Artificial Neural Networks and Genetic Algorithms
Instructor: Shu-Heng Chen
Department of Economics
National Chengchi University

Lecture 4: The Significance of Transfer Functions

This lecture is adapted from Sebald and Chellapilla (1998a,b).

Main References:


- A videotape about these two papers given by Sebald and Chellapilla will be displayed in the class.

- This videotape is also available from the Language Center, NCCU.

Overview

- Evolutionary Friendliness (Michael Conrad)

- Problem Representation and Ease of Evolution
Simultaneous evolution of the basis functions and parameters of a neural network using EP

Simulations and Results

Hypothesis

• Each response surface has a natural symmetry
• Each evolution is constrained to certain symmetries
• A good match between the above is important:
  – Simultaneously evolve
    * symmetry of parents/offspring
    * coordinate system for problem

Consider a Simple Example

• Data and Decision
• Input Space $\mathcal{X} \longrightarrow$ Output Space $\mathcal{Y} = \{0, 1\}$
• Symmetries in the decision boundaries
  – hyperplane symmetry
  – ellipsoidal symmetry
  – etc.

• Hyperplane Symmetry and Sigmoid Nodes $(W, slope)$

$$F(S,W.X) = tanh(q(b + \sum W_i X_i))$$  \hspace{1cm} (1)
• Band and Gaussian Nodes \((W, \sigma)\)

\[
F(S, W.X) = b + \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(\sum(W_iX_i)^2}{2\sigma^2}\right) \quad (2)
\]

• Ellipsoid Symmetry and RBF Nodes \((C_{i,f}, \sigma_{i,f}, \alpha_{i,j})\)

\[
F(S, W.X) = b + \frac{1}{\sqrt{2\pi} |\Sigma|} \exp\left(-(X - C)\Sigma^{-1}(X - C)\right) \quad (3)
\]

**Neural Networks for detection**

• Two layered Neural Network

• Basis function expansion

• Sigmoid (Tanh) node

• Gaussian node

• Radial Basis Function: (RBF) node

**Evolving NNs Using Evolutionary Programming**

• Initialization

• Repeat till termination
  \begin{itemize}
  \item Mutation
  \item EP tournament
  \item Selection
  \end{itemize}

• Mutation Operators
  \begin{itemize}
  \item ability to change parameters
  \item ability to change representation
  \end{itemize}
• Self Adaptation

**Mutation Operators**

• Gaussian perturbation of parameters

\[ W_{new} = W_{old} + \eta W N(0, 1) \]  

\[ (4) \]

• Changing the node structure
  - reinitialize node with a “new” activation function and randomly selected parameters
  - local learning using steepest descent for 20 iterations

• Self adaptation of standard deviations

\[ \eta_{new} = \eta_{old} + 0.1 N(0, 1) \]  

\[ (5) \]

**Experiments**

• Four simple detection experiments
  - consisting of hyperplanes and circular symmetries

• EP parameters
  - Population Size: 50
  - Tournament Size: 10
  - Termination criterion: 100 generations

• Fifty independent trials

• Objective Function: Mean Squared Error

**Experiment 1: Circle Results**
• Circular detection boundary

\[(x - 3)^2 + (y - 2)^2 < 1 \quad (6)\]

• 1 hidden node

• RBF nodes dominate the NN

Experiment 2: Band Results

• hyperplane detection boundary

\[|y - 0.5x| < 1 \quad (7)\]

• 1 hidden node

• RBF nodes dominate the NN!!!

Experiment 3: Plane and Circle Results

•

\[(x - 1.5)^2 + y^2 < 1 \quad (8)\]

and

\[x \leq 0 \quad (9)\]

• 2 hidden nodes

• Both sigmoids and RBF nodes dominate the NN

Experiment 4: Ring Results
• Ellipsoid detection boundary

\[ v \leq 2 \]  \hspace{2cm} (10)

or

\[ 3 \leq v \leq 4 \]  \hspace{2cm} (11)

where

\[ v = (x - 4)^2 + 4y^2 \]  \hspace{2cm} (12)

• 3 hidden nodes

• RBF Nodes dominate the NN

Summary

• Evolutionary friendliness has been evolved through problem specific representations

• Problem specific symmetries can be successfully exploited by simultaneously evolving both the representations and parameters, which finally results in better solutions.

Overview

• Detection problems

• Useful coordinate system

• Evolving useful coordinate systems that transforms the input space to simplify the design of detectors

• Simulations and Results

Problem representation
• Detection problems
  – Input Space $D \rightarrow$ Output Space $Y, \{0, 1\}$

• Selecting coordinate systems
  – simplify detector design

Useful Coordinate System

• Coordinate System
  A pair of transformations on the input space ($D$)

  – $G : D \rightarrow Y$ \hspace{1cm} (13)

  – $H : D \rightarrow X$ \hspace{1cm} (14)

• Useful Coordinate System

  – Clusters in $X \times Y$ occupy convex or linearly separable regions
  – Helpful for on-line monitoring of systems by human operators

Parse Tree Representation

•

  $G : D \rightarrow Y$ \hspace{1cm} (15)

  $G(D) = F(T_i - T_0)$ \hspace{1cm} (16)

  where $D = \{T_0, T_i, \}$

•

  $H : D \rightarrow Y$ \hspace{1cm} (17)

  $H(D) = \frac{0.2}{|T_0|}$ \hspace{1cm} (18)

  where $D = \{F_i, T_0, T_i, \}$
Linear Detector

\[ y = mx + c \]  \hspace{1cm} (19)

Designing Coordinate System

- Raw Data (Nonlinearly Separable regions) \((X, Y)\)
- Transformed Data (Linear Separable regions) \((F_1(X, Y), F_2(X, Y))\)
- Evolving transformations \(F_1(X, Y)\) and \(F_2(X, Y)\)
- Design linear detector through gradient descent

Evolving Coordinate System Using Evolutionary Programming

- Initialization
- Repeat till termination
  - Mutation
  - EP tournament
  - Selection
- Mutation Operators
  - ability to change content of the parse tree
  - ability to change the shape of the parse tree

Mutation Operators

- Parse tree mutation operators
  - OneNode, AllNodes, Swap, Grow, Turnc, Promote, Demote, and Gaussian
Gaussian perturbation of constants

\[ X_{new} = X_{old} + \eta_x N(0, 1) \]  

(20)

- Allow a transition from a given point to any other in the search space
- Enable an asymptotic convergence to the global optimum under elitise selection

Experiments

- Five detection experiments
  - consisting of linear and nonlinear detection boundaries
- EP Parameters
  - Pop Size: 500, Tournament Size: 10
  - Termination criterion: 250 generations
  - Parse tree sizes ≤ 30 nodes
- Fitness: Number of correct detections
- Fifty independent trials

Experiment 1: Band Results

- Linear detection boundaries
- Can be solved using two linear detectors
- All runs successful on initialization

Experiment 2: Circle Results
• Circular boundary detection boundary
• All runs successful on initialization
• All runs found perfect solution by Gen. 9

Experiment 3: Ring Results
• Circular detection boundary
• 49 runs successful by 72 Generations
• 40 runs found perfect solution by 136 Gen. 136

Experiment 4: Sin Results
• Detection boundary
  \[ | y | = \sin(4\pi x) \]  
  \[ (21) \]
• 49 runs successful by generation 239
• 38 runs found perfect solution by gen. 239

Experiment 5: 2 Ring Results
• Circular detection boundaries
• 49 runs successful by generation 44
• No perfect solution found on the random data set

Summary
• Evolutionary friendliness has been characterized through useful coordinate representations

• Problem specific coordinate systems can be evolved that simplify the design of detectors

• Future work: Co-evolving the detectors as well as the coordinate systems