Adaptive Resonance Theory (ART)
Basic ART Architecture

Layer 1

Orienting Subsystem

Layer 2

Gain Control

Expectation

Input

Reset
ART Subsystems

Layer 1
  Normalization
  Comparison of input pattern and expectation

L1-L2 Connections (Instars)
  Perform clustering operation.
  Each row of $W^{1:2}$ is a prototype pattern.

Layer 2
  Competition, contrast enhancement

L2-L1 Connections (Outstars)
  Expectation
  Perform pattern recall.
  Each column of $W^{2:1}$ is a prototype pattern

Orienting Subsystem
  Causes a reset when expectation does not match input
  Disables current winning neuron
Layer 1

\[ \varepsilon \frac{dn^1}{dt} = -n^1 + (b^1 - n^1) \{ p + W_{2:1} a^2 \} - (n^1 + b^1) [-W] a^2 \]
Layer 1 Operation

Shunting Model

\[ \varepsilon \frac{dn_1(t)}{dt} = -n_1(t) + (b^1 - n_1(t))\{p + W^{2:1}a^2(t)\} - (n_1(t) + b^1)[W^1]a^2(t) \]

- Excitatory Input (Comparison with Expectation)
- Inhibitory Input (Gain Control)

\[ a^1 = \text{hardlim}^+(n^1) \]

\[ \text{hardlim}^+(n) = \begin{cases} 
1, & n > 0 \\
0, & n \leq 0 
\end{cases} \]
Excitatory Input to Layer 1

\[ p + W^{2:1} a^2(t) \]

Suppose that neuron \( j \) in Layer 2 has won the competition:

\[
W^{2:1} a^2 = \begin{bmatrix}
  w_1^{2:1} & w_2^{2:1} & \ldots & w_j^{2:1} & \ldots & w_{s^2}^{2:1}
\end{bmatrix} \begin{bmatrix}
  0 \\
  0 \\
  \vdots \\
  0 \\
  1 \\
  \vdots
\end{bmatrix} = w_j^{2:1} \quad (j\text{th column of } W^{2:1})
\]

Therefore the excitatory input is the sum of the input pattern and the L2-L1 expectation:

\[ p + W^{2:1} a^2 = p + w_j^{2:1} \]
Inhibitory Input to Layer 1

Gain Control

\[ [-W^1]a^2(t) \]

\[-W^1 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \]

The gain control will be one when Layer 2 is active (one neuron has won the competition), and zero when Layer 2 is inactive (all neurons having zero output).
Steady State Analysis: Case I

\[ \frac{\varepsilon}{dt} \frac{dn_i}{dt} = -n_i + (b^1 - n_i) \left\{ p_i + \sum_{j=1}^{S^2} w_{i,j} a_j \right\} - (n_i + b^1) \sum_{j=1}^{S^2} a_j \]

Case I: Layer 2 inactive (each \( a_{2,j} = 0 \))

\[ \frac{\varepsilon}{dt} \frac{dn_i}{dt} = -n_i + (b^1 - n_i) \{ p_i \} \]

In steady state:

\[ 0 = -n_i + (b^1 - n_i) p_i = -(1 + p_i)n_i + b^1 p_i \quad \iff \quad n_i = \frac{+b^1 p_i}{1 + p_i} \]

Therefore, if Layer 2 is inactive:

\( a^1 = p \)
Steady State Analysis: Case II

Case II: Layer 2 active (one \( a^2_j = 1 \))

\[
\varepsilon \frac{dn_i^1}{dt} = -n_i^1 + (b^1_i - n_i^1)\{p_i + w_{i,j}^{2:1}\} - (n_i^1 + b^1_i)
\]

In steady state:

\[
0 = -n_i^1 + (b^1_i - n_i^1)\{p_i + w_{i,j}^{2:1}\} - (n_i^1 + b^1_i)
= - (1 + p_i + w_{i,j}^{2:1})n_i^1 + (b^1_i (p_i + w_{i,j}^{2:1}) - b^1_i)
\]

\[
n_i^1 = \frac{+b^1_i (p_i + w_{i,j}^{2:1}) - b^1_i}{2 + p_i + w_{i,j}^{2:1}}
\]

We want Layer 1 to combine the input vector with the expectation from Layer 2, using a logical AND operation:

\[
n^1_i < 0, \text{ if either } w_{i,j}^{2:1} \text{ or } p_i \text{ is equal to zero.} \quad +b^1_i (2) - b^1_i > 0
n^1_i > 0, \text{ if both } w_{i,j}^{2:1} \text{ or } p_i \text{ are equal to one.} \quad +b^1_i - b^1_i < 0
\]

Therefore, if Layer 2 is active, and the biases satisfy these conditions:

\[
a^1 = p \cap w_j^{2:1}
\]
Layer 1 Summary

If Layer 2 is inactive (each $a^{2}_j = 0$)

$$a^1 = p$$

If Layer 2 is active (one $a^{2}_j = 1$)

$$a^1 = p \cap w^2_{j}$$
Layer 1 Example

\[ \varepsilon = 1, +b^1 = 1 \text{ and } -b^1 = 1.5 \]

\[ W^{2:1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad p = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \]

Assume that Layer 2 is active, and neuron 2 won the competition.

\[
(0.1) \frac{dn^1_1}{dt} = -n^1_1 + (1 - n^1_1)\{p_1 + w^{2:1}_{1,2}\} - (n^1_1 + 1.5) \\
= -n^1_1 + (1 - n^1_1)\{0 + 1\} - (n^1_1 + 1.5) = -3n^1_1 - 0.5
\]

\[
\frac{dn^1_1}{dt} = -30n^1_1 - 5
\]

\[
(0.1) \frac{dn^2_2}{dt} = -n^1_2 + (1 - n^1_2)\{p_2 + w^{2:1}_{2,2}\} - (n^1_2 + 1.5) \\
= -n^1_2 + (1 - n^1_2)\{1 + 1\} - (n^1_2 + 1.5) = -4n^1_2 + 0.5
\]

\[
\frac{dn^1_2}{dt} = -40n^1_2 + 5
\]
Example Response

\[ n_1(t) = -\frac{1}{6} [1 - e^{-30t}] \]

\[ n_2(t) = \frac{1}{8} [1 - e^{-40t}] \]

\[ p \cap w_2^{2:1} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cap \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = a^1 \]
\[
\varepsilon \frac{dn^2}{dt} = -n^2 + (b^2 - n^2) \left\{ [W^2] f^2(n^2) + W^{1:2} a^1 \right\} - (n^2 + b^2) [-W^2] f^2(n^2)
\]
Layer 2 Operation

Shunting Model

\[ \epsilon \frac{dn^2(t)}{dt} = -n^2(t) \]

On-Center Feedback

Adaptive Instars

\[ + (\hat{b}^2 - n^2(t)) \left\{ [\hat{\boldsymbol{W}}^2] \hat{f}^2(n^2(t)) + \hat{\boldsymbol{W}}^{1:2} \hat{a}^1 \right\} \]

Excitatory Input

Off-Surround Feedback

Inhibitory Input

\[ - (n^2(t) + \hat{b}^2) [\hat{\boldsymbol{W}}^2] \hat{f}^2(n^2(t)) \]
16

Layer 2 Example

\[ \varepsilon = 0.1 \quad +b^2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad -b^2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad W^{1:2} = \begin{bmatrix} (1^{1:2})^T \\ (2^{1:2})^T \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 1 & 0 \end{bmatrix} \]

\[ f^2(n) = \begin{cases} 10(n)^2, & n \geq 0 \\ 0, & n < 0 \end{cases} \quad \text{Faster than linear, winner-take-all} \]

\[ (0.1) \frac{dn_1^2(t)}{dt} = -n_1^2(t) + (1 - n_1^2(t)) \left\{ f^2(n_2^2(t)) + (1^{1:2})^T \mathbf{a} \right\} - (n_1^2(t) + 1) f^2(n_2^2(t)) \]

\[ (0.1) \frac{dn_2^2(t)}{dt} = -n_2^2(t) + (1 - n_2^2(t)) \left\{ f^2(n_2^2(t)) + (2^{1:2})^T \mathbf{a} \right\} - (n_2^2(t) + 1) f^2(n_1^2(t)) . \]
Example Response

\[ a^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \]

\[ a^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \]
Layer 2 Summary

\[
a^2_i = \begin{cases} 
1, & \text{if}((iw^{1:2})^T a^1 = max[(jw^{1:2})^T a^1]) \\
0, & \text{otherwise}
\end{cases}
\]
Purpose: Determine if there is a sufficient match between the L2-L1 expectation ($a^1$) and the input pattern ($p$).
When the excitatory input is larger than the inhibitory input, the Orienting Subsystem will be driven on.
Steady State Operation

\[ 0 = -n^0 + (+b^0 - n^0) \left\{ \alpha \|p\|^2 \right\} - (n^0 + -b^0) \left\{ \beta \|a^1\|^2 \right\} \]

\[ = - (1 + \alpha \|p\|^2 + \beta \|a^1\|^2) n^0 + (+b^0 (\alpha \|p\|^2) - (-b^0 (\beta \|a^1\|^2) \]

\[ n^0 = \frac{+b^0 (\alpha \|p\|^2) - (-b^0 (\beta \|a^1\|^2)}{(1 + \alpha \|p\|^2 + \beta \|a^1\|^2)} \]

Let \( +b^0 = -b^0 = 1 \)

\[ n^0 > 0 \text{ if } \frac{\|a^1\|^2}{\|p\|^2} < \frac{\alpha}{\beta} = \rho \]

\[ \text{Vigilance} \]

Since \( a^1 = p \cap w_{j}^{2:1} \), a reset will occur when there is enough of a mismatch between \( p \) and \( w_{j}^{2:1} \).
Orienting Subsystem Example

\( \varepsilon = 0.1, \alpha = 3, \beta = 4 (\rho = 0.75) \quad \mathbf{p} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \mathbf{a}^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \)

\[
(0.1) \frac{dn^0(t)}{dt} = -n^0(t) + (1 - n^0(t))\{3(p_1 + p_2)\} - (n^0(t) + 1)\{4(a_1^1 + a_2^1)\}
\]

\[
\frac{dn^0(t)}{dt} = -110n^0(t) + 20
\]
Orienting Subsystem Summary

\[ a^0 = \begin{cases} 
1, & \text{if } \frac{\|a^1\|^2}{\|p\|^2} < \rho \\
0, & \text{otherwise}
\end{cases} \]
The ART1 network has two separate learning laws: one for the L1-L2 connections (instars) and one for the L2-L1 connections (outstars).

Both sets of connections are updated at the same time - when the input and the expectation have an adequate match.

The process of matching, and subsequent adaptation is referred to as resonance.
Suppose that $W^{1:2} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ so the prototypes are $w^{1:2}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $w^{1:2}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

We say that $w^{1:2}_1$ is a subset of $w^{1:2}_2$, because $w^{1:2}_2$ has a 1 wherever $w^{1:2}_1$ has a 1.

If the output of layer 1 is $a^1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, then the input to Layer 2 will be

$$W^{1:2}a^1 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Both prototype vectors have the same inner product with $a^1$, even though the first prototype is identical to $a^1$ and the second prototype is not. This is called the Subset/Superset dilemma.
Subset/Superset Solution

Normalize the prototype patterns.

\[ W^{1:2} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 0 \\ 1 & 1 & 1 \\ 3 & 3 & 3 \end{bmatrix} \]

\[ W^{1:2} a^1 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 0 \\ 1 & 1 & 1 \\ 3 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \]

Now we have the desired result; the first prototype has the largest inner product with the input.
L1-L2 Learning Law

Instar Learning with Competition

\[
\frac{d[\mathbf{w}^{1:2}(t)]}{dt} = a_i^2(t)\left[\{^+\mathbf{b} - \mathbf{w}^{1:2}(t)\} \zeta[^+\mathbf{W}]\mathbf{a}^1(t) - \{\mathbf{w}^{1:2} + ^-\mathbf{b}\}[-\mathbf{W}]\mathbf{a}^1(t)\right],
\]

where

\[
^+\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad ^-\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad ^+\mathbf{W} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \quad ^-\mathbf{W} = \begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 0 \end{bmatrix}
\]

Upper Limit Bias, Lower Limit Bias, On-Center Connections, Off-Surround Connections

When neuron \(i\) of Layer 2 is active, \(\mathbf{w}^{1:2}\) is moved in the direction of \(\mathbf{a}^1\). The elements of \(\mathbf{w}^{1:2}\) compete, and therefore \(\mathbf{w}^{1:2}\) is normalized.
Fast Learning

\[
\frac{dw_{i,j}^{1:2}(t)}{dt} = a_i^2(t) \left[ (1 - w_{i,j}^{1:2}(t)) \zeta a_j^1(t) - w_{i,j}^{1:2}(t) \sum_{k \neq j} a_k^1(t) \right]
\]

For fast learning we assume that the outputs of Layer 1 and Layer 2 remain constant until the weights reach steady state.

Assume that \(a^2_j(t) = 1\), and solve for the steady state weight:

\[
0 = \left[ (1 - w_{i,j}^{1:2}) \zeta a_j^1 - w_{i,j}^{1:2} \sum_{k \neq j} a_k^1 \right]
\]

Case I: \(a_j^1 = 1\)

\[
0 = (1 - w_{i,j}^{1:2}) \zeta - w_{i,j}^{1:2} (\|a\|^2 - 1) = - (\zeta + \|a\|^2 - 1) w_{i,j}^{1:2} + \xi \]

\[
w_{i,j}^{1:2} = \frac{\zeta}{\zeta + \|a\|^2 - 1}
\]

Case II: \(a_j^1 = 0\)

\[
0 = -w_{i,j}^{1:2} \|a\|^2 \]

\[
w_{i,j}^{1:2} = 0
\]

Summary

\[
w_{i,j}^{1:2} = \frac{\zeta a_j^1}{\zeta + \|a\|^2 - 1}
\]
Learning Law: L2-L1

Outstar

\[
\frac{d[w_{j}^{2:1}(t)]}{dt} = a_{j}(t)[-w_{j}^{2:1}(t) + a_{1}(t)]
\]

Fast Learning

Assume that \(a_{j}^{2}(t) = 1\), and solve for the steady state weight:

\[
0 = -w_{j}^{2:1} + a_{1} \quad \text{or} \quad w_{j}^{2:1} = a_{1}
\]

Column \(j\) of \(W^{2:1}\) converges to the output of Layer 1, which is a combination of the input pattern and the previous prototype pattern. The prototype pattern is modified to incorporate the current input pattern.
ART1 Algorithm Summary

0) All elements of the initial $W^{2:1}$ matrix are set to 1. All elements of the initial $W^{1:2}$ matrix are set to $\zeta/(\zeta+S^{1}-1)$.

1) Input pattern is presented. Since Layer 2 is not active,

$$a^1 = p$$

2) The input to Layer 2 is computed, and the neuron with the largest input is activated.

$$a_i^2 = \begin{cases} 1, & \text{if}((i \cdot w^{1:2})^T a^1 = max[(j \cdot w^{1:2})^T a^1]) \\ 0, & \text{otherwise} \end{cases}$$

In case of a tie, the neuron with the smallest index is the winner.

3) The L2-L1 expectation is computed.

$$W^{2:1} a^2 = w^{2:1}_j$$
4) Layer 1 output is adjusted to include the L2-L1 expectation.
\[ a^1 = p \cap w_{j}^{2:1} \]

5) The orienting subsystem determines match between the expectation and the input pattern.
\[ a^0 = \begin{cases} 
1, & \text{if} \|a^1\|^2 / \|p\|^2 < \rho \\
0, & \text{otherwise} 
\end{cases} \]

6) If \( a^0 = 1 \), then set \( a^2_j = 0 \), inhibit it until resonance, and return to Step 1. If \( a^0 = 0 \), then continue with Step 7.

7) Resonance has occurred. Update row \( j \) of \( W^{1:2} \).
\[ jw^{1:2} = \frac{\zeta a^1}{\zeta + \|a^1\|^2 - 1} \]

8) Update column \( j \) of \( W^{2:1} \).
\[ w_{j}^{2:1} = a^1 \]

9) Remove input, restore inhibited neurons, and return to Step 1.