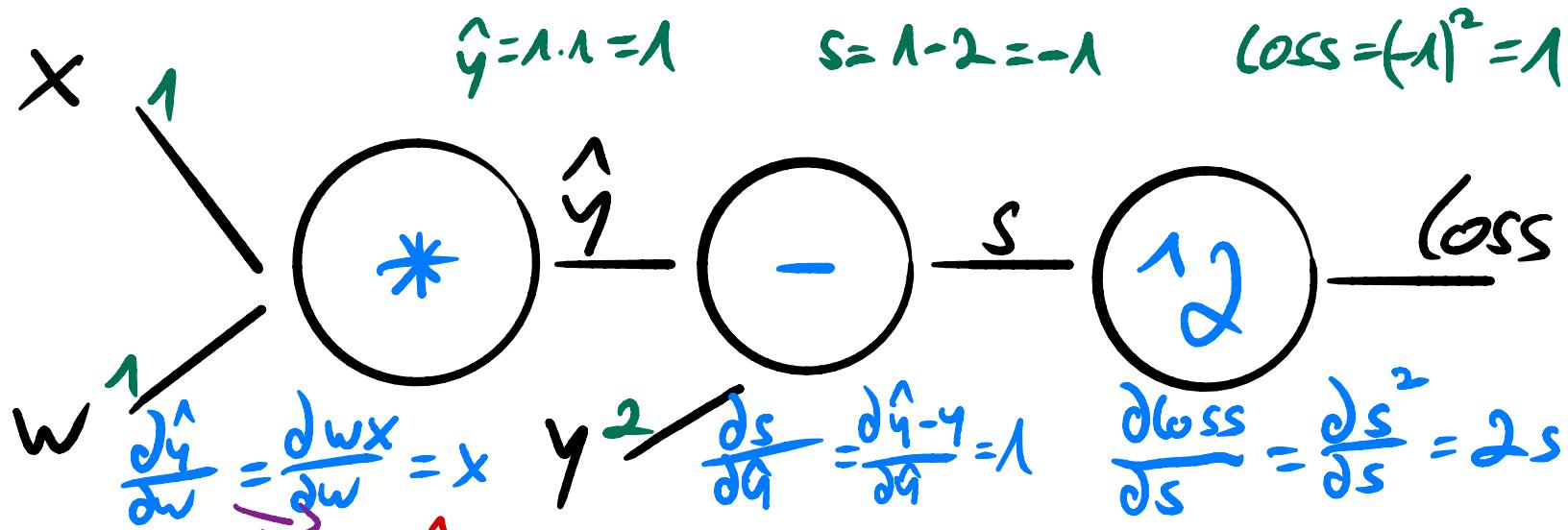
Backward pass

Forward pass $x=1$ $y=2$ $w=1$



Backward pass

Forward pass $x=1$ $y=2$ $w=1$

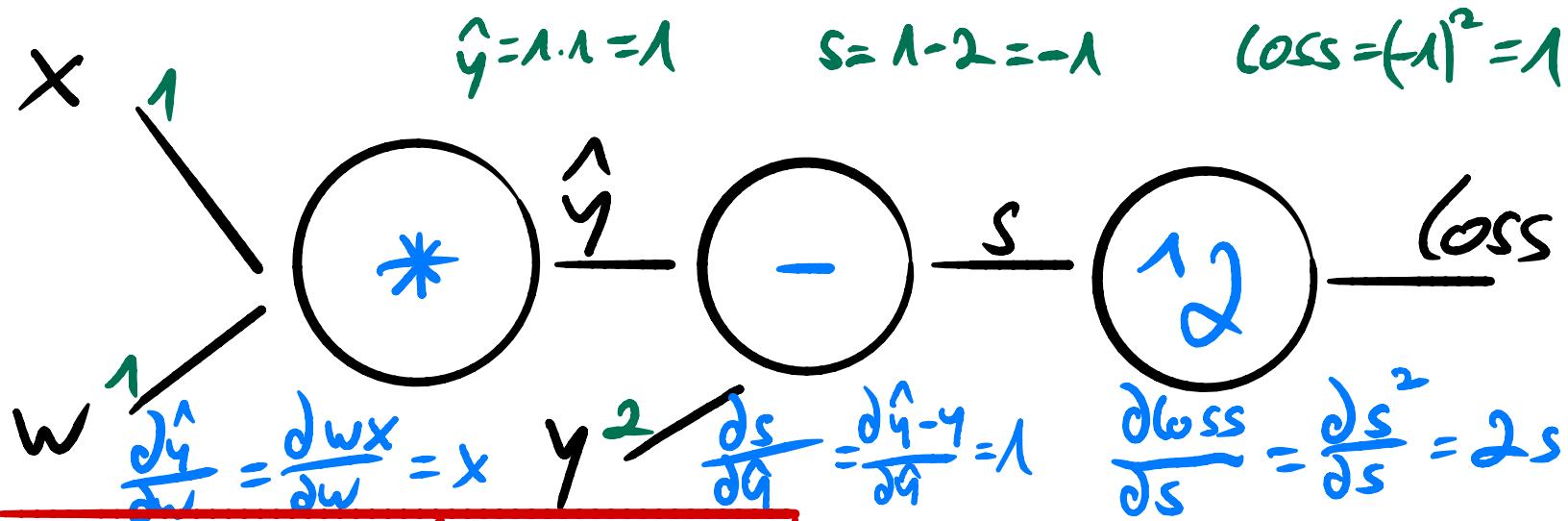


$$\frac{\partial \text{loss}}{\partial w} = \frac{\partial \text{loss}}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial w} = -2 \cdot x = -2$$

$$\frac{\partial \text{loss}}{\partial \hat{y}} = \frac{\partial \text{loss}}{\partial s} \cdot \frac{\partial s}{\partial \hat{y}} = 2 \cdot s \cdot 1 = -2$$

Backward pass

Forward pass $x=1$ $y=2$ $w=1$



$$\frac{\partial \text{loss}}{\partial w} = \frac{\partial \text{loss}}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial w} = -2 \cdot x = -2$$

$$\frac{\partial \text{loss}}{\partial \hat{y}} = \frac{\partial \text{loss}}{\partial s} \cdot \frac{\partial s}{\partial \hat{y}} = 2 \cdot s \cdot 1 = -2$$

Backward pass