Preserving Interpretability in the Optimization of Fuzzy Systems: A Topological Approach

Ricardo DE ALDAMA (Sony CSL)

Michaël AUPETIT (CEA LIST)

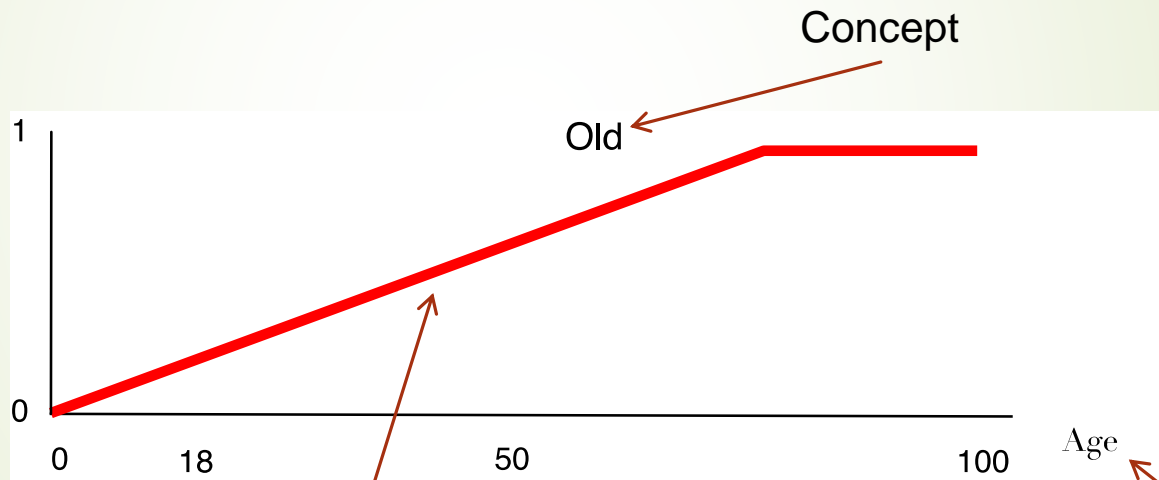DataSense Research day, July 8th 2014

list

INSTITUT CARNOT CEA LIST digiteo

FUZZY SYSTEMS AND OPTIMIZATION

➡ FUZZY SET

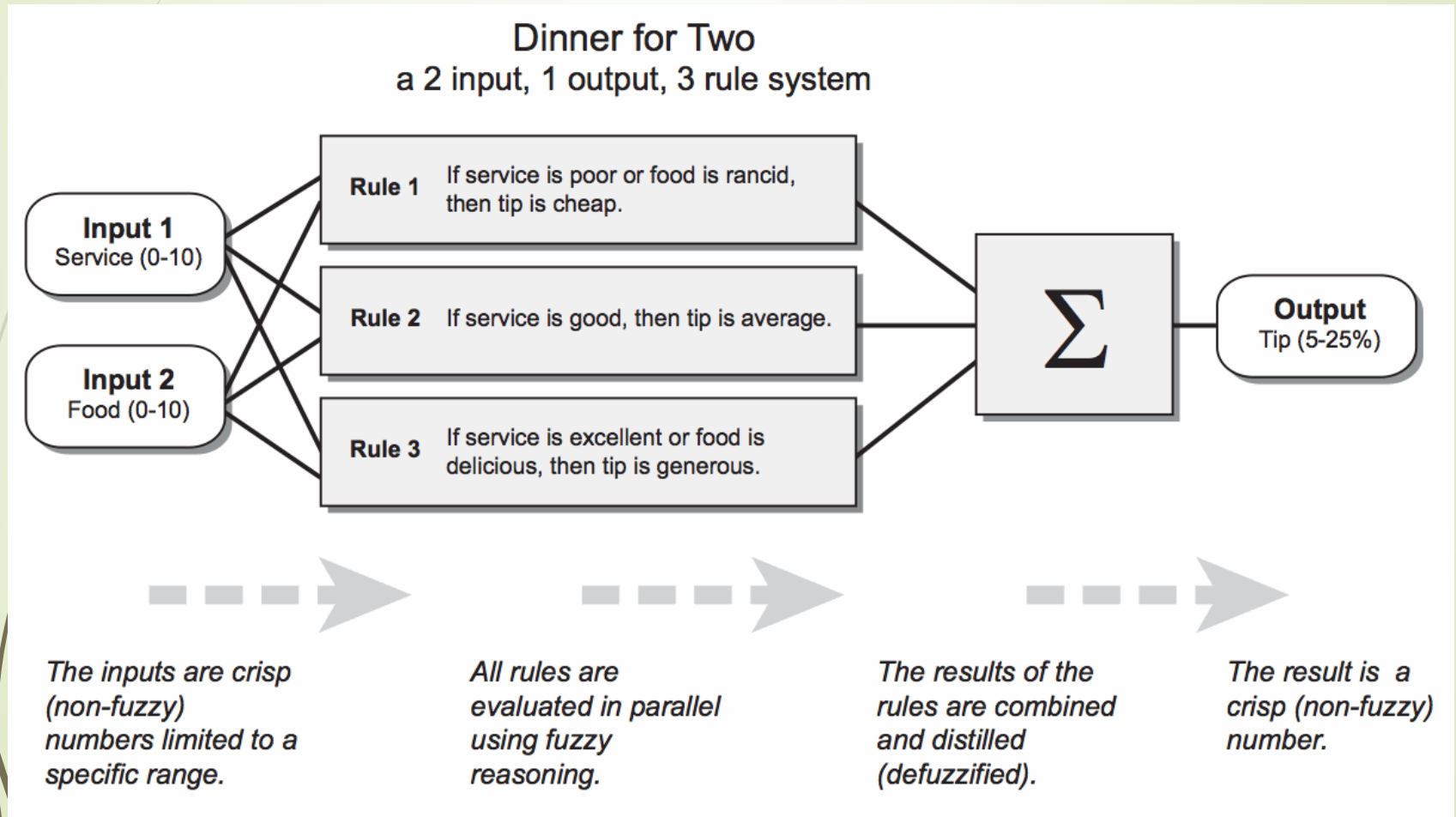• Easy way to integrate expert domain knowledge into the models

Truth value

Concept

Old

Fuzzy membership function

Linguistic variable

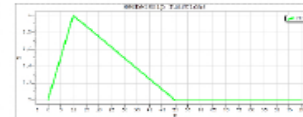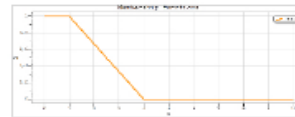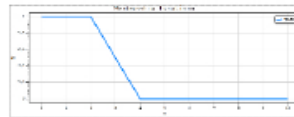# FUZZY SYSTEMS AND OPTIMIZATION

➡ FUZZY RULE-BASED SYSTEM

(Image from MathWorks website)
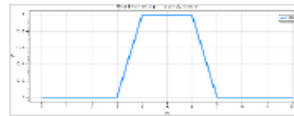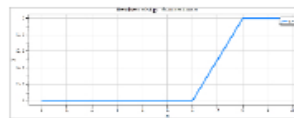
FUZZY SYSTEMS AND OPTIMIZATION

- FUZZY RULE-BASED SYSTEM

Fuzzy rules

FUZZY SYSTEMS AND OPTIMIZATION

- FUZZY RULE-BASED SYSTEM

FUZZY SYSTEMS AND OPTIMIZATION
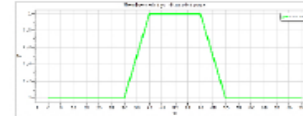
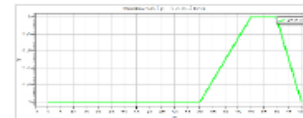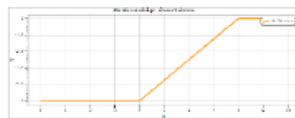- FUZZY RULE-BASED SYSTEM

FUZZY SYSTEMS AND OPTIMIZATION

- FUZZY RULE-BASED SYSTEM
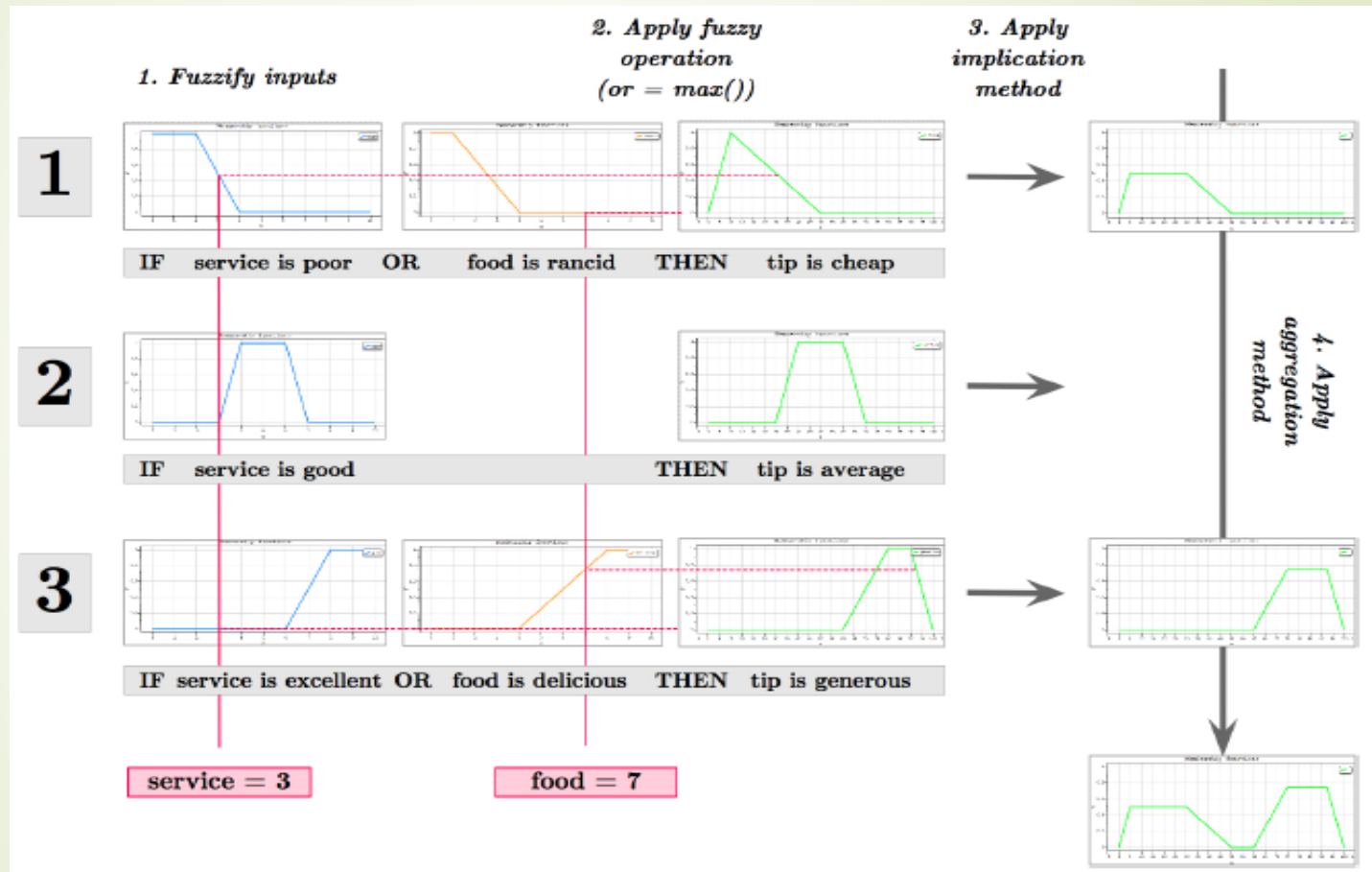
FUZZY SYSTEMS AND OPTIMIZATION

- FUZZY RULE-BASED SYSTEM
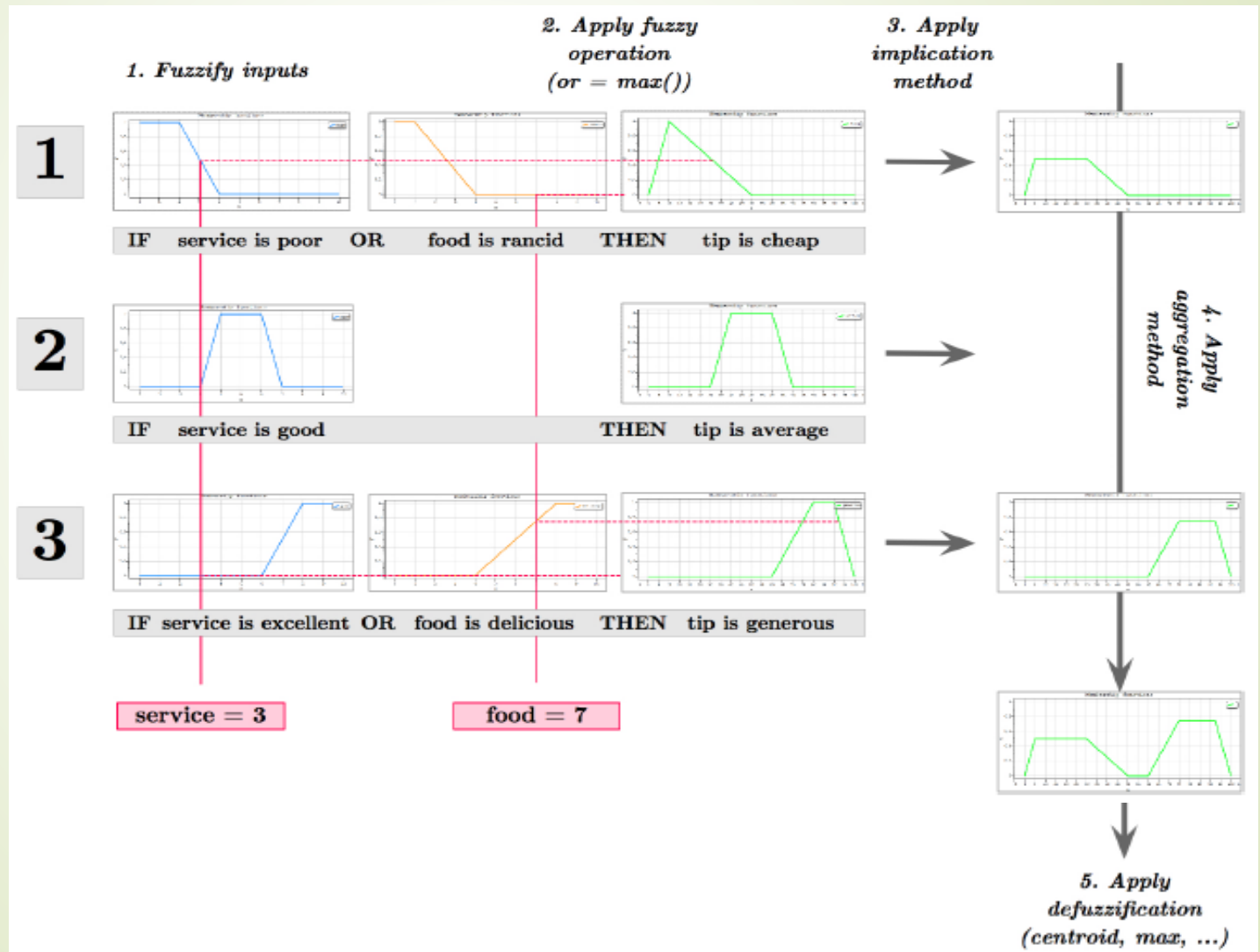
FUZZY SYSTEMS AND OPTIMIZATION

➡ FUZZY RULE-BASED SYSTEM

- Example: fuzzy controller for home energy management
    - ➡ **If** temperature is high, **then** reduce the hot water flow
    - ➡ **If** weather is sunny **and** room temperature is chilly, **then** open the blind
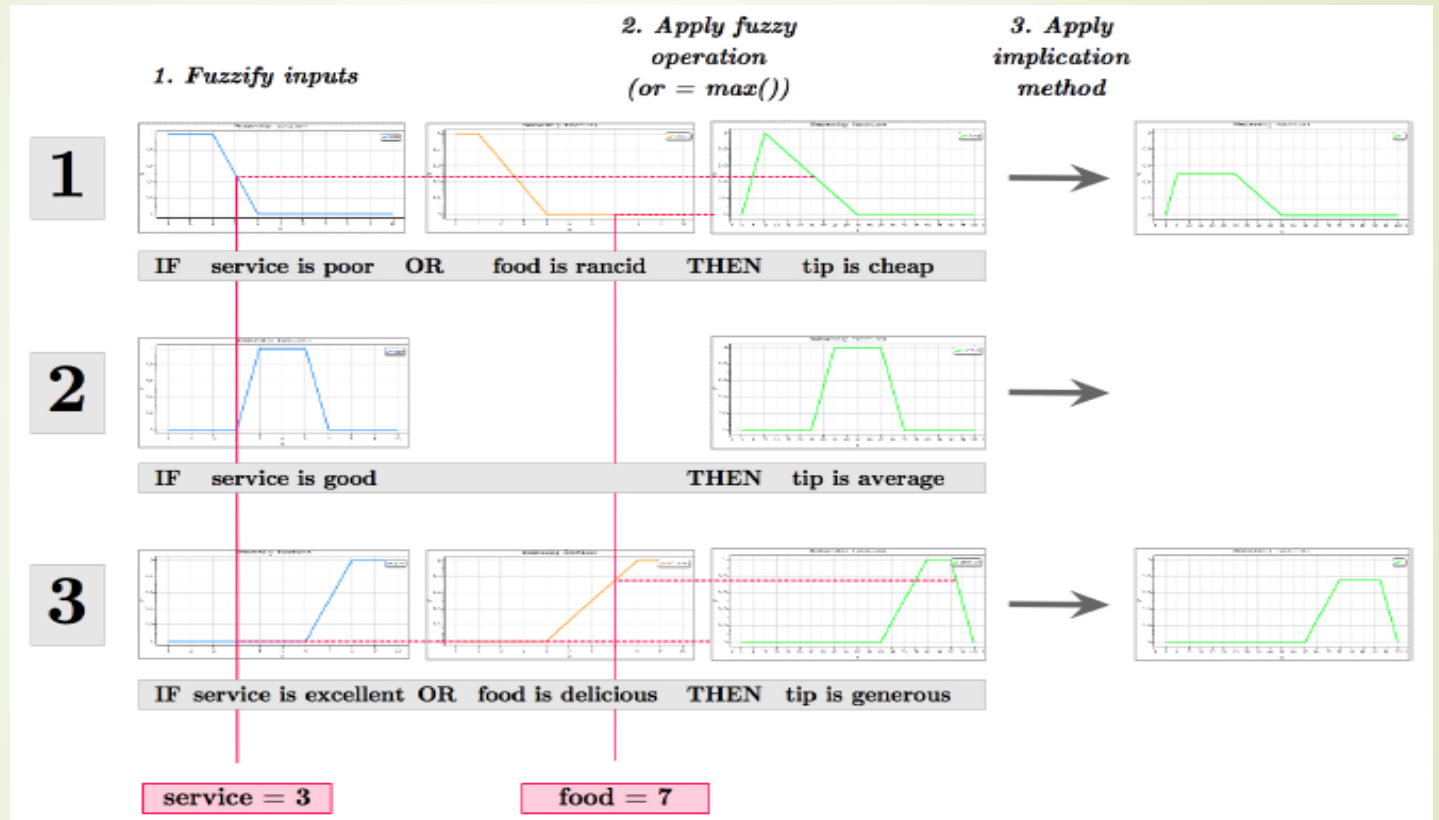   - ➡ **If** weather is sunny **and** room temperature is chilly, **then** close the blind
    - ➡ …

- Problem
    - ➡ Fuzzy sets provided by experts are not necessarily optimal

- Solution
    - ➡ Optimize the parameters of fuzzy sets to minimize energy consumption

FUZZY SYSTEMS AND OPTIMIZATION

- OPTIMIZATION

- Structural optimization (rule learning)
  - Number of rules, linguistic variables, rule complexity, etc.

- Parametric optimization (tuning)
  - Semantics: membership functions and fuzzy partitions

- ▶ OPTIMIZATION

- Problem
  - ▶ **Automatic** modifications of parameters may lead to **loss of interpretability**

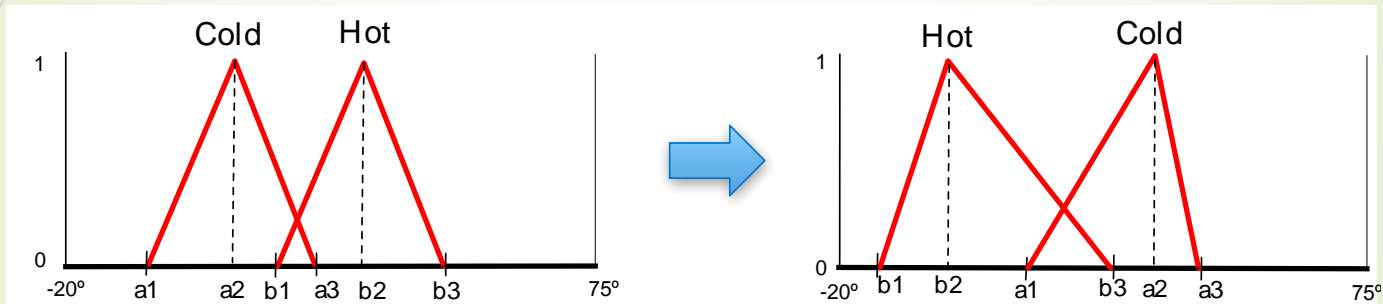- Solution
  - ▶ Force the optimization process to **preserve interpretability**

- How?
  - ▶ No formal definition!

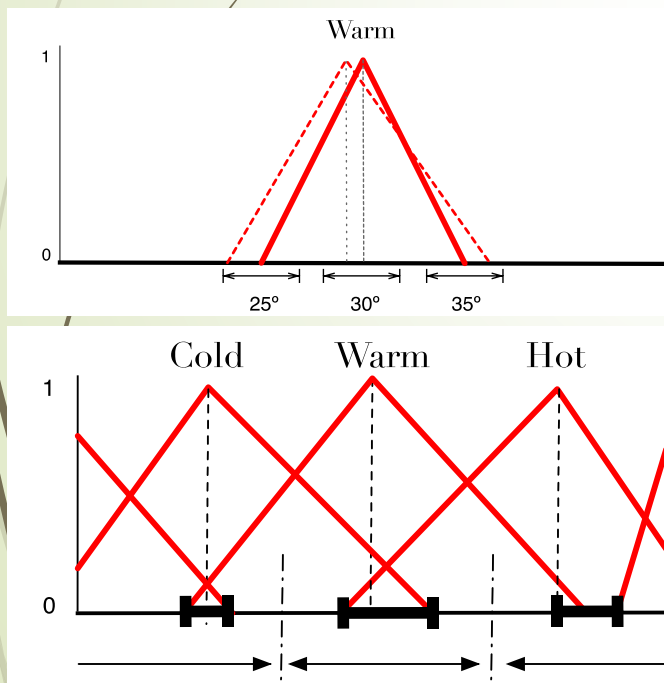# INTERPRETABILITY
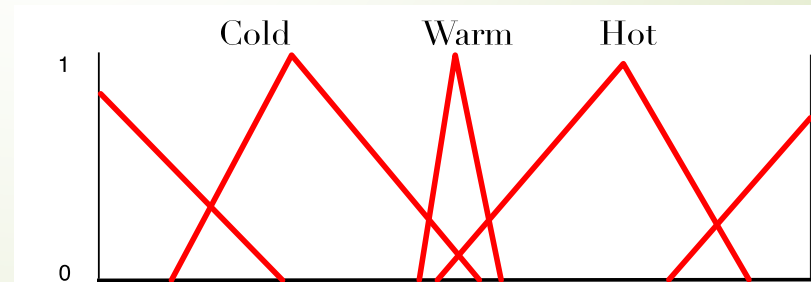
 WE FOCUS ON SEMANTICS (FUZZY PARTITIONS)

- Some common conditions (from the state of the art)
    - A different concept for each membership function
    - Minimal covering: each point is well represented by at least one function
    - Significant intersection between neighboring functions

- Example

This solution seems interpretable, but it violates the above conditions

OUR APPROACH

- FOCUSED ON TOPOLOGY

  - We do not fix a priori the definition of interpretability

  - We suppose the expert/user considers the initial solution(s) as being interpretable

  - We propose a formal framework in which the expert/user can express precisely the information to be preserved during the optimization process

  - The notion of interpretability is then relative to the expert and operationally defined

THE FRAMEWORK

- ☛ GEOMETRIC SIGNATURE (ASSOCIATED TO A LINGUISTIC VARIABLE)

- Regions Ri are maximal connected subspaces in which the order of the membership functions is constant
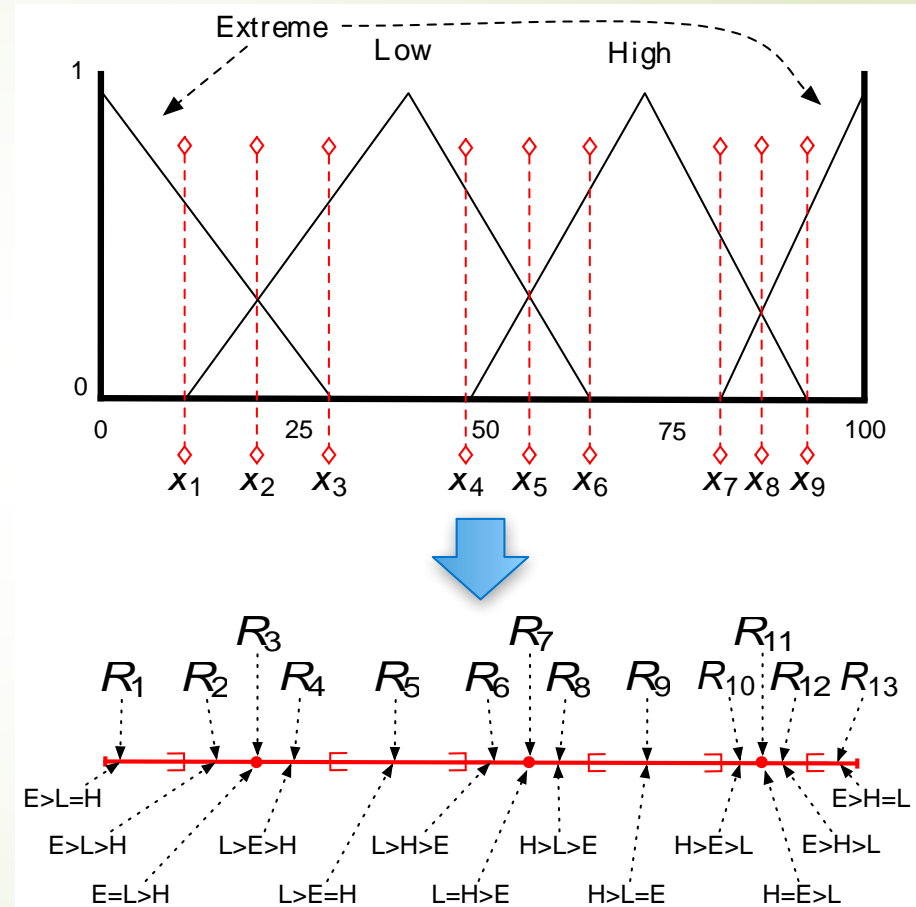
  - ☛ In R1: E>L=H

  - ☛ In R2 : E>L>H

  - ☛ …

- The geometric signature codifies this information

  - ☛ R1= [0, 10] and its order is E>L=H,

  - ☛ R2 = [10, 20] and its order is E>L>H
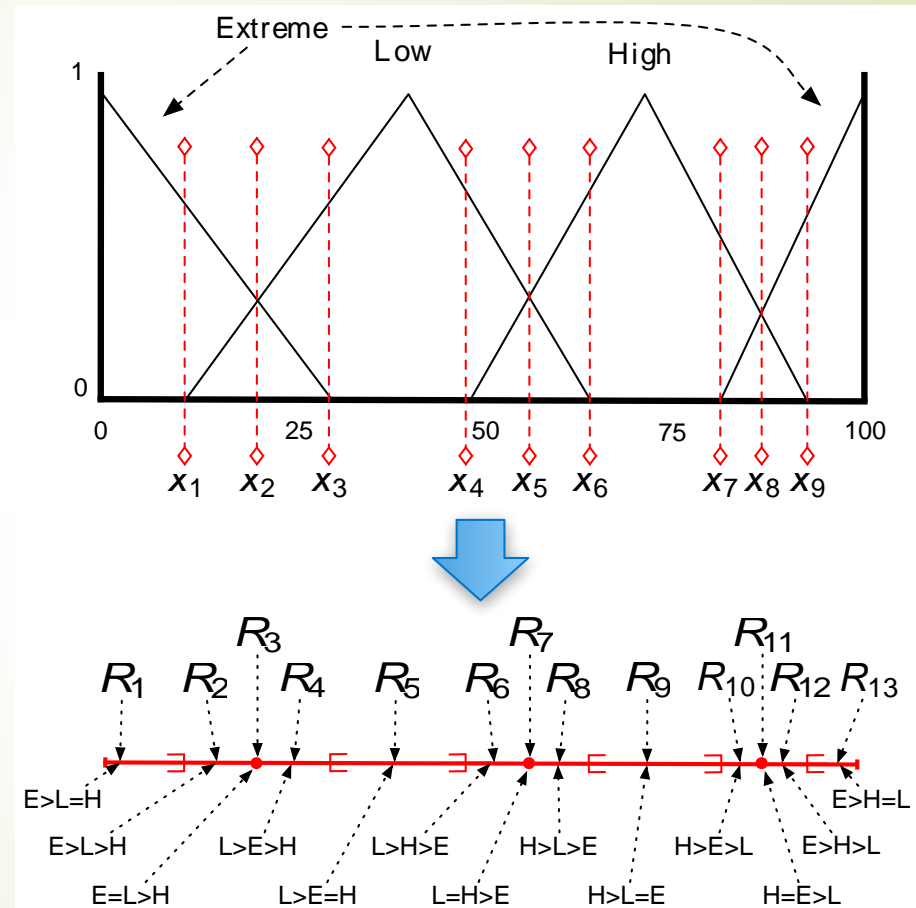
  - ☛ …

TOPOLOGICAL SIGNATURE (ASSOCIATED TO A LINGUISTIC VARIABLE)

- It is determined by the geometric signature

- It codifies its topological content:

  - R1 has exactly one neighbor: R2

  - R5 has exactly two neighbors: R4 and R6

  - …

- Together with the labels:

  - In R1: E>L=H

  - In R2 : E>L>H

  - …

- It does not codify the actual points belonging to the regions

THE FRAMEWORK

- ➡ WHAT TO PRESERVE?

- • First idea

    - ➡ Given an initial solution provided by the user, compute its topological signature and force the new solutions to have the same topological signature

    - ➡ Two main drawbacks:

        - • In most cases this condition does not seem sufficient for interpretability

        - • No room for different notions of interpretability

- • Second idea: operational definition of interpretability

    - ➡ Let the user interact, using geometric and topological signatures, to precisely defined what should be preserved

➡ RELAXATION OF THE TOPOLOGICAL CONDITIONS

- Example: regions could be merged by only taking into account the highest valued function
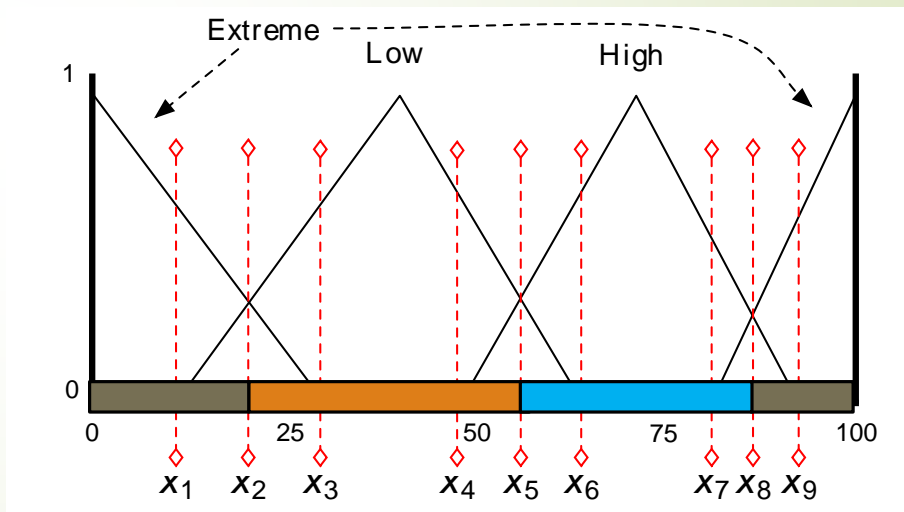
    ➡ RA = R1 U R2

    ➡ RB = R3 U R4 U R5 U R6 U R7

    ➡ RC = R8 U R9 U R10 U R11

    ➡ RD = R12 U R13

- This induces topological conditions **TC**:

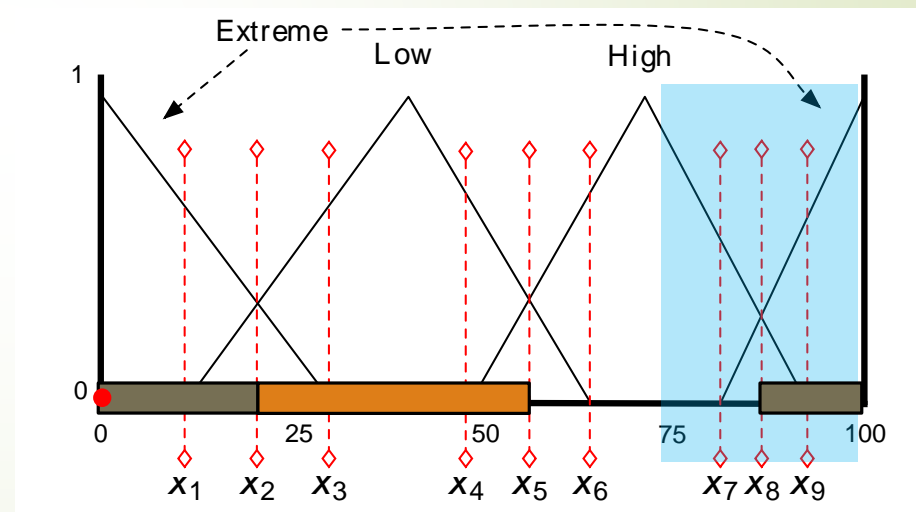   · RA has label "E > …" and exactly one neighbor: RB

    · …

USER INTERACTION

➡ ADDITION OF GEOMETRIC CONDITIONS

- The expert can impose some geometric conditions on RA, RB, RC, RD

    ➡ RA (one of the two regions with E > ..) must contain the point x=0

    ➡ RB (the region with L > ..) must not contain any point greater than 75

    ➡ …

- This induces some geometric conditions **GC** on the signatures of a potential solution

USER INTERACTION

- OPTIMIZATION WITH INTERPRETABILITY CONSTRAINTS: WHAT SHOULD BE SATISFIED BY A POTENTIAL SOLUTION?

- We say that a solution P is interpretable relative to a user U if

  - The topological signature of P satisfies **TC** (topological conditions after user interaction)

  and

  - The geometrical signature of P satisfies **GC** (geometrical conditions after user interaction)

ALGORITHM

OPTIMIZATION WITH INTERPRETABILITY PRESERVATION
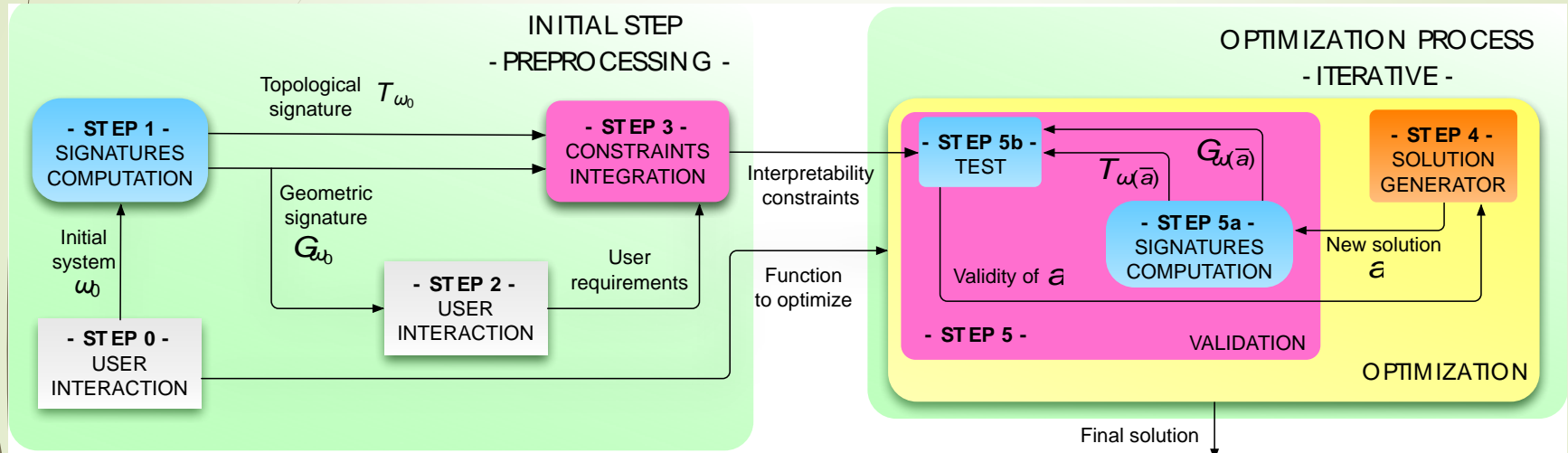
- ## Initialization

  - 0 – Initial FRBS proposed by the user

  - 1 – Characterization of this FRBS (geometric and topological signatures)

  - 2 – Extra user requirements

  - 3 – Integration of constraints

- ## Optimization

  - 4 – Generation of a new FRBS

  - 5a – Compute signatures

  - 5b – Test whether the interpretability constraints are met
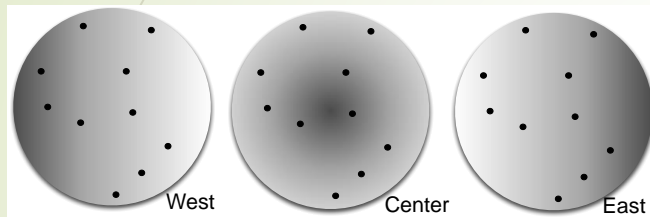
  - 6 – go to 4 or STOP

EXAMPLE

- A 2-D EXAMPLE

- Let's consider a fuzzy controller that produces policy decisions (e.g. public transports, taxes, etc.) for towns in a certain area, following rules of the type "If town T is in region East then apply policy P to T"

  - The expert defines geographical regions by means of fuzzy membership functions
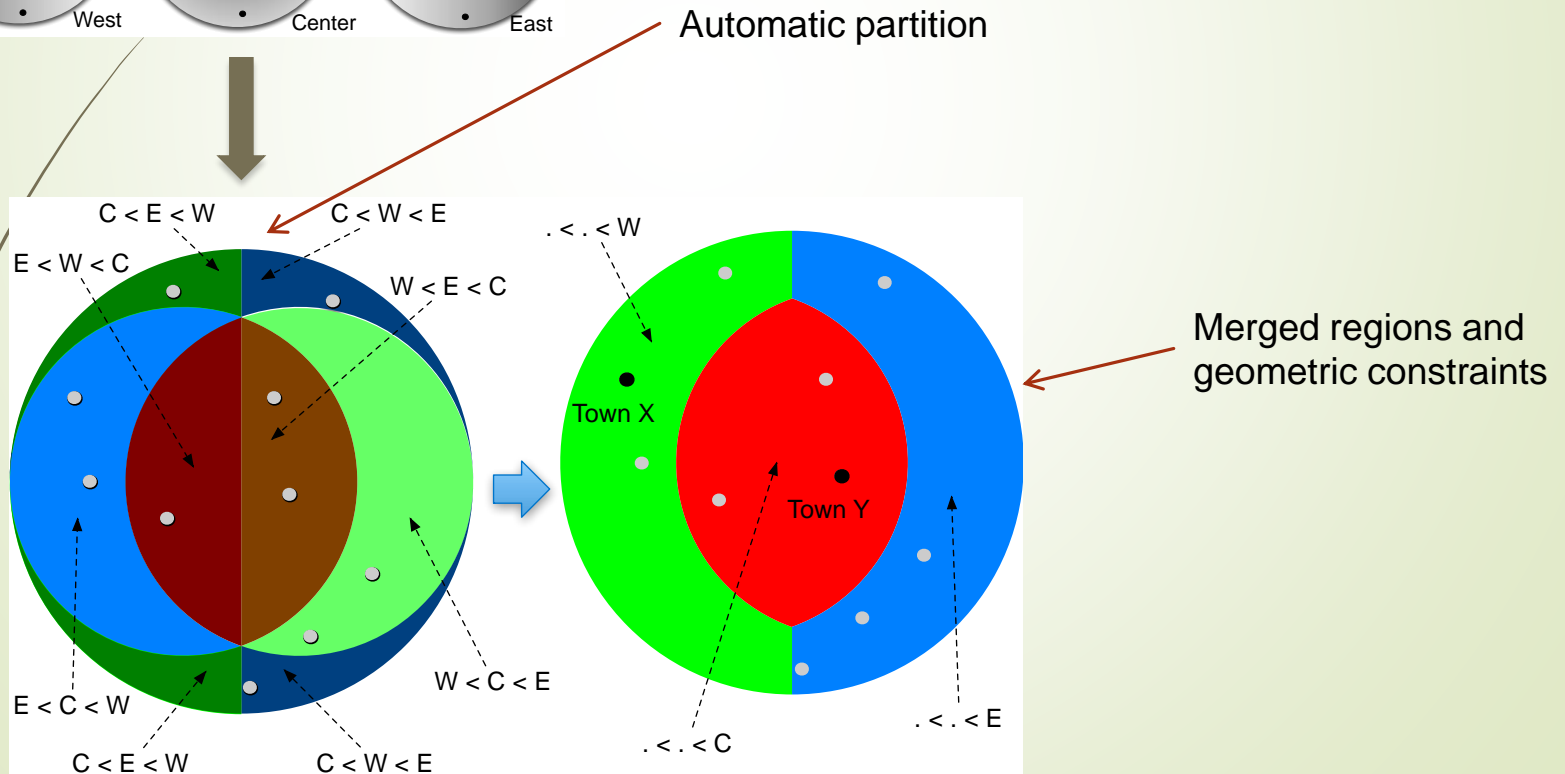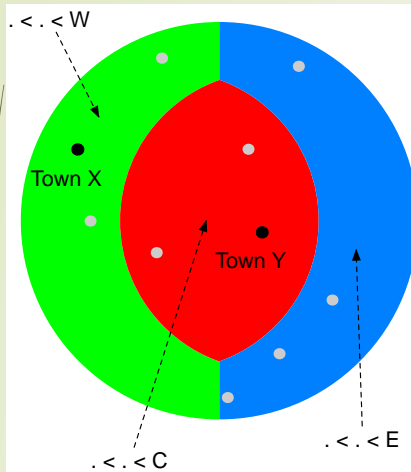

West   Center   East
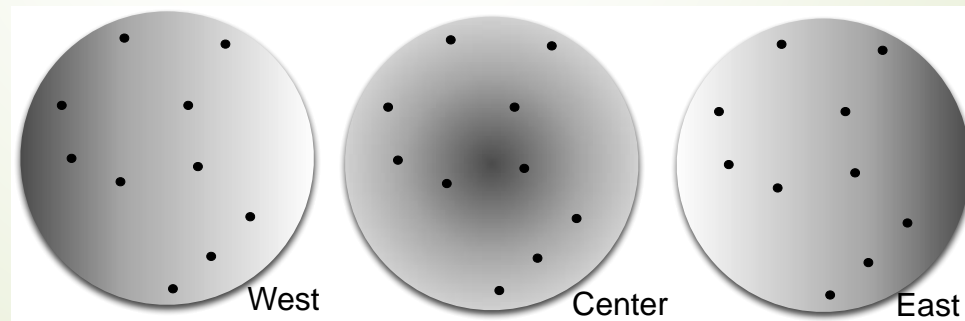
EXAMPLE

USER INTERACTION

- Automatic partition
- Merge of certain regions
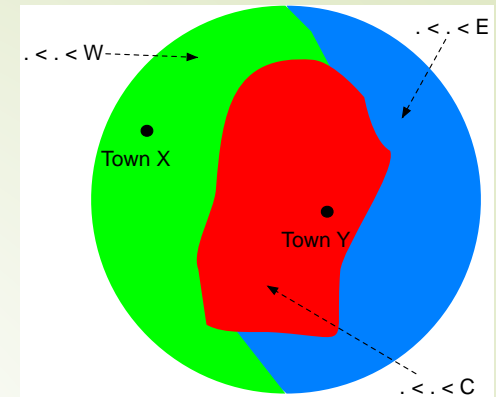- Addition of geometric constraints (X in green region, Y in red region)

Automatic partition
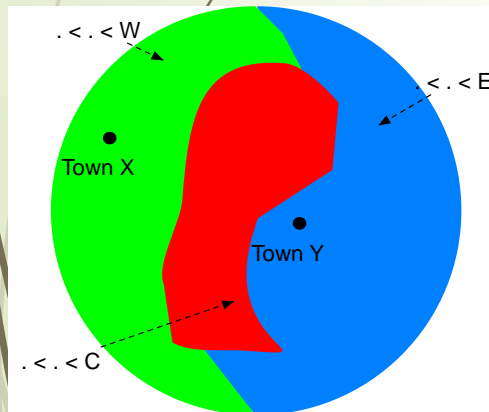
Merged regions and geometric constraints

OPTIMIZATION

Initial solution:
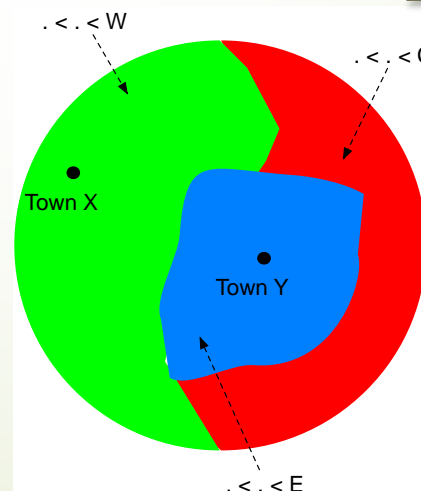**interpretable** but
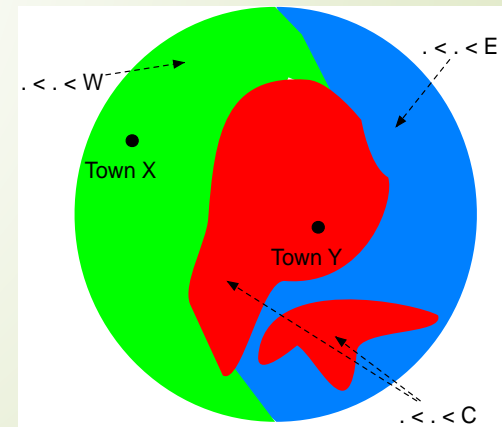**not optimal**

Optimal and **interpretable**

Optimal but **not interpretable**

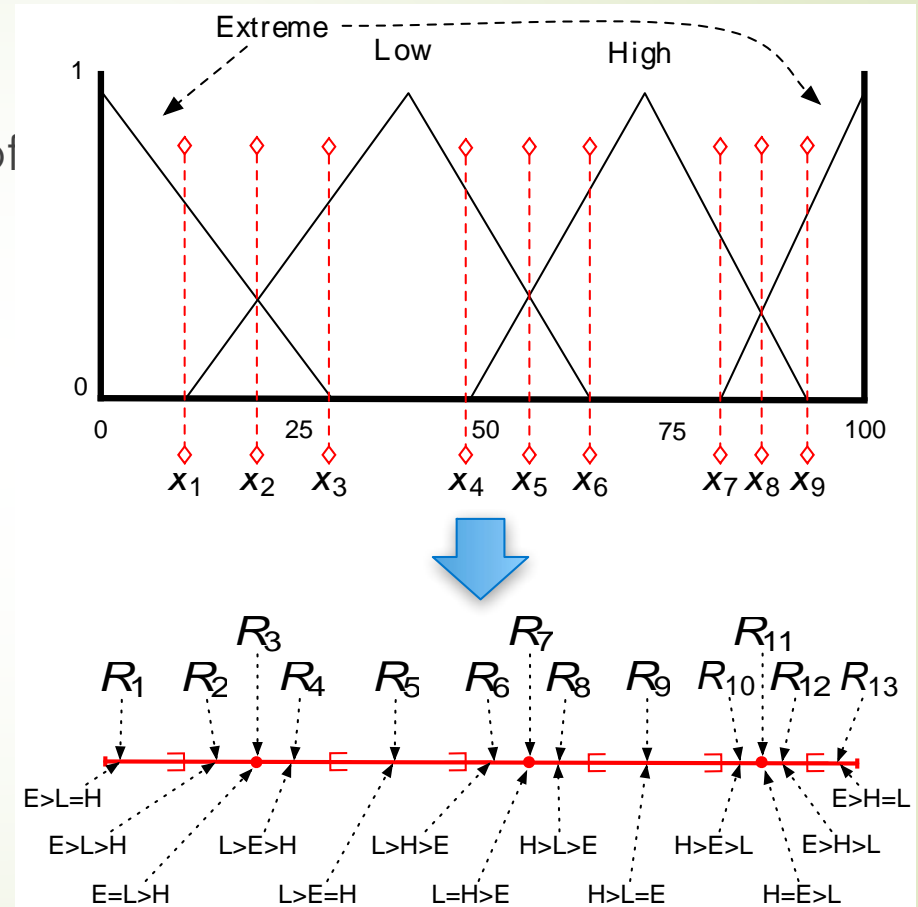Violation of **geometric constraints**

Violation of **topological constraints**

FEASEABILITY

▶ TECHNICAL ASPECTS

- 1D: easy technical solutions

  ▶ Partition: from intersections of membership functions (analytically o numerically computed)

  ▶ Verification of geometric constraints: testing interval membership

  ▶ Verification of topological constraints: comparing the relative order of regions (R1 before R2, R2 between R1 and R3, etc.)

FEASEABILITY

- TECHNICAL ASPECTS

- 2D or more

  - Partition: **adaptive mesh** for detection boundaries of regions

  - Verification of geometric constraints: test whether some points belong to particular **connected components** (of same label)

  - Verification of topological constraints: compare **Betti numbers**
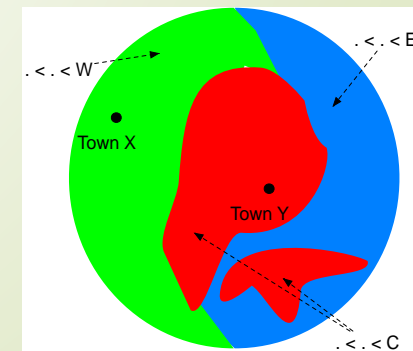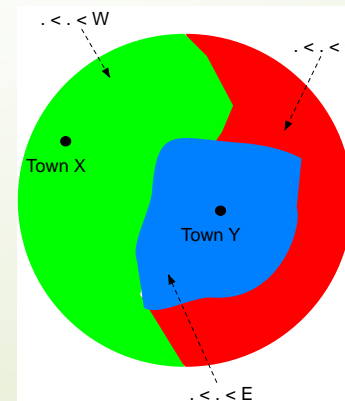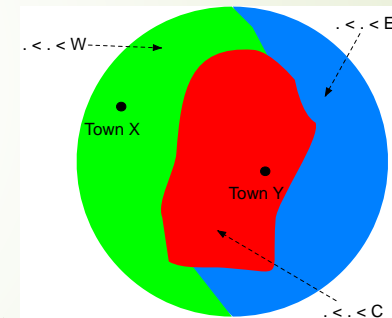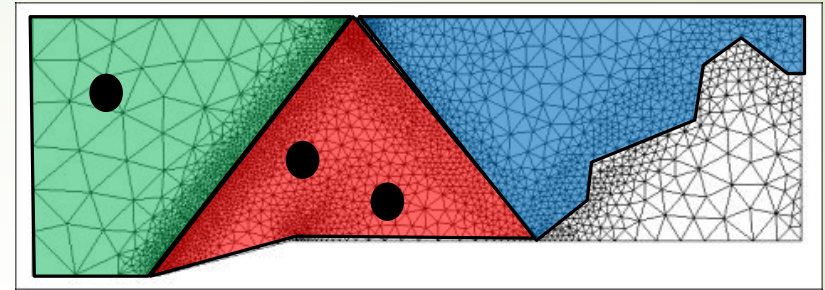
FINAL REMARKS

- CONCLUSION

  - A formal framework for expressing and manipulating interpretability

    - Focused on topological and geometric notions

  - A generic algorithm to preserve it

    - No need for experts in fuzzy systems optimization

- PERSPECTIVES

  - Implement and evaluate in concrete examples

  - Analyze the limits of the framework

  - Improve the computation/codification of multi-labeled topology