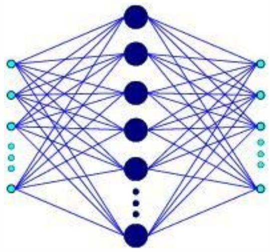
Ciência da Computação

REDE NEURAIS

Semestre: 2010/1

AULA 12



Professor: Max Pereira

http://paginas.unisul.br/max.pereira

UNISUL

Aqui seu futuro acontece

Conteúdo

- Redes ART
- Exercícios

ART
Adaptive Resonance Theory

- Capenter e Grossberg (1987)
- ART1 projetada para clusterização de vetores binários.
- ART2 aceita valores contínuos
- Aprendizado não-supervisionado / competitivo
- Controlam o grau de similaridade dos padrões colados no mesmo *cluster*.
- Outros modelos: Fuzzy-ART, ARTMAP, Fuzzy-ARTMAP

ART
Adaptive Resonance Theory

- Utiliza um mecanismo de realimentação entre a camada competitiva e a camada de entrada.
- O nome ART vem da forma de aprendizado e a recuperação da informação.
- A informação, na forma de saída dos neurônios, se propaga para frente e para trás entre as camadas.
- Se um determinado padrão é *clusterizado*, aparece uma oscilação estável, o que é equivalente ao conceito de ressonância para a rede neural.

Arquitetura Básica

- Envolve três grupos de neurônios: uma unidade de processamento (camada *F1*), unidades *cluster* (camada *F2*), e um mecanismo para controlar o grau de similaridade dos padrões colocados no mesmo *cluster* (um mecanismo de *reset*).
- A camada *F1* consiste de duas partes: a parte de *entrada F1(a)* e a parte de *interface F1(b)*.

Arquitetura Básica

- Para controlar a similaridade entre os padrões do mesmo *cluster*, há dois conjuntos de conexões (cada um com seus próprios pesos) entre cada unidade da camada de interface e cada unidade *cluster*.
- A camada *F1(b)* se conecta com a camada *F2* através de pesos *bottom-up*; os pesos *bottom-up* da conexão da i-ésima unidade *F1* com a j-ésima unidade *F2* são designados como b_{ij}

Arquitetura Básica

- A camada *F2* se conecta com a camada *F1(b)* através de pesos *top-down*; os pesos *top-down* da conexão da j-ésima unidade *F2* como a i-ésima unidade *F1* são designados como t_{ji}
- A camada *F2* é competitiva. A unidade *cluster* com o maior valor de entrada se torna o candidato para aprender o padrão de entrada.
- As ativações de todas as outras unidades de *F2* são definidas como zero.
- As unidades de interface combinam informações da camada de entrada e unidades *cluster*.

Arquitetura Básica

- Se é permitido ou não à unidade *cluster* aprender o padrão de entrada depende de quão similar é o seu vetor *top-down* com o vetor de entrada.
- Essa decisão é tomada pela unidade de *reset*, baseado nos sinais que ela recebe da entrada (*a*) e interface (*b*) da camada *F1*.

Arquitetura

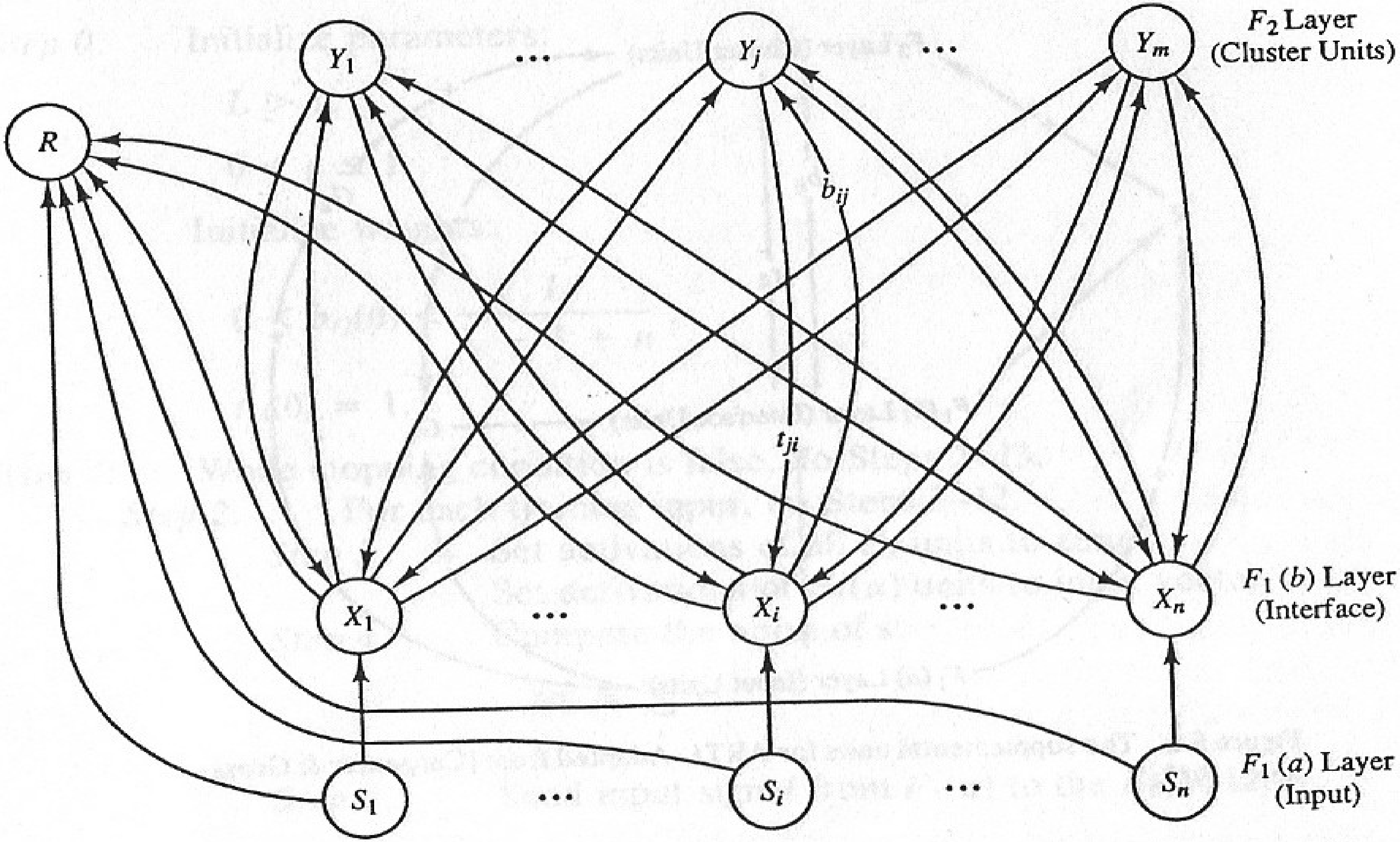


Figure 5.1 Basic structure of ART1.

Algoritmo

- Notação Utilizada:

n – número de componentes do vetor de entrada
m – número máximo de *clusters*
b_{ij}– pesos *bottom-up*
t_{ji} – pesos *top-down*
ρ - parâmetro de vigilância
s – vetor de entrada binário
x – vetor de ativação para a camada *F1(b)*
$||x||$ - soma dos componentes do vetor **x** (x_i)
L – parâmetro para ajuste dos pesos bottom-up

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 **s**.

Step 4. Compute the norm of **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 **x** of $F_1(b)$:

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

Step 10. Compute the norm of vector **x**:

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

Step 11. Test for reset:

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

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

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

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.

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 $\mathbf{s}$:

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

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_J; 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. $\|x\| / \|s\| = 1.0$ 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.)

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 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, $\mathbf{t}_2 = (0, 0, 0, 1)$; therefore,

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

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

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

Step 11. Check the vigilance criterion:

$$\frac{\|\mathbf{x}\|}{\|\mathbf{s}\|} = 0.5 < 0.7;$$

therefore, reset is true, so inhibit Y_2:

$$y_2 = -1.0.$$

Proceed with Step 7.

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

Step 8. Now the values for the F_2 layer are

$$y_1 = 0.0,$$

$$y_2 = -1.0,$$

$$y_3 = 0.4.$$

So unit Y_3 has the largest net input, and

$$J = 3.$$

Step 9. Recompute the activation of the F_1 layer:

$$x_i = s_i t_{3i};$$

currently, $\mathbf{t}_3 = (1, 1, 1, 1)$; therefore,

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

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

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

Step 11. Test for reset:

$$\frac{\|\mathbf{x}\|}{\|\mathbf{s}\|} = 1.0 \geq 0.7;$$

therefore, reset is false. Proceed with Step 12.

Step 12. Update $\mathbf{b}_3$; the bottom-up weight matrix becomes

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

Update $\mathbf{t}_3$; the top-down weight matrix becomes

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

Step 13. Test stopping condition.