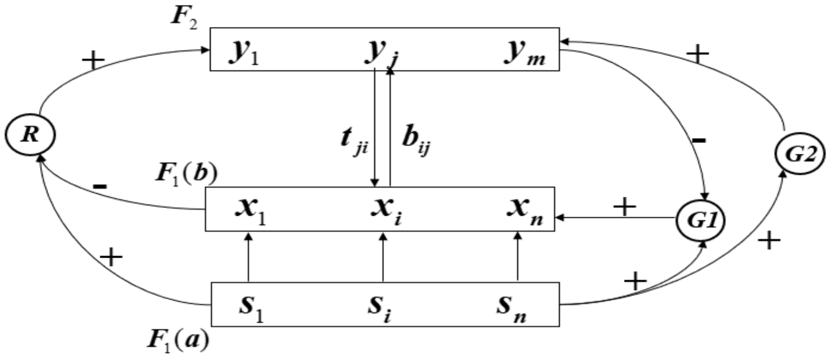
ART1 Architecture



$F_1(a)$: input units
$F(b)_1$: interface units
F_2 : cluster units
$F_1(a)$ to $F_1(b)$: pair - wise connection
between F_2 and $F_1(b)$: full connection

b_{ij} : bottom up weights from x_i to y_j (real value)
t_{ji} : top down weights from y_j to x_i (representing class j binary/bipolar)
R, $G1$, $G2$: control units

n	number of components in the input vector.
m	maximum number of clusters that can be formed.
b_{ij}	bottom-up weights (from $F_1(b)$ unit X_i to F_2 unit Y_j).
t_{ji}	top-down weights (from F_2 unit Y_j to F_1 unit X_i).
ρ	vigilance parameter.
$\mathbf{s}$	binary input vector (an n-tuple).
$\mathbf{x}$	activation vector for $F_1(b)$ layer (binary).
$\|\mathbf{x}\|$	norm of vector $\mathbf{x}$, defined as the sum of the components x_i.

Parameters

User-defined parameters with restrictions indicated for permissible values [Carpenter & Grossberg, 1987a] and sample values [Lippmann, 1987] are as follows:

PARAMETER	PERMISSIBLE RANGE	SAMPLE VALUE
L	$L > 1$	2
ρ	$0 < \rho \le 1$ (vigilance parameter)	.9
b_{ij}	$0 < b_{ij}(0) < \dfrac{L}{L - 1 + n}$ (bottom-up weights)	$\dfrac{1}{1 + n}$
t_{ji}	$t_{ji}(0) = 1$ (top-down weights)	1

Example 5.1 An ART1 net to cluster four vectors: low vigilance

The values and a description of the parameters in this example are:

n	$= 4$	number of components in an input vector;
m	$= 3$	maximum number of clusters to be formed;
ρ	$= 0.4$	vigilance parameter;
L	$= 2$	parameter used in update of bottom-up weights;
$b_{ij}(0)$	$= \dfrac{1}{1+n}$	initial bottom-up weights (one-half the maximum value allowed);
$t_{ji}(0)$	$= 1$	initial top-down weights.

The example uses the ART1 algorithm to cluster the vectors (1, 1, 0, 0), (0, 0, 0, 1), (1, 0, 0, 0), and (0, 0, 1, 1), in at most three clusters.

Application of the algorithm yields the following:

Step 0. Initialize parameters:

$$L = 2,$$

$$\rho = 0.4;$$

Initialize weights:

$$b_{ij}(0) = 0.2,$$

$$t_{ji}(0) = 1.$$

Step 1. Begin computation.

Step 2. For the first input vector, (1, 1, 0, 0), do Steps 3–12.

Step 3. Set activations of all F_2 units to zero.
Set activations of $F_1(a)$ units to input vector
$\mathbf{s} = (1, 1, 0, 0)$.

Step 4. Compute norm of **s**:

$$\|\mathbf{s}\| = 2.$$

Step 5. Compute activations for each node in the F_1 layer:

$$\mathbf{x} = (1, 1, 0, 0).$$

Step 6. Compute net input to each node in the F_2 layer:

$$y_1 = .2(1) + .2(1) + .2(0) + .2(0) = 0.4,$$

$$y_2 = .2(1) + .2(1) + .2(0) + .2(0) = 0.4,$$

$$y_3 = .2(1) + .2(1) + .2(0) + .2(0) = 0.4.$$

Step 7. While reset is true, do Steps 8–11.

Step 8. Since all units have the same net input,

$$J = 1.$$

Step 9. Recompute the F_1 activations:

$x_i = s_i t_{1i}$; currently, $t_1 = (1, 1, 1, 1)$;

therefore, $\mathbf{x} = (1, 1, 0, 0)$

Step 10. Compute the norm of $\mathbf{x}$:

$\|\mathbf{x}\| = 2.$

Step 11. Test for reset:

$$\frac{\|\mathbf{x}\|}{\|\mathbf{s}\|} = 1.0 \geq 0.4;$$

therefore, reset is false.
Proceed to Step 12.

Step 12. Update b_1; for $L = 2$, the equilibrium weights are

$$b_{i1}(\text{new}) = \frac{2x_i}{1 + \|\mathbf{x}\|}.$$

Therefore, the bottom-up weight matrix becomes

$$\begin{bmatrix} .67 & .2 & .2 \\ .67 & .2 & .2 \\ 0 & .2 & .2 \\ 0 & .2 & .2 \end{bmatrix}$$

Update $\mathbf{t}_1$; the fast learning weight values are

$t_{Ji}(\text{new}) = x_i$,

therefore, the top-down weight matrix becomes

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Step 2. For the second input vector, (0, 0, 0, 1), do Steps 3–12.

Step 3. Set activations of all F_2 units to zero.
Set activations of $F_1(a)$ units to input vector
$\mathbf{s} = (0, 0, 0, 1)$.

Step 4. Compute norm of $\mathbf{s}$:

$\|\mathbf{s}\| = 1.$

Step 5. Compute activations for each node in the F_1 layer:

$\mathbf{x} = (0, 0, 0, 1).$

Step 6. Compute net input to each node in the F_2 layer:

$$y_1 = .67(0) + .67(0) + 0(0) + 0(1) = 0.0,$$
$$y_2 = .2(0) + .2(0) + .2(0) + .2(1) = 0.2,$$
$$y_3 = .2(0) + .2(0) + .2(0) + .2(1) = 0.2.$$

Step 7. While reset is true, do Steps 8–11.

Step 8. Since units Y_2 and Y_3 have the same net input

$J = 2$.

Step 9. Recompute the activation of the F_1 layer:

$x_i = s_i t_{2i}$;

currently $\mathbf{t}_2 = (1, 1, 1, 1)$; therefore,

$\mathbf{x} = (0, 0, 0, 1)$.

Step 10. Compute the norm of $\mathbf{x}$:

$\|\mathbf{x}\| = 1$.

Step 11. Test for reset:

$$\frac{\|\mathbf{x}\|}{\|\mathbf{s}\|} = 1.0 \geq 0.4;$$

therefore, reset is false. Proceed to Step 12.

Step 12. Update $\mathbf{b}_2$; the bottom-up weight matrix becomes

$$\begin{bmatrix} .67 & 0 & .2 \\ .67 & 0 & .2 \\ 0 & 0 & .2 \\ 0 & 1 & .2 \end{bmatrix}$$

Update $\mathbf{t}_2$; the top-down weight matrix becomes

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Step 2. For the third input vector, (1, 0, 0, 0), do Steps 3–12.

Step 3. Set activations of all F_2 units to zero.

Set activations of $F_1(a)$ units to input vector

$\mathbf{s} = (1, 0, 0, 0)$.

Step 4. Compute norm of $\mathbf{s}$:

$\|\mathbf{s}\| = 1$.

Step 5. Compute activations for each node in the F_1 layer:

$\mathbf{x} = (1, 0, 0, 0)$.

Step 6. Compute net input to each node in the F_2 layer:

$$y_1 = .67(1) + .67(0) + 0(0) + 0(0) = 0.67,$$
$$y_2 = 0(1) + 0(0) + 0(0) + 1(0) = 0.0,$$
$$y_3 = .2(1) + .2(0) + .2(0) + .2(0) = 0.2.$$

Step 7. While reset is true, do Steps 8–11.

Step 8. Since unit Y_1 has the largest net input,

$$J = 1.$$

Step 9. Recompute the activation of the F_1 layer:

$$x_i = s_i t_{1i};$$

current, $t_1 = (1, 1, 0, 0)$; therefore,

$$\mathbf{x} = (1, 0, 0, 0).$$

Step 10. Compute the norm of $\mathbf{x}$:

$$\|\mathbf{x}\| = 1.$$

Step 11. $\|\mathbf{x}\| / \|\mathbf{s}\| = 1.0$ Proceed to Step 12.

Step 12. Update $\mathbf{b}_1$; the bottom-up weight matrix becomes

$$\begin{bmatrix} 1 & 0 & .2 \\ 0 & 0 & .2 \\ 0 & 0 & .2 \\ 0 & 1 & .2 \end{bmatrix}$$

Update $\mathbf{t}_1$; the top-down weight matrix becomes

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Step 2. For the fourth input vector, (0, 0, 1, 1), do Steps 3–12.

Step 3. Set activations of all F_2 units to zero.

Set activations of $F_1(a)$ units to input vector

$$\mathbf{s} = (0, 0, 1, 1).$$

Step 4. Compute norm of $\mathbf{s}$:

$$\|\mathbf{s}\| = 2.$$

Step 5. Compute activations for each node in the F_1 layer:

$$\mathbf{x} = (0, 0, 1, 1).$$

Step 6. Compute net input to each node in the F_2 layer:

$$y_1 = 1(0) + 0(0) + 0(1) + 0(1) = 0.0,$$
$$y_2 = 0(0) + 0(0) + 0(1) + 1(1) = 1.0,$$
$$y_3 = .2(0) + .2(0) + .2(1) + .2(1) = 0.4.$$

Step 7. While reset is true, do Steps 8–11.

Step 8. Since unit Y_2 has the largest net input,

$$J = 2.$$

Step 9. Recompute the activation of the F_1 layer:

$$x_i = s_i t_{2i};$$

currently, $t_2 = (0, 0, 0, 1)$; therefore,

$$\mathbf{x} = (0, 0, 0, 1).$$

Step 10. Compute the norm of $\mathbf{x}$:

$$\|\mathbf{x}\| = 1.$$

Step 11. Test for reset:

$$\frac{\|\mathbf{x}\|}{\|\mathbf{s}\|} = 0.5 \geq 0.4;$$

therefore, reset is false. Proceed to Step 12.

Step 12. Update $\mathbf{b}_2$; however, there is no change in the bottom-up weight matrix:

$$\begin{bmatrix} 1 & 0 & .2 \\ 0 & 0 & .2 \\ 0 & 0 & .2 \\ 0 & 1 & .2 \end{bmatrix}$$

Similarly, the top-down weight matrix remains

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Step 13. Test stopping condition.
(This completes one epoch of training.)

Training algorithm

The training algorithm for an ART1 net is presented next. A discussion of the role of the parameters and an appropriate choice of initial weights follows.

Step 0. Initialize parameters:

$$L > 1,$$

$$0 < \rho \leq 1.$$

Initialize weights:

$$0 < b_{ij}(0) < \frac{L}{L - 1 + n},$$

$$t_{ji}(0) = 1.$$

Step 1. While stopping condition is false, do Steps 2–13.

Step 2. For each training input, do Steps 3–12.

Step 3. Set activations of all F_2 units to zero.
Set activations of $F_1(a)$ units to input vector $\mathbf{s}$.

Step 4. Compute the norm of $\mathbf{s}$:

$$\|\mathbf{s}\| = \sum_i s_i.$$

Step 5. Send input signal from $F_1(a)$ to the $F_1(b)$ layer:

$$x_i = s_i.$$

Step 6. For each F_2 node that is not inhibited:
If $y_j \neq -1$, then

$$y_j = \sum_i b_{ij} x_i.$$

Step 7. While reset is true, do Steps 8–11.

Step 8. Find J such that $y_J \geq y_j$ for all nodes j.
If $y_J = -1$, then all nodes are inhibited and this pattern cannot be clustered.

Step 9. Recompute activation $\mathbf{x}$ of $F_1(b)$:

$$x_i = s_i t_{Ji}.$$

Step 10. Compute the norm of vector $\mathbf{x}$:

$$\|\mathbf{x}\| = \sum_i x_i.$$

Step 11. Test for reset:

If $\frac{\|\mathbf{x}\|}{\|\mathbf{s}\|} < \rho$, then

$y_J = -1$ (inhibit node J) (and continue, executing Step 7 again).

If $\frac{\|\mathbf{x}\|}{\|\mathbf{s}\|} \geq \rho$,

then proceed to Step 12.

Step 12. Update the weights for node J (fast learning):

$$b_{iJ}(\text{new}) = \frac{Lx_i}{L - 1 + \|\mathbf{x}\|},$$

$$t_{Ji}(\text{new}) = x_i.$$

Step 13. Test for stopping condition.

Question

The same vectors are presented to the ART1 net (in the same order) as in Example 5.1. The vigilance parameter is set at 0.7. The training for vectors (1, 1, 0, 0), (0, 0, 0, 1), and (1, 0, 0, 0) proceeds as before, giving the bottom-up weight matrix

$$\begin{bmatrix} 1 & 0 & .2 \\ 0 & 0 & .2 \\ 0 & 0 & .2 \\ 0 & 1 & .2 \end{bmatrix}$$

and the top-down weight matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$