Adaptive Resonance Theory

ART Networks

Stephen Grossberg &
Gail Carpenter

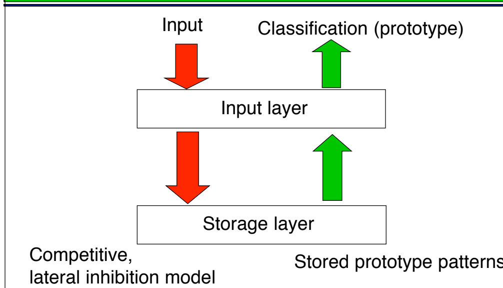
ART Networks

- Work by clustering + nuances
- Several varieties:
  - ART1: Discrete patterns
  - ART2: Continuous patterns …
  - Fuzzy ARTMAP
- Attempt to address the "stability/plasticity dilemma":
  - stability: recognized patterns should be insensitive to noise
  - plasticity: system should be capable of learning new patterns

ART Networks

- Combine supervised and unsupervised (competitive, clustering)
- Dynamically create new categories
- Biologically motivated by an ODE model
- Models short- and long-term memory

ART Layers



ART1

- Two kinds of explanations:
  - algorithmic
  - neural
- We concentrate on the first.
- The second is more complicated, since it involves ways to achieve the control aspects of the algorithmic approach.

Basic ART Operation

- Input pattern presented to input layer
- Storage layer indicates tentative hypothetical classification
- Input layer decides if hypothetical is **close enough**; if so, done.
- If not, storage layer indicates alternate hypothesis.
- The above two steps are repeated until the hypothetical classification is accepted.

ART Operation

- All hypotheses could be rejected; in this case, a **new** class is created in the storage layer.
- “resonance” = mutual reinforcement between input and storage layers
- “adaptive” = weights are adjusted when resonance occurs

ART1 (Discrete Patterns)

- training pattern $x \square \{0, 1\}^n$
- prototype patterns $w_j \square \{0, 1\}^n$
- Storage unit computes
  $y_j = w_j\, x \,/\, \|w_j\|^2$ for each prototype
- The winner is the prototype with the largest y_j
- For acceptance, $y_j > \|x\|^2 / n$, where n is the number of dimensions.
- This means that sufficiently-many bits must **match**.

ART1

- Assuming that the acceptance test is passed, it is also required that

  $$w_j\, x \,/\, \|x\|^2 > \square$$

  where $\square$ is an adjustable parameter called **vigilance**.

  This means that the input and the pattern share a sufficient fraction of 1’s.

ART1

- If the acceptance test is passed, but the vigilance test is not:
  - the prototype in question is temporarily omitted from consideration;
  - a new competition takes place
- until a prototype is found for which both tests are passed.
- If no prototype is found, a new class k is created with

  $$w_k = x$$

ART1

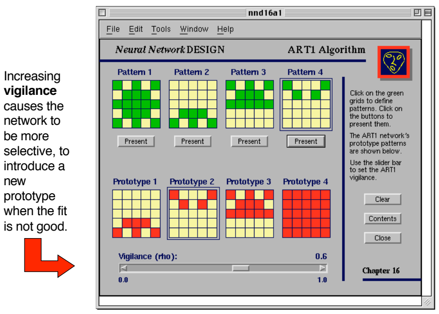
- The **higher** the **vigilance**, the more likely a new pattern is to be introduced.
- Lower vigilance will allow one input to pass as another pattern.

ART1 Issues

- Subset-Superset dilemma:
  - If one pattern is **contained in** another, then a given input may have the same inner product with two different prototypes.
  - Can be resolved by allowing weights other than $\{0,\square\}$ and **normalizing** the prototypes.
  - Neural normalization can be achieved by an on-center, off-surround competition.

ART1 Demo

Increasing **vigilance** causes the network to be more selective, to introduce a new prototype when the fit is not good.



Try different patterns

ART: Neural Version