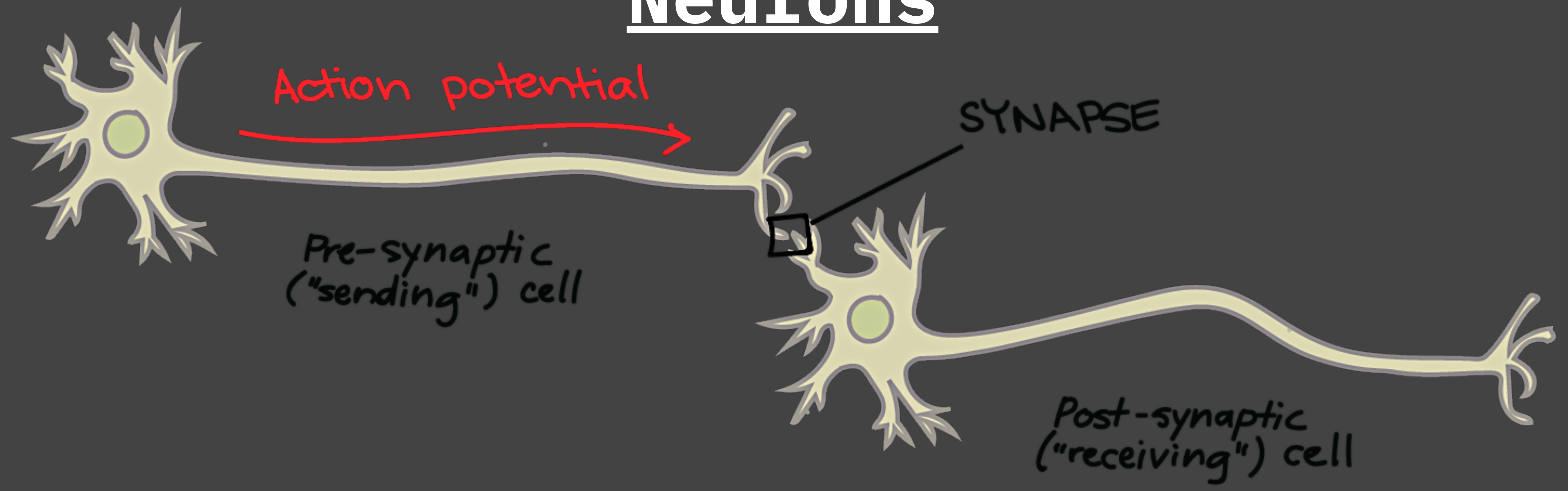
CS4047: Neural Nets (aNNs)

By Derek Rodriguez

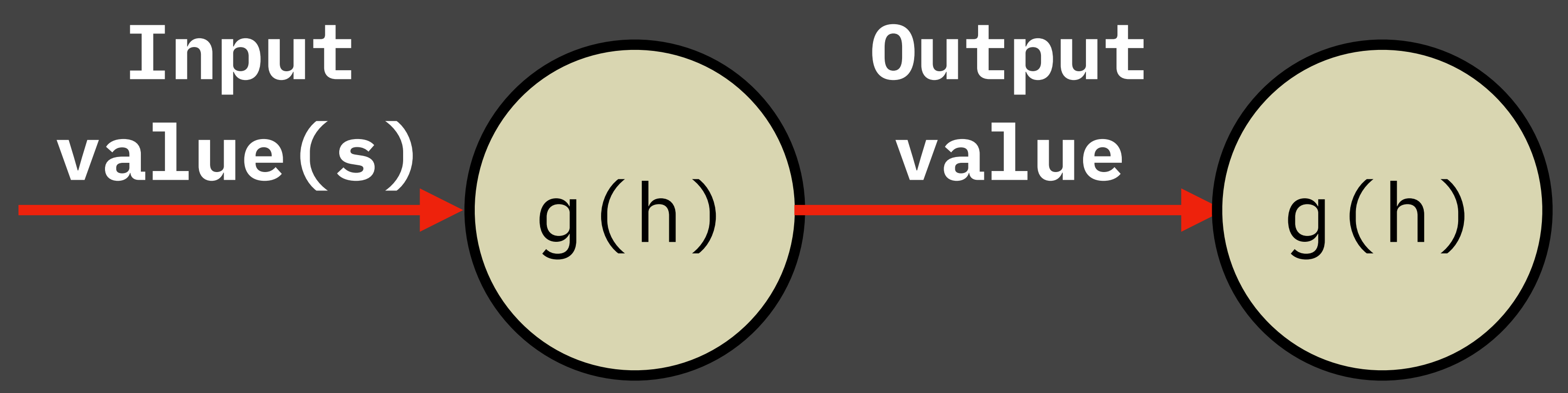
Origin of Perceptron

- Neurons produce electrical signals if certain chemical reactions happen inside them.
- Chemical reactions caused by signals coming from other neurons producing a synapse.
- Perceptrons are “digital neurons” that produce output based on input.

Neurons

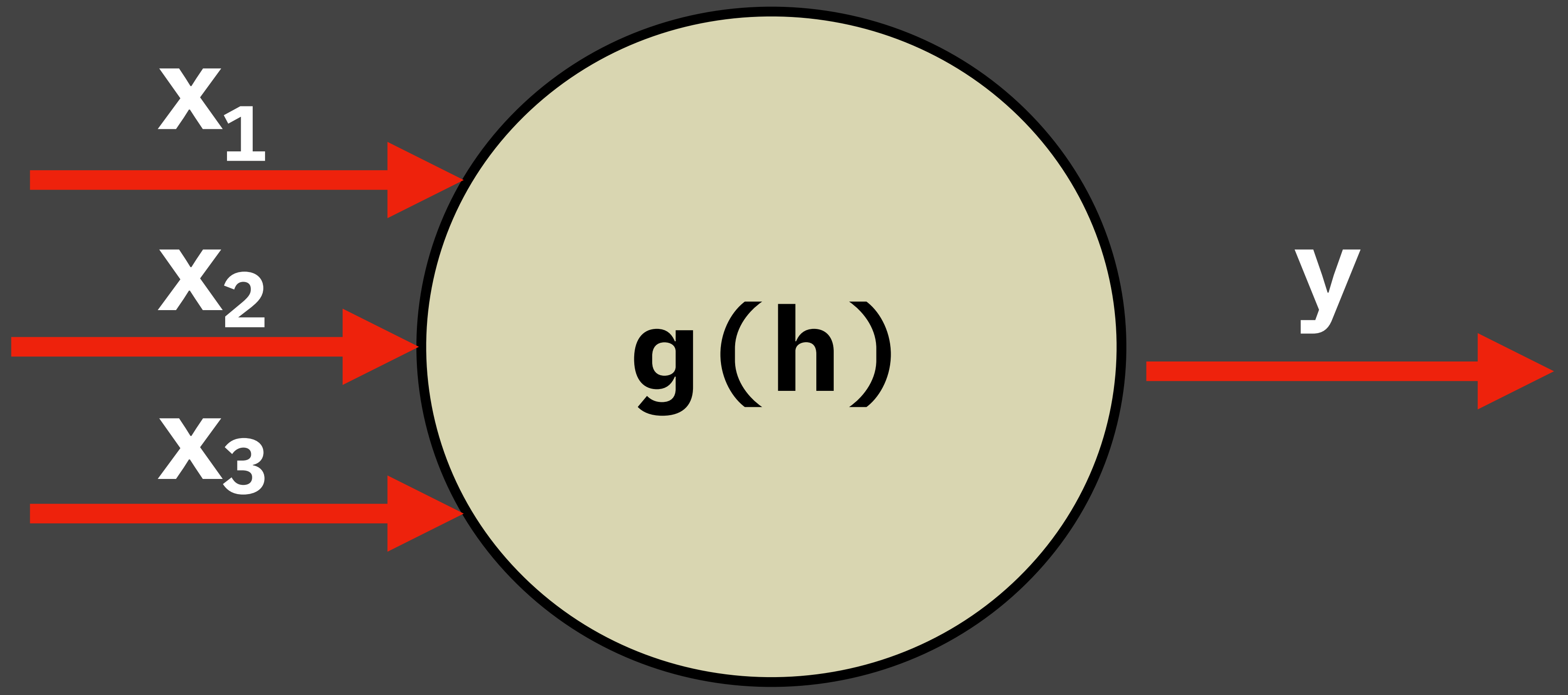


Perceptrons



Source:https://ka-perseus-images.s3.amazonaws.com/b404bfc9d6246ff3475f6fd2b34fbb7374fc4873.png

Basics of a Perceptron



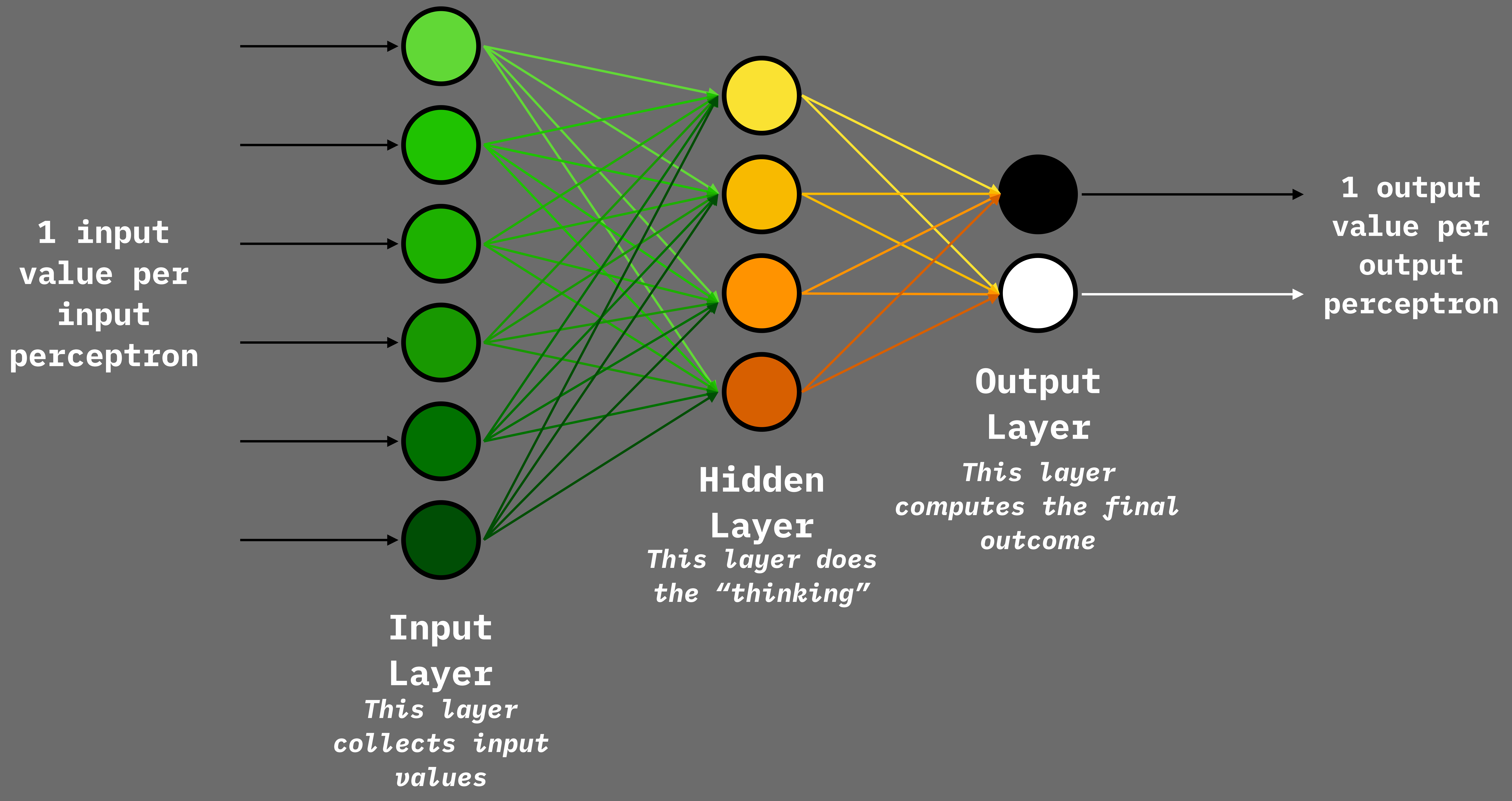
Symbol	Meaning
x	An input value
g()	Function applied to h
h	h = weighted avg. + bias
y	Output of g(h)
N	Amt. of inputs
W	A weight in the weighted avg.

g() can be any function, just needs to have a derivative.

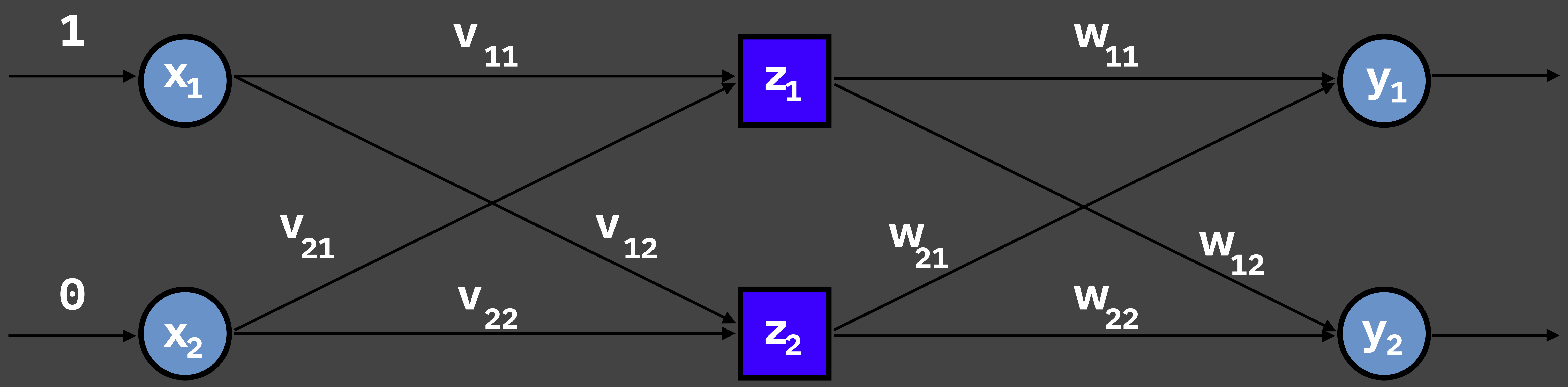
$$h = \sum_{n=1}^{N} (w_n x_n) + bias$$

Think of this as a fixed "threshold" for how hard it is to fire

Structure of Multi-layer aNN



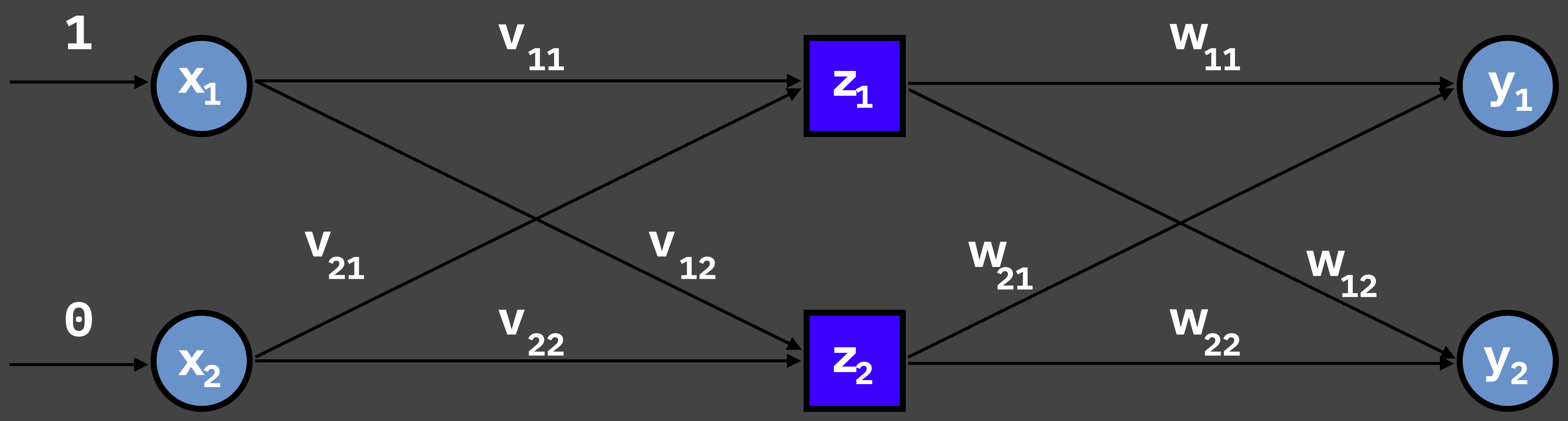
2015 Exam Problem



		Destination	
	v	1	2
Source	1	1	-1
	2	0	1
	w	1	2
Source	1	-1	0
	2	0	1

Given that we want $y_1=0$ and $y_2=1$, find all v and w after one iteration of back prop.

Deciphering variables



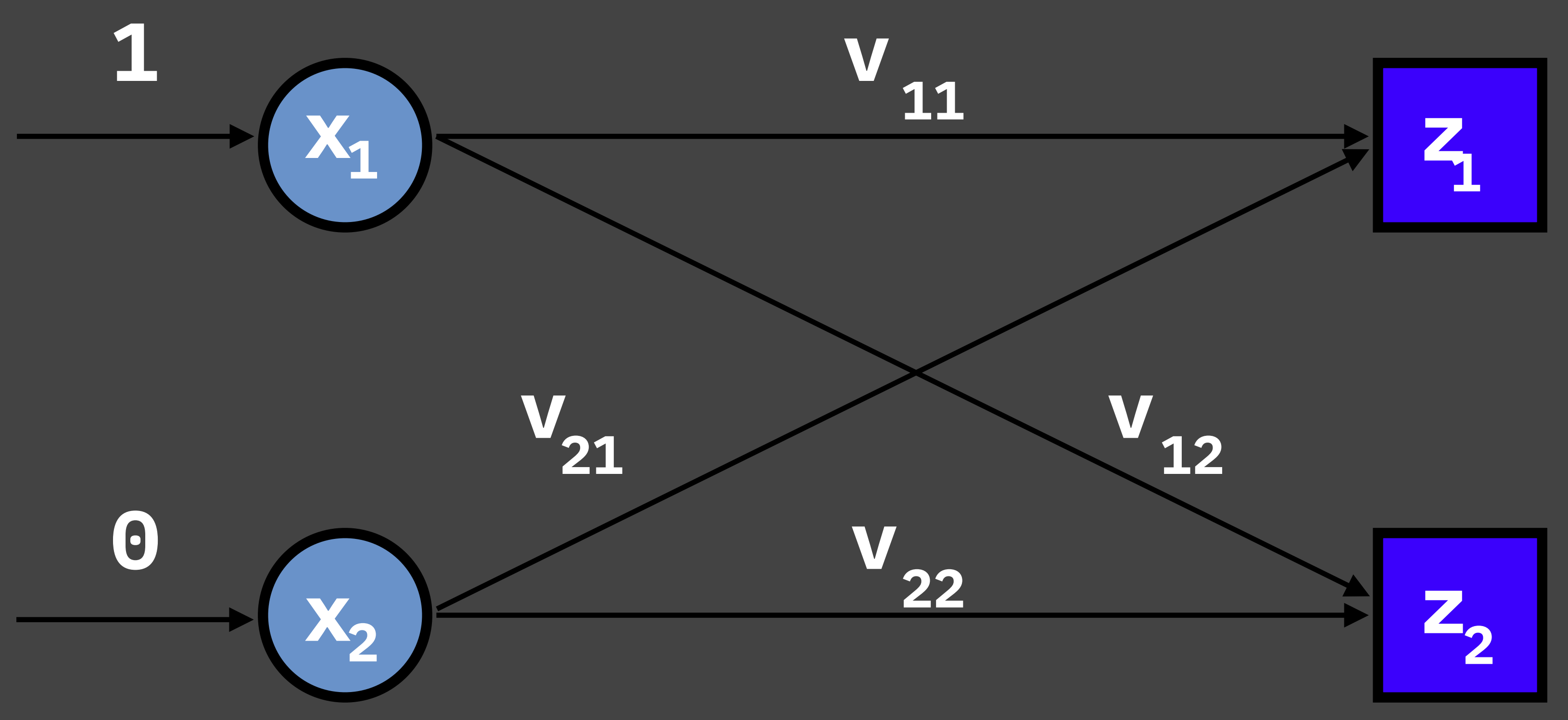
Symbol	Meaning
x	input perceptron
y	output perceptron
z	hidden perceptron
g()	Function applied to h
h	h = weighted avg. + bias
y	output perceptron
i	source index
j	destination index
v	weight into hidden layer
w	weight into output layer

- **For every perceptron, <u>bias is 1</u>**
- **For this problem, g() does nothing, i.e. g(h) = h**
- **For clarity, weights from input have renamed to v**
- **Final output of perceptron is named after perceptron**
- **Input perceptrons just push input values forward**
- **i and j don't appear explicitly, but we'll use them later for writing formulas.**
- **To make life easy, we never differentiate the bias**

Solution Strategy

1. Perform one iteration of forward prop to get current values.
2. Calculate output “change factor” for each perceptron(Δ).
3. Calculate changes to weights for output layer new w.
4. Use “change factor” from output layer to calculate “change factor” in hidden layer (δ).
5. Perform step 2 for hidden layer to get new v.

Step 1 Part 1: Forward Prop for z



Remember input layer just forwards values

Input formula for Z perceptron

$$h_{z_j} = \sum_{i=1}^{2}(v_{ij}x_i) + bias$$

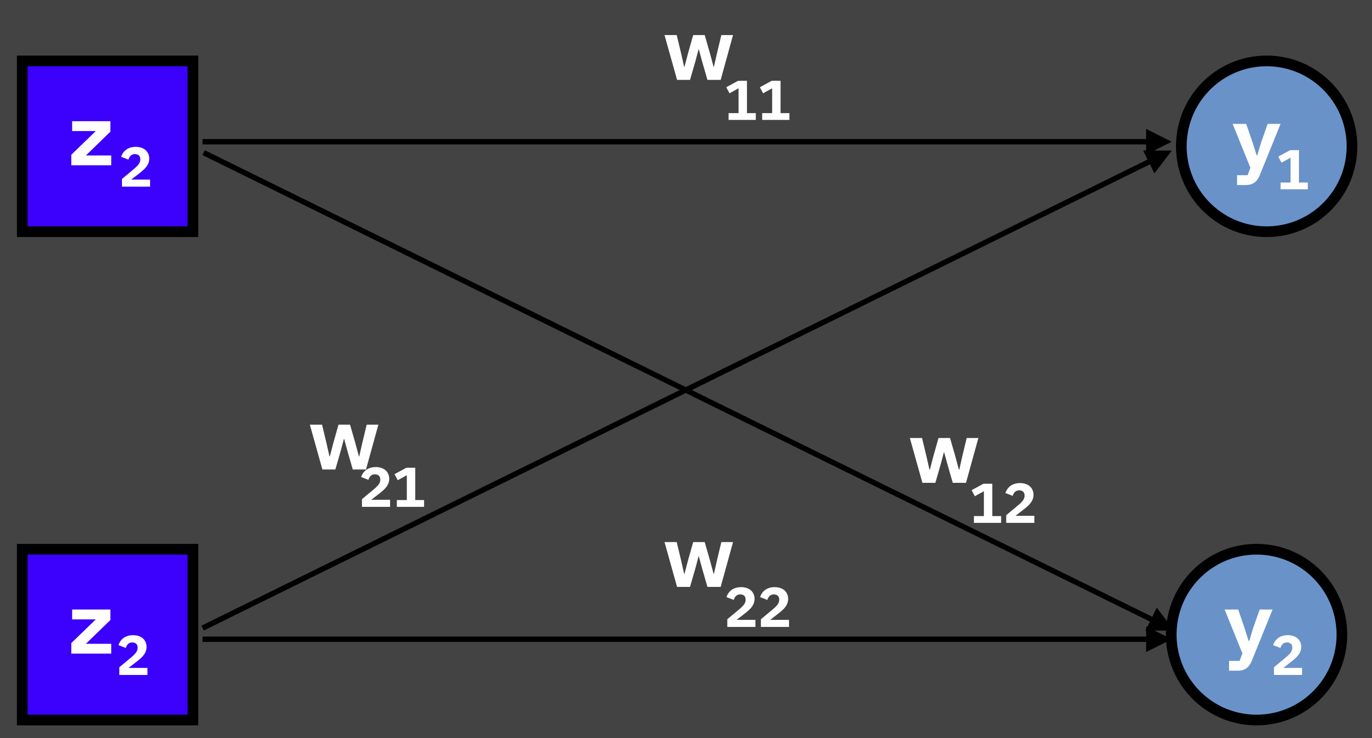
$$h_{z_1} = v_{11}x_1 + v_{21}x_2 + bias = (1*1)+(0*0)+1 = 2$$

This is how we keep track of which h is which

$$h_{z_2} = v_{12}x_1 + v_{22}x_2 + bias = (-1*1)+(1*0)+1 = 0$$

Remember that $g(h_z)=h_z=z$ so these will be the values passes forward

Step 1 Part 2: Forward Prop for y



Input formula for y perceptron

$$h_{y_j} = \sum_{i=1}^{2} \left(w_{ij} z_i \right) + bias$$

$$h_{y_1} = w_{11}z_1 + w_{21}z_2 + bias = (-1*0)+(0*0)+1 = \underline{1}$$

$$h_{y_2} = w_{12}z_1 + w_{22}z_2 + bias = (0*2)+(1*2)+1 = \underline{3}$$

This is roughly the same as the previous step

Step 2: Calculating Δs

- Let's call current output $\hat{y}$ and desired value y

Recall from question:

- $y_1 = 0$
- $y_2 = 1$
- $\hat{y}_1 = 1$
- $\hat{y}_2 = 3$

$$\Delta_1 = \hat{y}_1 - y_1 = 1-0 = \underline{1}$$

$$\Delta_2 = \hat{y}_2 - y_2 = 3-1 = \underline{2}$$

Step 3: update y perceptrons

- Let's call the new weight $w_{ij}^{(1)}$ and the old weight $w_{ij}^{(0)}$

$$w_{ij}^{(1)} = w_{ij}^{(0)} + \eta\Delta_j z_i$$

$$w_{11}^{(1)} = w_{11}^{(0)} + \eta\Delta_1 z_1 = -1 + 0.1*-1*1 = -1.1$$

$$w_{12}^{(1)} = w_{12}^{(0)} + \eta\Delta_2 z_1 = 0 + 0.1*3*1 = 0.3$$

$$w_{21}^{(1)} = w_{21}^{(0)} + \eta\Delta_1 z_2 = 0 + 0.1*-1*3 = -0.3$$

$$w_{22}^{(1)} = w_{22}^{(0)} + \eta\Delta_2 z_2 = 1 + 0.1*3*3 = 1.9$$

Step 4: Calculating δs

- This to calculate δ we use Δ and w

Recall:

- Δ_1 **= 1**
- Δ_2 **= 2**

$$\delta_1 = \Delta_1 w_{11} + \Delta_1 w_{12} = 1 * -1 + 2 * 0 = \underline{-1}$$

$$\delta_2 = \Delta_2 w_{21} + \Delta_2 w_{22} = 1 * 0 + 2 * 1 = \underline{2}$$

Step 5: update z perceptrons

- This is the same as step 3 but with v and δ instead of w and Δ.

$$v_{ij}^{(1)} = v_{ij}^{(0)} + \eta\delta_j x_i$$

$$v_{11}^{(1)} = v_{11}^{(0)} + \eta\delta_1 x_1 = -1 + 0.1*-1*1 = -1.1$$

$$v_{12}^{(1)} = v_{12}^{(0)} + \eta\delta_2 x_1 = 0 + 0.1*2*1 = 0.2$$

$$v_{21}^{(1)} = v_{21}^{(0)} + \eta\delta_1 x_2 = -1 + 0.1*-1*0 = -1.1$$

$$v_{22}^{(1)} = v_{22}^{(0)} + \eta\delta_2 x_2 = 1 + 0.1*2*0 = 1.2$$

And you're done!