Fuzzy Membership Function Design Using Information Theory Measures and Genetic Algorithms

Outlines

- Introduction
- Problem Statement
- Proposed Approach
- Results
- Conclusion

Introduction

Fuzzy systems
- Perfect operation with fuzzy data
- Precise data from measurement and interfaces
- Need to have fuzzy data from precise data
- Conversion from precise to fuzzy (fuzzification)
• Fuzzification
  - A gateway to any fuzzy system applications
introduction . fuzzy membership function

- Normalization

- Different types of fuzzy membership functions
introduction. fuzzy membership function

trapezoidal

Support
Core
Boundary

UD

triangular

Prototype
Boundary
Support

UD

sigmoidal

Support
Core
Boundary

UD

bell-shaped

Prototype
Boundary
Support

UD
• Support or fuzzy partition
  - An essential part of any fuzzy membership function

**Three parts of FMF**
- Support: \( \forall x_i \in X: \mu_{x_i} > 0 \)
- Boundary: \( \forall x_i \in X: 0 < \mu_{x_i} < 1 \)
- Core/Prototype: \( \forall x_i \in X: \mu_{x_i} = 1 \)
**Introduction. Fuzzy Membership Function**

- **FMF design factors**
  - **Support**: the domain in which the FMF is defined - domain of FMF or "a partition of desired information and our interest in which fuzzy information is defined"
  - **Shape**: determining the boundaries and core/prototype and fuzzy behavior of FMF
• FMF design factors
  - Number: number of fuzzy partitions assigned to a Linguistic Variable, influencing the size of fuzzy rule base,

Outlines
• Introduction
• Problem Statement
• Proposed Approach
• Results
• Conclusion

Problem statement
• How can IT measures help in designing FMF? Which parameters can be optimized by IT measures?

• Number
  - Estimating number of fuzzy partitions is a trade off with fuzzy rules, we can not estimate it independently
  - It can be finalized during optimization of fuzzy rules
  - The number of fuzzy rules is the bottleneck, not the number of fuzzy partitions
• Shape
  - Shape of FMF is still a heuristic issue
  - There is no proven relation between information domain and degree of fuzziness in that domain, completely related to intuition, expertise, and expert knowledge
  - Learning from examples can be a solution

• Support
  - we can just estimate informational parameters of FMF, not fuzzy issues
  - Support is a part of our information in which an uncertainty is happening
  - IT measures is suitable for estimating support of FMF, or fuzzy partitions

• Finding an optimum set of fuzzy partitions related to a given linguistic variable
• Optimum fuzzy partitions
• Optimization problem

Outlines
• Introduction
• Problem Statement
• Proposed Approach
• Results
• Conclusion
proposed approach . requirements

• Solution requirements
  – Set of data (simulation or real) for partitioning
  – Fuzzy partitions modeling and Optimization technique
  – FMF design
  – Evaluation procedure

proposed approach . data

• Data
  – Real data preferred
  – U of Toronto-Mississauga Meteorological Station
  – Temperature information for year 2000 and 2001

proposed approach . modeling & optimization

• Optimization
  – To search UD for the best set of support values
  – Genetic Algorithms (GA)

proposed approach . modeling & optimization

• Performance indices
  – Fitness function in GA optimization procedure
    • Shannon entropy
    • Mutual information
• How we relate FMF to information measure?
  - Mapping the FMF on the histogram of given data
  - Probability-statistics
    - PDF-histogram
  - Maximizing the entropy of partitioned histogram based on given number of partitions (n)

• overlaps
  - In a n-fuzzy-partitioned information, allowed overlaps just between two adjacent partitions, we have n-1 overlaps
  - How to model the overlaps between partitions?
• Two strategy:
  - Overlaps as independent partitions: maximize entropy of independent partitions
  - Overlaps as conjunction of two joint partitions: maximize entropy of joint partitions (considering mutual information)

• First: Overlaps as independent partitions – \((2n-1)\) partitions
Algorithm
- Do optimization for given number of partitions
  - Change width of partitions
  - Until maximum $H$

$$H = \sum_{i=1}^{2n-1} H_i$$

- Increased and enhanced overlaps
- A conservative strategy
  - In fuzzy control applications,
    - Longer rise time
    - Less overshoot
    - Smooth convergence

- Second: Overlaps as conjunction of two joint partitions
Algorithm

- Do optimization for given number of partitions
  - Change width of partitions
  - Until maximum $H$

$$H = \sum_{i=1}^{n} H_i - \sum_{i=1}^{n-1} I_{(i,i+1)}$$

- Decreased overlaps
  - In fuzzy control applications
    - Shorter rise time
    - More overshoot

- Ready to design FMFs
  - We have partitions
  - We need values of boundaries to have complete define of FMF
    - A criteria to choose right value for boundary is necessary
- Importance of boundary
  - Defining a range instead of an exact value

In two partitions A and B, if:

\[ X_{A1} < X_{B1} < X_{A3} < X_{B3} \]

Well-defined boundary, \( W_B \):

\[ W_B \leq (X_{A3} - X_{B1}) \]

\[ A \quad x_{A1} \quad x_{A3} \quad B \]

\[ A \quad x_{B1} \quad x_{B3} \]

**Proposed Approach: Design**

- Importance of boundary
  - Defining a range instead of an exact value

\[ W_B = X_{A3} - X_{A2} \]

\[ W_B > X_{A3} - X_{B1} \]
Outlines

- Introduction
- Problem Statement
- Proposed Approach
- Results
- Conclusion

proposed approach . design

\[ W_B = X_{A1} - X_{A2} \]
\[ W_B = X_{A3} - X_{A2} \]

proposed approach . evaluation

- Evaluation procedure
  - Testing membership function in a complete fuzzy system reacting to a process, compare the output with heuristic-defined membership function
  - Applying the algorithm on other set of data and study the behavior of membership function

results

- Conditions
  - Normalized data
  - Five partitions
  - Algorithm test in both two modes
• GA optimization parameters
  – Search space: 35,184,372,088,832
  – Population size: 400
  – Chromosome/string length: 45
  – P_{cross-over}: 0.3
  – P_{mutation}: 0.01
  – Minimum generation: 200

• First data set:
  – Hourly temperature of city of Toronto during year 2000
  • Max: 36.88
  • Min: -20.67
  • Mean: 8.90
  • STD: 10.30

• Temperature vs. time – year 2000
• Normalized temperature
**Results**

- Temperature vs. time - year 2000
- Normalized temperature
- Histogram

**Mode 1: Overlaps as independent partitions**

- Best strings over generations
- Mean of strings during convergence
- Resulted fuzzy memberships
Mode 2: Overlaps as conjunction of two joint partitions

- Best strings over generations
- Resulted fuzzy memberships

• Second data set:
  - Hourly temperature of city of Toronto during year 2001
    - Max: 32.23
    - Min: -23.34
    - Mean: 9.60
    - STD: 10.20
results

Mode 1: Overlaps as independent partitions

- Best strings over generations
Mode 1: Overlaps as independent partitions

- Best strings over generations
- Mean of strings during convergence
- Resulted fuzzy memberships

Mode 2: Overlaps as conjunction of two joint partitions

- Best strings over generations
- Mean of strings during convergence
Mode 2: Overlaps as conjunction of two joint partitions

- Best strings over generations
- Mean of strings during convergence
- Resulted fuzzy memberships

Mode 1: Overlaps as independent partitions

- Data set 2000
- Data set 2001

Outlines

- Introduction
- Problem Statement
- Proposed Approach
- Results
- Conclusion
A solution for designing fuzzy membership function

Besides fuzzy rules generation, a solution for designing fuzzy system by learning from example

The idea: having generic membership function for generic data