Neuro-Fuzzy-Evolutionary Computing (NeF-ECom)

Masoud Nikravesh*

BISC Program, EECS-UCB
&
Life Sciences
Lawrence Berkeley National Laboratory (LBNL)
U.S. Department of Energy (DOE)

http://www-bisc.cs.berkeley.edu/
Email: Nikravesh@cs.berkeley.edu
Tel: (510) 643-4522; Fax: (510) 642-5775

Genetic and Evolutionary Computation Conference
June 25-29
Washington DC
BISC - Decision Support System

Outline

- BISC Decision Support System
- Neuro-Fuzzy-Evolutionary Computing: NeF-ECom
  - Multi-Criteria Decision Analysis with Uncertain and Incomplete Information
- Application Areas
  - ASIS
BISC - Decision Support System
Develop soft-computing-based techniques for decision analysis

- Tools to assist decision-makers in assessing the consequences of decision made in an environment of imprecision, uncertainty, and partial truth and providing a systematic risk analysis;

- Tools to assist decision-makers answer “What if Questions”, examine numerous alternatives very quickly and find the value of the inputs to achieve a desired level of output;

- Tools to be used with human interaction and feedback to achieve a capability to learn and adapt through time;
BISC - Decision Support System

DECISION ENVIRONMENT

- Information (Can be uncertain)
- Granular (Scale and Precision)
- Query (Can be imprecise)
- Measure (Similarity)
- Aggregation (Can be fuzzy)
- Ranking (Provide Alternatives)
- Optimization (Multi-Objective & Multi-Criteria)
BISC - Decision Support System

BISC DSS: Components and Structure

- **Model Representation Including Linguistic Formulation**
  - Functional Requirements
  - Constraints
  - Goals and Objectives
  - Linguistic Variables Requirement

- **Evolutionary Kernel**
  - Genetic Algorithm, Genetic Programming, and DNA
  - Selection
  - Cross Over
  - Mutation

- **Model Management**
  - Query
  - Aggregation
  - Ranking
  - Fitness Evaluation

- **Expert's Knowledge**

- **Input From Decision Makers**

- **Data Management**

- **Model and Data Visualization**

- **Model Representation Including Linguistic Formulation**

- **Model Management**
  - Query
  - Aggregation
  - Ranking
  - Fitness Evaluation

- **Evolutionary Kernel**
  - Genetic Algorithm, Genetic Programming, and DNA
  - Selection
  - Cross Over
  - Mutation

- **Input From Decision Makers**

- **Data Management**

- **Model and Data Visualization**
Query (Request): $Q$

$$Q = f \left( \nu_1 \left\{ (\mu_1, \mu_2, \ldots), w_1 \right\}, \nu_2 \left\{ (\mu_1, \mu_2, \ldots), w_2 \right\}, \ldots \right)$$

$\nu_i$: Variables

$\mu_i$: Degree in which $\nu_i$ belong to a certain grade

$w_i$: Preferences

- find if such query exists $\Rightarrow$ degree of match $\Rightarrow$ rank $\Rightarrow$ decision (i.e. resource allocation)

- compare queries $\Rightarrow$ rank $\Rightarrow$ decision (task allocation)

- Use Fuzzy Min-Max with degree of preferences
Objective function: Cost Function/ Fitness Function

\[ J = \sum_{k} \left[ \sum_{i}^{n} \left( f(v_i \{ \mu_1, \mu_2, \ldots, w_i \}) \cdot \hat{f}(v_i \{ \mu_1, \mu_2, \ldots, w_i \}) \right) \right] \]

This may involve multi-objective, multi-criteria optimization with conflict and fuzzy variables. Therefore, use fuzzy-GA to solve the objective function.
Neuro-Fuzzy-Evolutionary Computing
Multi-Criteria Decision Analysis with Uncertain and Incomplete Information
BISC - Decision Support System

BISC – DSS Software: Architecture

- Aggregation operators
- Similarity measures
  - Norm-Pairs
  - Fuzzy sets

Diagram:
- User Interface
- Application Template
- Fuzzy Search Engine (FSE)
- DB
- Evolutionary Computing Kernel
BISC - Decision Support System

Basic concepts

Fuzzy sets/ Membership Functions (MFs)

- Triangular
- Gaussian
- Trapezoidal

Low diversity  diverse  High diversity
**BISC - Decision Support System**

**Basic concepts**

**Fuzzy similarity measures**

Simple matching: \(|X \cap Y| / |X| + |Y|\)

Dice: \(2 \frac{|X \cap Y|}{|X| + |Y|}\)

Jaccard: \(\frac{|X \cap Y|}{|X \cap Y| + |X \cap Y^c|}\)

Cosine: \(\frac{|X \cap Y|}{\sqrt{|X|} \times \sqrt{|Y|}}\)

Overlap: \(\frac{|X \cap Y|}{\min(|X|, |Y|)}\)

\(X\) and \(Y\) are fuzzy measures defined over the same fuzzy sets with MFs:

\(\mu_1, \mu_2, \ldots, \mu_m\)

Norm-Pair operators \(\cap\) et \(\cup\) (norm-conorm)
## Norm-Pairs

<table>
<thead>
<tr>
<th></th>
<th>Fuzzy AND $[\cap]$</th>
<th>Fuzzy OR $[\cup]$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MinMax</strong></td>
<td>$\min(x, y)$</td>
<td>$\max(x, y)$</td>
</tr>
<tr>
<td><strong>Algebraic</strong></td>
<td>$x \times y$</td>
<td>$x + y - x \times y$</td>
</tr>
<tr>
<td><strong>Bounded</strong></td>
<td>$\max(0, x + y - 1)$</td>
<td>$\min(1, x + y)$</td>
</tr>
<tr>
<td><strong>Drastic</strong></td>
<td>$\min(x, y)$ if $\max(x, y) = 1$, $0$ else</td>
<td>$\max(x, y)$ if $\min(x, y) = 1$, $1$ else</td>
</tr>
<tr>
<td><strong>Einstein</strong></td>
<td>$(x \times y)/(2 - (x + y - x \times y))$</td>
<td>$(x + y)/(1 + (x \times y))$</td>
</tr>
<tr>
<td><strong>Hamacher</strong></td>
<td>$(x \times y)/(x + y - x \times y)$</td>
<td>$(x + y - 2 \times x \times y)/(1 - (x \times y))$</td>
</tr>
</tbody>
</table>

$x$ and $y$ are MF values in $[0,1]$. 
Aggregation Operators

**Arithmetic Mean:** \( \frac{1}{n} \sum_{i=1}^{n} x_i \)

**Geometric Mean:** \( \left( \prod_{i=1}^{n} x_i \right)^{\frac{1}{n}} \)

**Harmonic Mean:** \( \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} \)

**Minimum:** \( \min(x_1, x_2, \ldots, x_n) \)

**Maximum:** \( \max(x_1, x_2, \ldots, x_n) \)
Weighted Aggregation Operators

Weighted Mean: \( \sum_{i=1}^{n} w_i \times x_i \)

Weighted Geometric Mean: \( \prod_{i=1}^{n} x_i^{w_i} \) with \( \sum_{i=1}^{n} w_i = 1 \)

Weighted Harmonic Mean: \( \frac{1}{\sum_{i=1}^{n} w_i \times \frac{1}{x_i}} \)

Weighted Minimum: \( \min\left(\max\left(1 - w_i, x_i\right)\right) \) with \( \max_{i=1}^{n} (w_i) = 1 \)

Weighted Maximum: \( \max\left(\min\left(w_i, x_i\right)\right) \)
Advanced Multi-Aggregator Model

- **Parameters**
  - aggregators
  - weights
  - tree structure.
### BISC - Decision Support System

**BISC-DSS Software**

![Software Interface]

#### Data File Management

<table>
<thead>
<tr>
<th>dataset</th>
<th>attr1</th>
<th>attr2</th>
<th>attr3</th>
<th>attr4</th>
<th>attr5</th>
<th>attr6</th>
<th>attr7</th>
<th>attr8</th>
</tr>
</thead>
<tbody>
<tr>
<td>POS1</td>
<td>309</td>
<td>918</td>
<td>320</td>
<td>1579</td>
<td>137</td>
<td>380</td>
<td>929</td>
<td>902</td>
</tr>
<tr>
<td>POS2</td>
<td>208</td>
<td>643</td>
<td>329</td>
<td>996</td>
<td>138</td>
<td>312</td>
<td>585</td>
<td>832</td>
</tr>
<tr>
<td>POS3</td>
<td>243</td>
<td>782</td>
<td>334</td>
<td>1410</td>
<td>133</td>
<td>350</td>
<td>824</td>
<td>877</td>
</tr>
<tr>
<td>POS4</td>
<td>256</td>
<td>1063</td>
<td>363</td>
<td>2276</td>
<td>442</td>
<td>1039</td>
<td>1473</td>
<td>1359</td>
</tr>
<tr>
<td>POS5</td>
<td>257</td>
<td>1103</td>
<td>303</td>
<td>2276</td>
<td>442</td>
<td>1039</td>
<td>1473</td>
<td>1359</td>
</tr>
<tr>
<td>POS6</td>
<td>222</td>
<td>646</td>
<td>396</td>
<td>1650</td>
<td>220</td>
<td>893</td>
<td>990</td>
<td>1366</td>
</tr>
<tr>
<td>POS7</td>
<td>159</td>
<td>641</td>
<td>434</td>
<td>1270</td>
<td>200</td>
<td>1361</td>
<td>888</td>
<td>2291</td>
</tr>
<tr>
<td>POS8</td>
<td>161</td>
<td>719</td>
<td>407</td>
<td>1466</td>
<td>213</td>
<td>1117</td>
<td>1020</td>
<td>2269</td>
</tr>
<tr>
<td>POS9</td>
<td>244</td>
<td>1019</td>
<td>436</td>
<td>2016</td>
<td>360</td>
<td>1312</td>
<td>1446</td>
<td>2853</td>
</tr>
<tr>
<td>POS10</td>
<td>132</td>
<td>313</td>
<td>262</td>
<td>261</td>
<td>0</td>
<td>-30</td>
<td>160</td>
<td>100</td>
</tr>
<tr>
<td>POS11</td>
<td>155</td>
<td>444</td>
<td>248</td>
<td>403</td>
<td>0</td>
<td>0</td>
<td>220</td>
<td>325</td>
</tr>
<tr>
<td>POS12</td>
<td>155</td>
<td>444</td>
<td>248</td>
<td>403</td>
<td>0</td>
<td>0</td>
<td>220</td>
<td>325</td>
</tr>
<tr>
<td>COS</td>
<td>123</td>
<td>333</td>
<td>222</td>
<td>111</td>
<td>0</td>
<td>0</td>
<td>220</td>
<td>325</td>
</tr>
</tbody>
</table>

- **Advanced Neuro-Fuzzy Model run...**
- **Fuzzy Query-Aggregation Classifier run...**

- **Clear**
- **Save to file >>**
BISC - Decision Support System

EC: Genetic Algorithms

Requirements

- Individual : problem representation
- Fitness function: for evaluation
- Termination criterion

Principle:

- Create randomly an initial population of individuals
- Evolve the population:
  - evaluate and select individuals
  - use them in genetic operators (crossover, mutation)
  - generate new generation
- Stop if termination criterion satisfied
Genetic Operators

Crossover

mutation

parent 1

0 1 0 1 1 1 0 1

parent 2

1 1 0 0 1 1 1 0

child

0 1 0 1 1 1 1 0

parent

0 1 0 1 1 1 0 1

child

0 1 0 0 1 1 0 1
• Individual = Computer program

• Most common representation: tree encoding (nodes = functions, leaves = terminals)

• Fitness function = returned value by the root node
BISC - Decision Support System

EC: Genetic Programming

**Parent 1**

- Chosen node

**Parent 2**

- Chosen node

**Child 1**

**Child 2**

**Crossover**
BISC - Decision Support System

BISC-DSS: Interaction and Optimization

Comparison, Aggregation, Scoring

MODEL based on

- Aggregation operators,
- Similarity measures
- Norm-Pairs
- Fuzzy sets

User Interface

DB

Fuzzy Search Engine (FSE)

QUERY

ANSWERS

User preferences: (re-ranking, selection)

Evolutionary Computing Kernel

OPTIMIZATION
Multi-Attribute Query: $K$ attributes $A_1, A_2, \ldots, A_K$

Database

$\begin{array}{cccc}
  x_{11} & x_{12} & \cdots & x_{1K} \\
  x_{21} & x_{22} & \cdots & x_{2K} \\
  \vdots & \vdots & \ddots & \vdots \\
  x_{N1} & x_{N2} & \cdots & x_{NK} \\
\end{array}$

Scores

$\begin{array}{c}
  S_1 \\
  S_2 \\
  \vdots \\
  S_N \\
\end{array}$

Query

$\begin{array}{cccc}
  y_1 & y_2 & \cdots & y_K \\
\end{array}$

Similarity calculation

Query Answering

- Ranking based (criteria: number top answers)
- Selection based (criteria: threshold)
For each attribute

Query

Data

Fuzzification

Fuzzy similarity calculation

aggregation

Fuzzy similarity measures

Norm-pairs [\cap, \cup]

Fuzzy sets

Aggregation model

Scoring

Ranking or Selecting Answers
Data: \( X_i = (x_{i1}, x_{i2}, \ldots, x_{iK}) \),

Query: \( Q = (y_1, y_2, \ldots, y_k) \)

\( K \) attributes: \( A_1, A_2, \ldots, A_K \)

For each attribute \( A_j \):

- \( r_j \) fuzzy sets \( \mu_1(A_j, \cdot), \mu_2(A_j, \cdot), \ldots, \mu_{r_j}(A_j, \cdot) \)

- \( s_j = \text{similarity}(x_{ij}, y_j) \), \( j = 1, 2, \ldots, K \)

\( \text{Score} = \text{SIM}(Q, X_i) = \text{Aggregation}(s_1, s_2, \ldots, s_k) \)
BISC - Decision Support System

First Order Aggregation Model (1)

- Norm-pair: Min/Max
- Fuzzy similarity measure: Jaccard
- Aggregation operator: Weighted Mean

\[
SIM(Q, X_i) = \sum_{j=1}^{M} w_j \times Jaccard(y_j, x_{ij}), \text{ with } \sum_{j=1}^{M} w_j = 1
\]

\[
Jaccard(y_j, x_{ij}) = \frac{|y_j \cap x_{ij}|}{|y_j \cup x_{ij}|}
\]

\[
y_j \cap x_{ij} = \left[\min(\mu_k(A_j, y_j), \mu_k(A_j, x_{ij}))\right]_{k=1, \ldots, r_j}
\]

\[
y_j \cup x_{ij} = \left[\max(\mu_k(A_j, y_j), \mu_k(A_j, x_{ij}))\right]_{k=1, \ldots, r_j}
\]
Aggregation model = simple weighted aggregation operator

user preferences = attribute weighting
(Degree of importance of each attribute)

Aggregation model parameters = weighting vector

Optimization process: find the optimal weights Using GA.
Model parameters learning using GA

GA-based learning module
- Individuals: weight vectors
- Genetic operators: crossover, Mutation
- Fitness function
- Termination criterion

Optimal weights

Specific fitness function

Problem specification
• **parameters**
  - similarity measures
  - norm-pairs
  - aggregation operators
  - weights
  - aggregation model structure

**Representation of user/expert preferences**
Model description

Parameters
- aggregators
- weights
- tree structure.
Model parameters learning using GP

- Aggregators set, Attributes set, Model constraints

- Specific DNA encoding

- Specific fitness function

- GP-based learning module

- Optimal multi-aggregation model
Fitness function combining:

- accuracy rates to **maximize**
- distance $\Delta$ to **maximize**
- model structure size to **minimize**
<table>
<thead>
<tr>
<th>Application</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finance</td>
<td>• stock prices and characteristics, credit scoring, credit card ranking</td>
</tr>
<tr>
<td>Military</td>
<td>• battlefield simulation and decision making</td>
</tr>
<tr>
<td>Medicine</td>
<td>• diagnosis</td>
</tr>
</tbody>
</table>
| Marketing   | • store and product display  
               • electronic shopping |
| Internet    | • provide knowledge and advice to large numbers of user |
| Education   | • university admission |
| Banking     | • fraud detection |
Automated Sensory Inspection System
What can We Do with Time-Series?

- Clustering
- Classification
  - Normal
  - Ischemia
- Motif Discovery
- Rule Discovery
  - 10
  - $s = 0.5$
  - $c = 0.3$
- Query by Content
- Novelty Detection
BISC - Decision Support System

Chromosome Representation

It is not part of current ASIS-BISC-DSS

Feature Vector, Automated

Feature Vector, Expert Knowledge

$x_1$, $x_i$, $x_n$

$x_{1E}$, $x_{iE}$, $x_{nE}$
Chromosome Representation

- Composed of primitive statistical, fuzzy set, aggregator, similarity, arithmetic, and signal processing operators.
- Each gene (or algorithm) is represented as a tree, accepts both scalar and series input, and outputs scalar features.
- The chromosome produces a feature vector set.

Fuzzy Label, Set Value, Scalar & Series Input

Scalar & Fuzzy Label Features
BISC - Decision Support System

Front/Back-end Architecture

Classifier Generator

- Feature Extraction
- Front-end Stochastic Search
- Classification
- Back-end Classification

- Genetic Algorithm
- Feature Selection Based on ASIS

- Support Vector Machine
- BISC-DSS

- Time Series Classifier
BISC - Decision Support System

Classifier Architecture

Time Series

Classifier Generator

Time Series Classifier

Feature Extractor

Feature Set

Model

Prediction Result

Feature Selection Based On ASIS

BISC-DSS
BISC - Decision Support System

Fitness: In-sample Rate

Steps:
A. Run Feature Extractor
B. Produce Training Set
C. Train SVM/BISC-DSS
D. Produce Model
E. Run Classifier
F. Produce Result Set
G. Calculate Score
BISC - Decision Support System

Fitness: N-Fold Cross Valid

Steps:
A. Run Feature Extractor
B. Produce Training Set
C. Produce Testing Partitions
D. Train on Complement
E. Produce Model
F. Predict Labels of Test Set
G. Score if finished, otherwise, goto step D.
Projection of the cases on the factor-plane (1 x 2)

Cases with sum of cosine square >= 0.00

Factor 2: 12.18%

Factor 1: 71.84%
Projection of the cases on the factor-plane (1 x 2)
Cases with sum of cosine square >= 0.00

Factor 1: 76.84%
Factor 2: 18.26%
BISC - Decision Support System

BISC-DSS Clustering-Based ANSIS

Input
All the Attributes

Unsupervised Clustering
K-Mean, PCA, SOM, ...

Selected Attributes

Final Set of Attributes selected using Unsupervised Clustering

Modeling Techniques
Principal Component Analysis (PCA)
Singular Value decomposition (SVD)
Mahalanobis’ Distance (MD)
One Class Support vector Machine (1CSVM)

Final Set of Attributes selected using Modeling Technique

Marked Data as Clusters, Anomaly, ...

In Progress

Sets of Ranked Attributes;
AS1, AS2, AS3, AS4

AS1
Att1, Att2, Att3, Att4, ...

AS2
Att11, Att12, Att13, Att14

AS3
Att21, Att22, Att23, Att24, ...

AS4
Att31, Att32, Att33, Att34, ...

Cluster, Classification, Anomaly, Rules, Decision Tree, ...

BISC-DSS Software
Projection of the cases on the factor-plane (1 x 2)

Cases with sum of cosine square >= 0.00

Factor 2: 18.26%
Factor 1: 76.84%
Projection of the cases on the factor-plane (1 x 2)

Cases with sum of cosine square >= 0.00

Factor 1: 75.75%
Factor 2: 15.35%
BISC - Decision Support System

Projection of the cases on the factor-plane (1 x 2)

Cases with sum of cosine square >= 0.00

Factor 1: 76.84%
Factor 2: 18.26%
Projection of the cases on the factor-plane (1 x 2)

Cases with sum of cosine square >= 0.00

Factor 1: 75.75%

Factor 2: 15.35\%