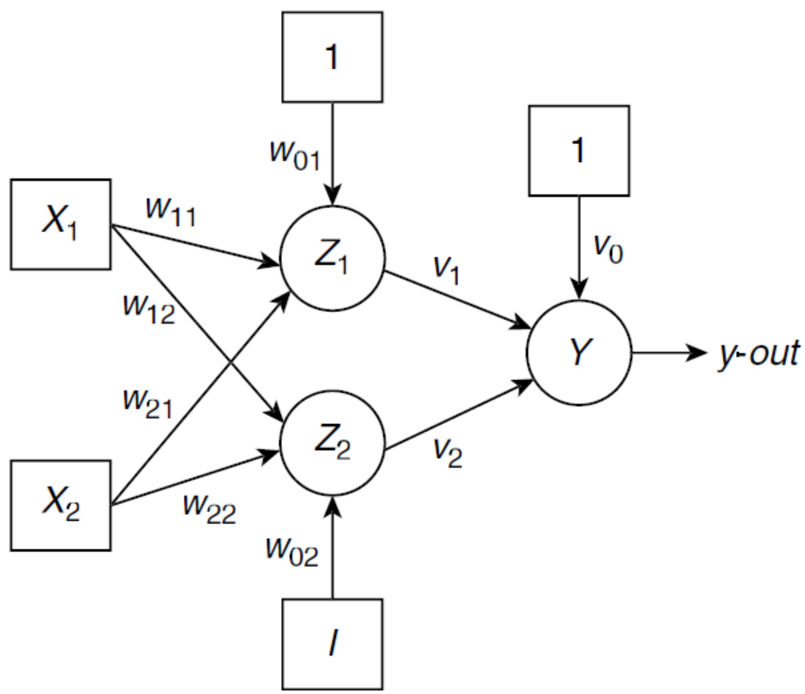
# MADALINE

Dr Dayal Kumar Behera

# MADALINE

- Several *ADALINE*s arranged in a multilayer net is known as *Many ADALINES*, or *MADALINE* *(Many Adaptive Linear Neurons*)

# Architecture




*A two input, one output, one hidden layer with two hidden units MADALINE*

# Learning

- There are two training algorithms for *MADALINE*, *viz*., *MR-I* and *MR-II*.

- In *MR-I* algorithm, only the weights of the hidden units are modified during the training.(weights for the inter-connections from the hidden units to the output unit are kept unaltered)

- However, in case of *MR-II*, all weights are adjusted, if required.

# Procedure *MADALINE-MR-I-Learning*

**Procedure** *MADALINE-MR-I-Learning*

**Step 1.** Initialize $v_0$, $v_1$, $v_2$ with 0.5 and other weights $w_{01}$, $w_{11}$, $w_{12}$, $w_{02}$, $w_{12}$ and $w_{22}$ by small random values. All bias inputs are set to 1.

**Step 2.** Set the learning rate $h$ to a suitable value.

**Step 3.** For each bipolar training pair $s : t$, do Steps 4-6

**Step 4.** Activate the input units: $x_1 = s_1$, $x_2 = s_2$, all biases are set to 1 permanently.

# Procedure *MADALINE-MR-I-Learning*

**Step 5.** Propagate the input signals through the net to the output unit $Y$.

5.1 Compute net inputs to the hidden units.

$$z\_in_1 = 1 \times w_{01} + x_1 \times w_{11} + x_2 \times w_{21}$$

$$z\_in_2 = 1 \times w_{02} + x_1 \times w_{12} + x_2 \times w_{22}$$

5.2 Compute activations of the hidden units $z\_out_1$ and $z\_out_2$ using the bipolar step function

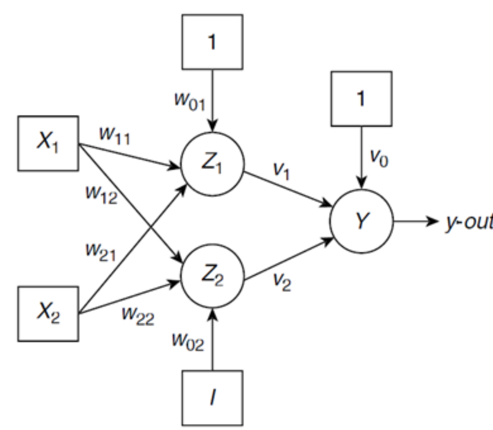
$$z\_out = \begin{cases} 1, & \text{if } z\_in \geq 0 \\ -1, & \text{if } z\_in < 0. \end{cases}$$

5.3 Compute net input to the output unit

$$y\_in = 1 \times v_0 + z\_out_1 \times v_0 + z\_out_2 \times v_2$$

5.4 Find the activation of the output unit *y-out* using the same activation function as in Step 5.2, *i.e.*,

$$y\_out = \begin{cases} 1, & \text{if } y\_in \geq 0 \\ -1, & \text{if } y\_in < 0. \end{cases}$$



# Procedure *MADALINE-MR-I-Learning*

**Step 6.** Adjust the weights of the hidden units, if required, according to the following rules:

i) If ($y\_out = t$) then the net yields the expected result. Weights need not be updated.

ii) If ($y\_out \neq t$) then apply one of the following rules whichever is applicable.

**Case I:** $t = 1$
Find the hidden unit $z_j$ whose net input $z\_in_j$ is closest to 0. Adjust the weights attached to $z_j$ according to the formula

$$w_{ij} \text{ (new)} = w_{ij} \text{ (old)} + h \times (1 - z\_in_j) \times x_i, \text{ for all } i.$$

**Case II:** $t = -1$
Adjust the weights attached to those hidden units $z_j$ that have positive net input.

$$w_{ij} \text{ (new)} = w_{ij} \text{ (old)} + h \times (-1 - z\_in_j) \times x_i, \text{ for all } i.$$

**Step 7.** Test for stopping condition. It can be any one of the following:

i) No change of weight occurs in Step 6.
ii) The weight adjustments have reached an acceptable level.
iii) A predefined number of iterations have been carried out.

If the stopping condition is satisfied then stop. Otherwise go to Step 3.

# Example

- Let us train a *MADALINE* net through the *MR-I* algorithm to realize the two-input *XOR* function by assuming initial weights and learning rate as below.

**Table 7.9.** *Bipolar training set for XOR function*

| $x_0$ | $x_1$ | $x_2$ | $t$ |
|---|---|---|---|
| 1 | 1 | 1 | −1 |
| 1 | 1 | −1 | 1 |
| 1 | −1 | 1 | 1 |
| 1 | −1 | −1 | −1 |

**Table 7.10.** *Initial weights and the fixed learning rate*

| $w_{01}$ | $w_{11}$ | $w_{21}$ | $w_{02}$ | $w_{12}$ | $w_{22}$ | $\eta$ |
|---|---|---|---|---|---|---|
| .2 | .3 | .2 | .3 | .2 | .1 | .5 |

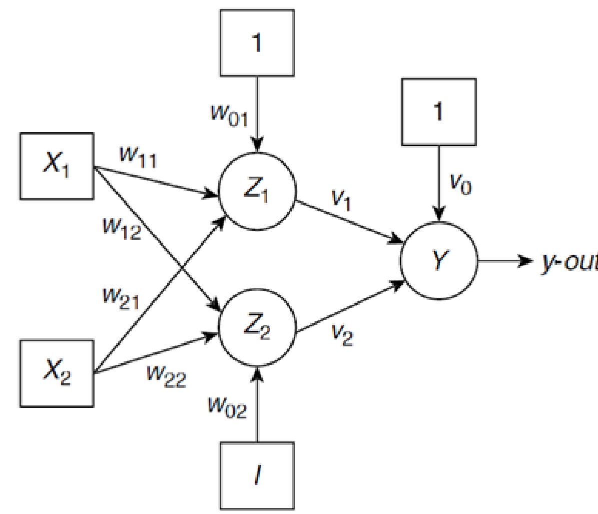
# Calculation

- $z\_in1 = 1 \times w01 + x1 \times w11 + x2 \times w21$
  $= 1 \times .2 + 1 \times .3 + 1 \times .2 = .7$

- $z\_out1 = 1$

- $z\_in2 = 1 \times w02 + x1 \times w12 + x2 \times w22$
  $= 1 \times .3 + 1 \times .2 + 1 \times .1 = .6$

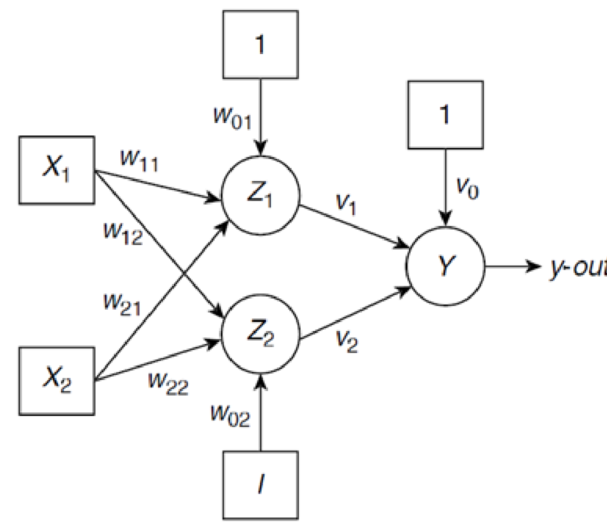
- $z\_out2 = 1$



- $y\_in = 1 \times v0 + z\_out1 \times v1 + z\_out2 \times v2$

  $= 1 \times .5 + 1 \times .5 + 1 \times .5 = 1.5$

- $y\_out = 1$



# Weight Adjustment

- $w01\ (new) = w01\ (old) + h \times (-1 - z\_in1)$

$= .2 + .5 \times (-1 - .7)$

$= .2 - .85$

$= -.65$

- $w11\ (new) = w11\ (old) + h \times (-1 - z\_in1)$

$= .3 - .85$

$= -.55$

# Weight Adjustment

- $w21\ (new) = w21\ (old) + h \times (-1 - z\_in1)$

$= .2 - .85$

$= -.65$

- $w02\ (new) = w02\ (old) + h \times (-1 - z\_in2)$

$= .3 + .5 \times (-1 - .6)$

$= .3 - .8$

$= -.5$

# Weight Adjustment

- $w12\ (new) = w12\ (old) + h \times (-1 - z\_in2)$

$= .2 - .8$

$= -.6$

- $w22\ (new) = w22\ (old) + h \times (-1 - z\_in2)$

$= .1 - .8$

$= -.7$

# Weight Adjustment

- Hence the new set of weights after training with the first training pair (1, 1) : −1 in the **first epoch** is obtained as

$$W = \begin{bmatrix} w_{01} & w_{02} \\ w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} = \begin{bmatrix} -.65 & -.5 \\ -.55 & -.6 \\ -.65 & -.7 \end{bmatrix}$$

# MR-I: XOR Function

**Table 7.11.** *MADALINE Learning of XOR Function through MR-I algorithm*

| # | $x_0$ | $x_1$ | $x_2$ | t | $z\_in_1$ | $z\_in_2$ | $z\_out_1$ | $z\_out_2$ | $y\_in$ | $y\_out$ | $w_{01}$ | $w_{11}$ | $w_{21}$ | $w_{02}$ | $w_{12}$ | $w_{22}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | | | | | | | | | | .2 | .3 | .2 | .3 | .2 | .1 |
| 1 | 1 | 1 | 1 | −1 | .7 | .6 | 1 | 1 | 1.5 | 1 | −.65 | −.55 | −.65 | −.5 | −.6 | −.7 |
| 2 | 1 | 1 | −1 | 1 | −.55 | −.4 | −1 | −1 | −.5 | −1 | −.65 | −.55 | −.65 | .2 | .1 | −1.4 |
| 3 | 1 | −1 | 1 | 1 | −.75 | −1.3 | −1 | −1 | −.5 | −1 | .23 | −1.43 | .13 | .2 | .1 | −1.4 |
| 4 | 1 | −1 | −1 | −1 | 1.56 | 1.5 | 1 | 1 | 1.5 | 1 | −1.05 | −.15 | 1.41 | −1.05 | 1.35 | −.15 |

Epoch #1

# MR-I: XOR Function

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | | | | | | | | | | −1.05 | −.15 | 1.41 | −1.05 | 1.35 | −.15 |
| 1 | 1 | 1 | 1 | −1 | .21 | .15 | 1 | 1 | 1.5 | 1 | −1.66 | −.76 | .8 | −1.63 | .77 | −.73 |
| 2 | 1 | 1 | −1 | 1 | −3.22 | −.13 | −1 | −1 | −.5 | −1 | −1.66 | −.76 | .8 | −2.07 | .33 | −.29 |
| 3 | 1 | −1 | 1 | 1 | −.1 | −1.45 | −1 | −1 | −.5 | −1 | −2.11 | −.31 | .35 | −2.07 | .33 | −.29 |
| 4 | 1 | −1 | −1 | −1 | −2.15 | −2.11 | −1 | −1 | −.5 | −1 | | | | | - | |

Epoch #2

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | | | | | | | | | | −2.11 | −.31 | .35 | −2.07 | .33 | −.29 |
| 1 | 1 | 1 | 1 | −1 | −2.07 | −2.03 | −1 | −1 | −.5 | −1 | | | | | | |
| 2 | 1 | 1 | −1 | 1 | −2.77 | −1.45 | −1 | −1 | −.5 | −1 | −2.11 | −.31 | .35 | −.84 | 1.56 | −1.52 |
| 3 | 1 | −1 | 1 | 1 | −1.45 | −3.92 | −1 | −1 | −.5 | −1 | −.88 | −1.54 | .88 | −.84 | 1.56 | −1.52 |
| 4 | 1 | −1 | −1 | −1 | −.22 | −.88 | −1 | −1 | −.5 | −1 | | | | | | |

Epoch #3

# MR-I: XOR Function

| # | $x_0$ | $x_1$ | $x_2$ | $t$ | $z\_in_1$ | $z\_in_2$ | $z\_out_1$ | $z\_out_2$ | $y\_in$ | $y\_out$ | $w_{01}$ | $w_{11}$ | $w_{21}$ | $w_{02}$ | $w_{12}$ | $w_{22}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | | | | | | | | | | −.88 | −1.54 | .88 | −.84 | 1.56 | −1.52 |
| 1 | 1 | 1 | 1 | −1 | −1.54 | −.8 | −1 | −1 | −.5 | −1 | | | | | | |
| 2 | 1 | 1 | −1 | 1 | −3.3 | 2.24 | −1 | 1 | .5 | 1 | | | | | | |
| 3 | 1 | −1 | 1 | 1 | 1.54 | −3.92 | 1 | −1 | .5 | 1 | | | | | | |
| 4 | 1 | −1 | −1 | −1 | −.22 | −.88 | −1 | −1 | −.5 | −1 | | | | | | |

Epoch #4