ALTANA Chair for
Bioinformatics and Information Mining

Tutorial: Fuzzy Logic

Advanced Course on Knowledge Discovery, Ljubljana, June 2005

Michael R. Berthold
Konstanz University

Michael.Berthold@uni-konstanz.de

Overview

- Fuzzy Logic
  - Motivation
  - Degrees of Membership and Fuzzy Sets
  - Linguistic Values and Variables
  - Operators on Fuzzy Sets
  - Fuzzy Implication
- Fuzzy Rules
  - Inference
  - Construction of Fuzzy Rules from Data
- Fuzzy Arithmetic
  - Fuzzy Numbers
  - Operations on Fuzzy Numbers
  - Extension Principle
- Conclusion

Fuzzy Logic: Motivation

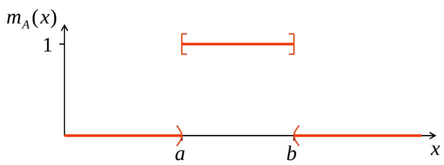
- Modeling of imprecise concepts:
  - Age, Weight, Height, …
- Modeling of imprecise dependencies (e.g. rules):
  - If Temperature is low and Oil is cheap then crank up the heating system
- Origin of Information:
  - Modeling of Expert Knowledge
  - Representation of information extracted from inherently imprecise data

Characteristic Functions: Crisp Sets

- Classical Sets can be described by a characteristic function:

$$m_A(x) := \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases} \qquad m_A(x) \in \{0,1\}$$

- Example: $A = \{x \mid a \le x \le b\}$



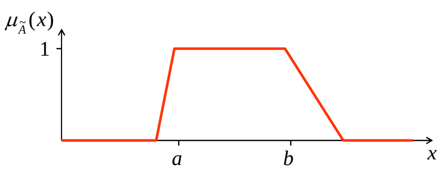
Characteristic Functions: Fuzzy Sets

- Fuzzy Sets are described by a membership function:

$$\mu_{\tilde{A}}(x) \in [0,1]$$

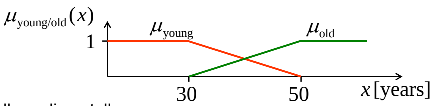
- Example:

$$\tilde{A} = x \text{ is roughly in } [a,b]$$

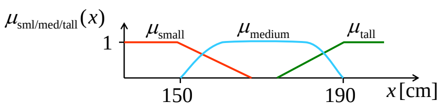


Linguistic Variables and Values

- Associating meaning (semantic) with fuzzy sets results in
  - Linguistic Variables: the (labeled!) domain of the fuzzy sets
  - Linguistic Values: a (labeled!) collection of fuzzy sets on this domain
- Examples:
  - Age: young, old

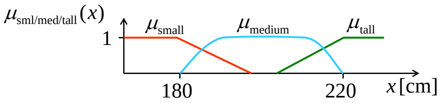


  - Size: small, medium, tall

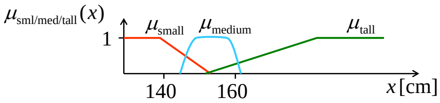


Linguistic Values & Context

- what about basketball players?

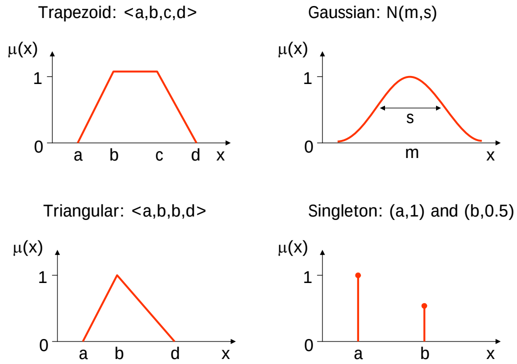


- …and jockeys?



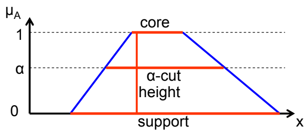
⇒ linguistic values are inherently context dependent!

Types of Membership Functions



Fuzzy Membership Function: Basic Concepts

- **Support**: elements having non-zero degree of membership
- **Core**: set with elements having degree of 1
- **α-Cut**: set of elements with degree >= α
- **Height**: maximum degree of membership



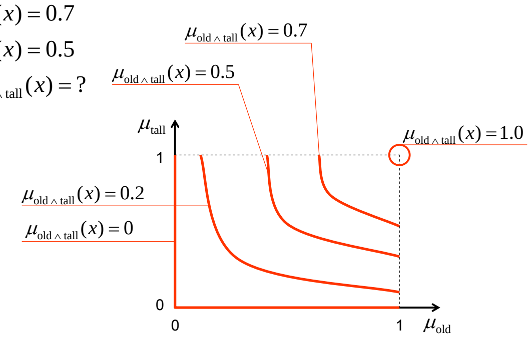
Operators on Fuzzy Sets

- Set of old and tall people (Conjunction)

$\mu_{old}(x) = 0.7$

$\mu_{tall}(x) = 0.5$

$\mu_{old \wedge tall}(x) = ?$



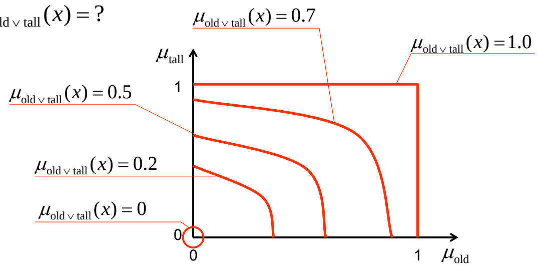
Operators on Fuzzy Sets

- Set of old or tall people (Disjunction)

$\mu_{old}(x) = 0.7$

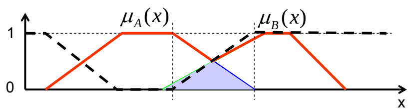
$\mu_{tall}(x) = 0.5$

$\mu_{old \vee tall}(x) = ?$



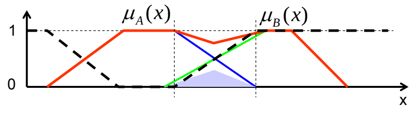
Min / Max-Norm

- Classical Fuzzy Operators: Min/Max-Norm
  - Conjunction: $\mu_{A \wedge B}(x) := \min\{\mu_A(x), \mu_B(x)\}$
  - Disjunction: $\mu_{A \vee B}(x) := \max\{\mu_A(x), \mu_B(x)\}$
  - Negation: $\mu_{\neg A}(x) := 1 - \mu_A(x)$



Product / Bounded-Sum

- Classical Fuzzy Operators: Product / Bounded-Sum
  - Conjunction: $\mu_{A\wedge B}(x) := \mu_A(x)\cdot\mu_B(x)$
  - Disjunction: $\mu_{A\vee B}(x) := \mu_A(x) + \mu_B(x) - \mu_A(x)\cdot\mu_B(x)$
  - Negation: $\mu_{\neg A}(x) := 1 - \mu_A(x)$



Lukasiewicz Norm

T-norms and S-norms

S(u,v) = 1 – T(1-u,1-v)   using De Morgans Law
T(u,v) = 1 – S(1-u,1-v)   (A∧B = ¬(¬A∨¬B))

min/max: max(u,v) = 1 – min(1-u,1-v)

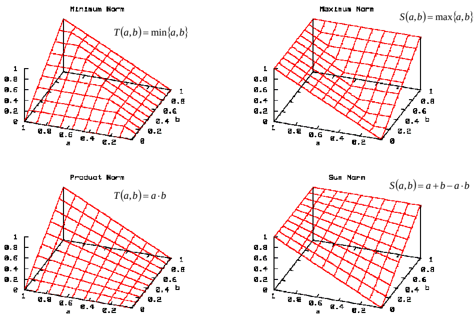
...= 1-1-min(-u,-v) = max(u,v)

Fuzzy T- and S-Norms

T-Norm   S-Norm

Min-Max

Product-Sum

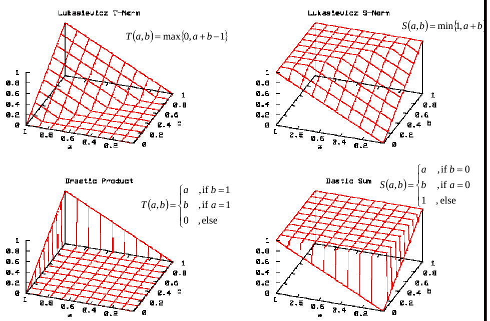


Fuzzy T- and S-Norms

T-Norm   S-Norm

Lukasiewicz Norm

Drastic Product/Sum



Operator - Spectrum

T-norms   S-norms (T-conorms)

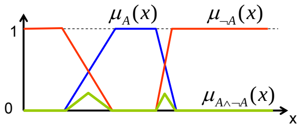
drastic product | bounded product | min | max | bounded sum | drastic sum

drastic product : $T(u,v)=\begin{cases} u & \text{if } v=1 \\ v & \text{if } u=1 \\ 0 & \text{otherwise}\end{cases}$

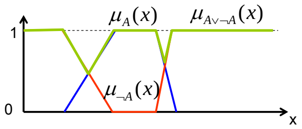
drastic sum : $S(u,v)=\begin{cases} u & \text{if } v=0 \\ v & \text{if } u=0 \\ 1 & \text{otherwise}\end{cases}$

Fuzzy Norms: Issues…

- Interesting effects:
  - A ∧ ¬A = ?



  - A ∨ ¬A = ?



Fuzzy Implication

- One possibility:
  - derive Implication via tautology „A→B = ¬A∨(A∧B)“ and min/max norm.

$$\mu_{A\to B}(x) := \underbrace{\max}_{\text{"v"}}\{\underbrace{1-\mu_A(x)}_{\neg A}, \underbrace{\min\{\mu_A(x),\mu_B(x)\}}_{A\wedge B}\}$$

Fuzzy Implication

- Or vice versa:
  - start with Lukasiewicz-Implication:

$$\mu_{A\to B}(x) := \min\{1, 1-\mu_A(x)+\mu_B(x)\}$$

and derive disjunction and conjunction using

  - A∨B = ¬A→B

$$\mu_{A\vee B}(x) := \min\{1, \mu_A(x)+\mu_B(x)\}$$

  - A∧B = ¬ (¬A∨¬B)

$$\mu_{A\wedge B}(x) := 1-\min\{1, 1-\mu_A(x)+1-\mu_B(x)\}$$
$$= \max\{0, \mu_A(x)+\mu_B(x)-1\}$$

Imprecise Reasoning

- Classic Modus Ponens

$$\frac{\begin{array}{c}A\\ A\to B\end{array}}{B} \qquad \frac{\begin{array}{c}\neg A\\ A\to B\end{array}}{?}$$

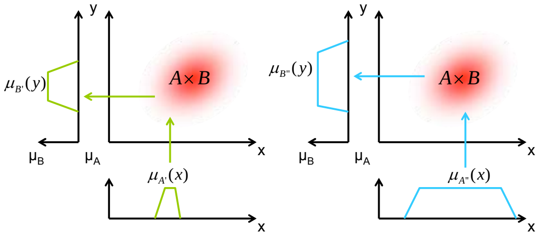
- Imprecise:
  Generalized Modus Ponens

$$\frac{\begin{array}{c}\mu_{A'}(x)\\ A\to B\end{array}}{\mu_{B'}(y)}$$

Joint Constraint (support distribution)

- Using the Min/Max Norm:
  Implication results in a constraint on (x,y)
  ⇒ Cartesian Product A×B: $\mu_{A\times B}(x,y) = \min\{\mu_A(x), \mu_B(y)\}$

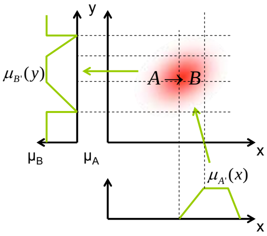
$$\mu_{B'}(y) = \sup_x\{\min\{\mu_{A'}(x), \mu_{A\times B}(x,y)\}\}$$



Conditional Constraint (possibility distribution)

- Alternatively (using Lukasiewicz Norm):
  Implication results in relation expressing possibilities:

$$\mathrm{Poss}(x,y) = \min\{1, 1-\mu_A(x)+\mu_B(y)\}$$
$$\mu_{B'}(y) = \sup_x\{\min\{\mu_{A'}(x), 1-\mu_A(x)+\mu_B(x)\}\}$$



Joint vs. Conditional Constraint

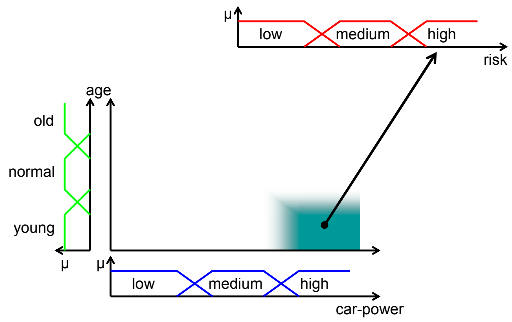
- Joint Constraints:
  - express positive knowledge (facts are supported)
- Conditional Constraints:
  - express negative knowledge (facts are excluded)

Fuzzy Rules

- Rule: IF <Antecedent> THEN <Consequent>
- Fuzzy Version 1: Mamdani Rules
  - Antecedent: Conjunction of fuzzy memberships
  - Consequent: Fuzzy Set

Example: Mamdani Rule

IF age IS young AND car-power IS high THEN risk IS high

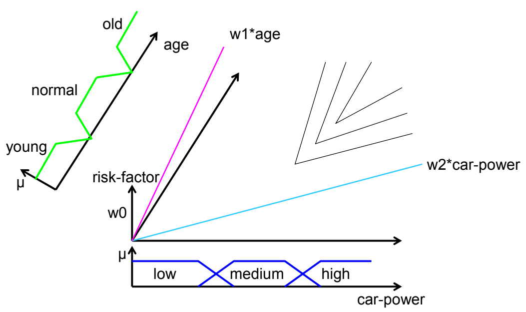


Fuzzy Rules

- Rule: IF <Antecedent> THEN <Consequent>
- Fuzzy Version 1: Mamdani Rules
  - Antecedent: Conjunction of fuzzy memberships
  - Consequent: Fuzzy Set
- Fuzzy Version 2: Takagi-Sugeno -(Kang)- Rules
  - Antecedent: Conjunction of fuzzy memberships
  - Consequent: (usually) real-valued functions of degree 0-2.

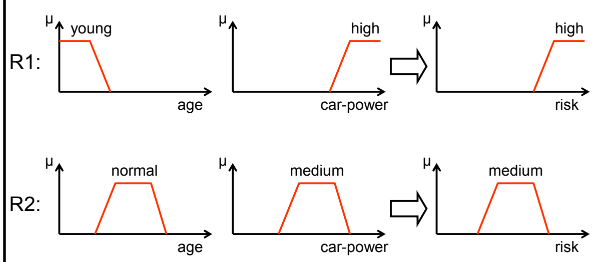
Example: Takagi-Sugeno Rule

IF age IS young AND car-power IS high
THEN risk-factor = w0+w1*age+w2*car-power



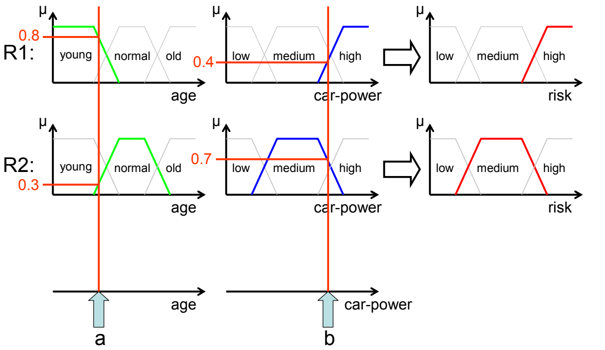
Fuzzy Rule System (Mamdani)

R1: IF age IS young AND car-power IS high THEN risk IS high
R2: IF age IS normal AND car-power IS medium THEN risk IS medium



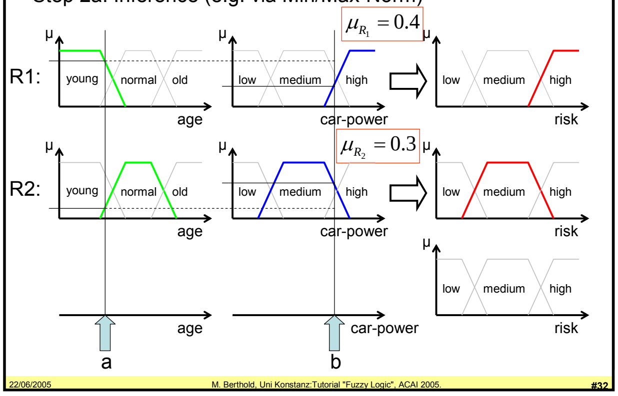
Fuzzy Rule Systems: Example 1

- Step 1: Fuzzification of crisp inputs



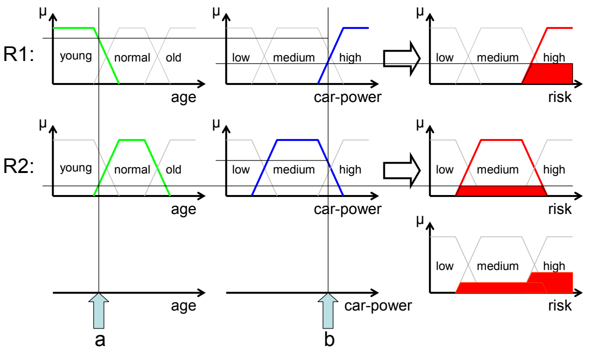
Fuzzy Rule Systems: Example 1

- Step 2a: Inference (e.g. via Min/Max-Norm)



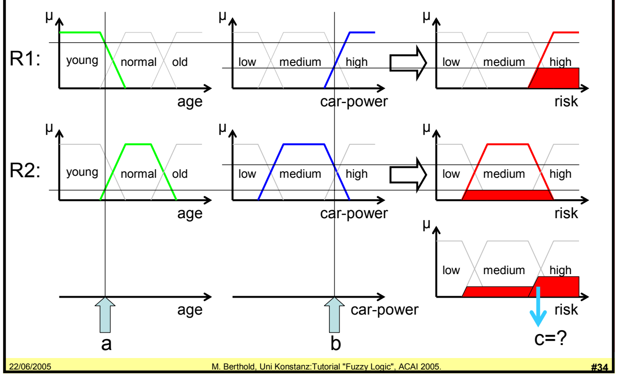
Fuzzy Rule Systems: Example 1

- Step 2b: Inference (e.g. via Maximum-Norm)



Fuzzy Rule Systems: Example 1

- Step 3: Defuzzification



Defuzzification: Center of Gravity

- Center of Gravity:

$$y = \frac{\int y \cdot \mu_{\text{risk}}(y)\,\mathrm{d}y}{\int \mu_{\text{risk}}(y)\,\mathrm{d}y}$$

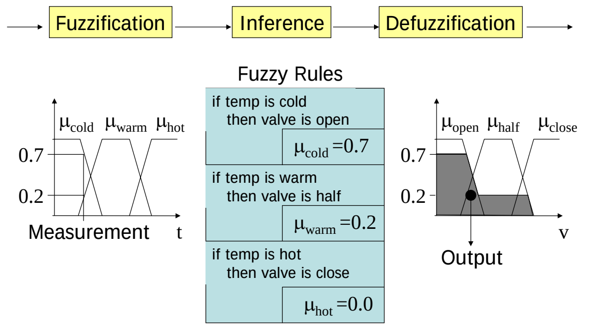
- An easier approximation (with some imprecision…):

$$y = \frac{\sum_{j=1}^{r} \mu_j \cdot s_j}{\sum_{j=1}^{r} \mu_j}$$

Defuzzification: Other Methods

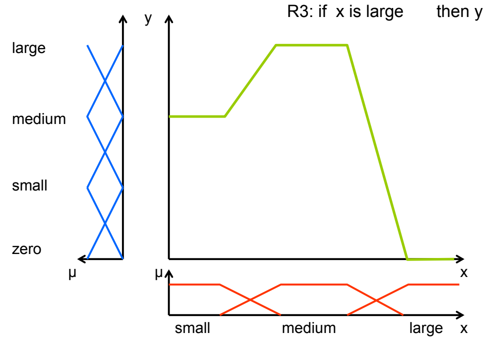
Fuzzy Inference (Mamdani)

- Fuzzy Inference



Fuzzy Rule System (Mamdani)

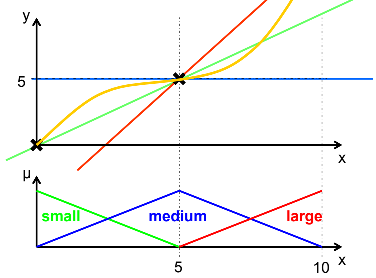
R1: if x is small then y is medium
R2: if x is medium then y is large
R3: if x is large then y is zero



Fuzzy Rule System (Takagi Sugeno)

R1: IF x IS small THEN y = x
R2: IF x IS medium THEN y = 5
R3: IF x IS large THEN y = 2*x-5

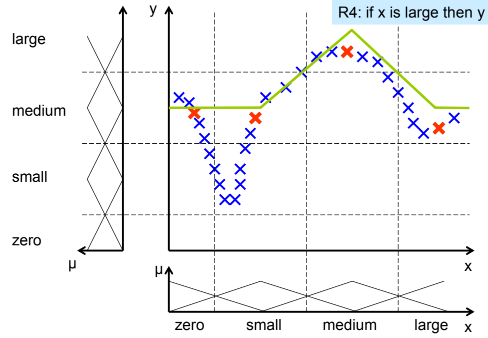
$$y = \frac{\sum_{i=1}^{r} \mu_{R_i} \cdot y_i(\vec{x})}{\sum_{i=1}^{r} \mu_{R_i}}$$



Construction of Fuzzy Rule Systems

- Wang&Mendel Algorithm

R1: if x is zero then y is medium
R2: if x is small then y is medium
R3: if x is medium then y is large
R4: if x is large then y is medium



Grid Based Algorithms

- Exponentially many rules in high dimensional spaces
- If grid is chosen fine enough: arbitrarily good approximation possible (but at what cost!)
- Wrong choice of grid: skipping of extrema
- One extension: Higgins&Goodman Algorithm (see IDA-book)
  - pre-define threshold for desired approximation error
  - finds "best" partitioning (=grid)
  - Disadvantages:
    - concentrates on outliers
    - interpretation difficult: granulation solely data driven
- Other Constructive Algorithms (coming up next…)
  - Local membership functions
  - partially predefined membership functions/granulations possible
  - somewhat tolerant against outliers through local caching

Free Fuzzy Rules

- No global dependence on granulation:
  - individual (per rule) membership functions
  - better modeling of local properties
- Not all attributes used for all rules:
  - individual choice of constraints on few attributes per rule
  - better interpretability in high dimensions
  - no exponential explosion of rules with dimensionality

Formation of Free Fuzzy Rules

Algorithm FRL

1. FORALL training examples (**x**,c) DO
   1. IF correct rule of class c exists:
      - COVERED:
        - increase weight +1
        - adjust core region of rule to cover **x**
   2. ELSE:
      - COMMIT:
        - insert new rule with core=**x**
        - Support = infinite (i.e. rule is not constrained)
   3. SHRINK:
      - Reduce Support of all Rules of conflicting class that cover **x**.
2. Until no more changes occured.

Formation of Free Fuzzy Rules

Algorithmus FRL

Steps:

- COVERED: easy
- COMMIT: easy.
- SHRINK: heuristic
  one solution: volume-based.

Demo.

Formation of Free Fuzzy Rules

Observations:

- FRL finds set of rules completely describing the data
  (as long as training data is conflict free…)
- Each rule is a *partial hypothesis* for a subset of training data
  - core: most specific hypothesis covering subset of training data
  - support: (one of the) most general hypotheses covering subset of data.
  - support is_more_general_than core.
- Core and Support regions can also be seen as:
  - smallest area with highest degree of confidence (we have evidence)
  - largest area without conflict (we have not seen any counter-examples)

Other Types of Fuzzy Rule Learning Methods

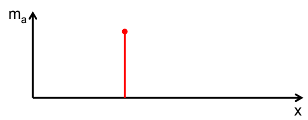
- Constructive:
  - Finding fuzzy rules by growing from singletons
    (⇔ FRL shrinks from most general until it fits to data)
- GRID:
  - merge grid cells (or rows/columns) if no points covered or same class predicted
- Adaptive:
  - Initialize rules randomly (or via expert knowledge) and optimize rule parameters (location, sometime also number of membership functions) iteratively (gradient descent, heuristic hill climbing algorithms).
- Neuro-Fuzzy (later…):
  - inject fuzzy rules into neural network and use NN training algorithms.

Fuzzy Numbers: Motivation

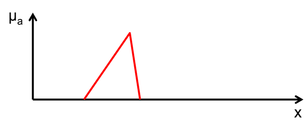
- Often it is desirable to preprocess imprecise inputs
  - add or multiply fuzzy numbers
  - apply function (normalization, …) to fuzzy numbers
- Fuzzy Controller can be extended to process imprecise inputs via use of fuzzy numbers.
- Fuzzy Numbers: Model Imprecision in e.g. measurement.

Fuzzy Numbers

- Classic ("crisp") Numbers: $a \in R$
  - Characteristic Function $m_a(x) \in \{0,1\}$



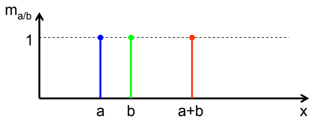
- Fuzzy Number („about a"):
  - membership function $\mu_a(x) \in [0,1]$



Membership functions of fuzzy numbers are (usually):
- normalized
- monotone (left/right)
- singular

Operations on Fuzzy Numbers

- Addition of two crisp Numbers:



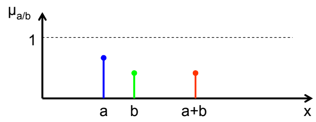
$$m_{a+b}(z) = \begin{cases} 1 & z = a+b \\ 0 & \text{else} \end{cases}$$

$$m_{a+b}(z) = \max_{x,y\in R}\{m_a(x) \cdot m_b(y) \mid x+y=z\}$$

= "and"?...

Operations on Fuzzy Numbers

- Addition of two Singletons:



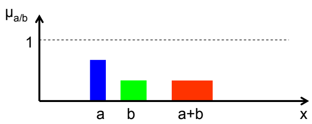
$$\mu_{a+b}(z) = \begin{cases} \min\{\mu_a(a), \mu_b(b)\} & z = a+b \\ 0 & \text{else} \end{cases}$$

$$\mu_{a+b}(z) = \max_{x,y\in R}\{\min\{\mu_a(x), \mu_b(y)\} \mid x+y=z\}$$

= and

Operations on Fuzzy Numbers

- Addition of two (fuzzy) intervals:

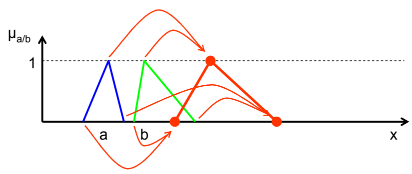


$$\mu_{a+b}(z) = \max_{x,y\in R}\{\min\{\mu_a(x), \mu_b(y)\} \mid x+y=z\}$$

= and

Operations on Fuzzy Numbers

- And finally: Addition of two Fuzzy Numbers:

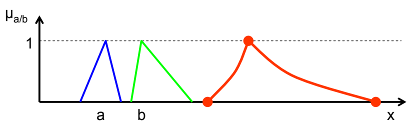


$$\mu_{a+b}(z) = \max_{x,y\in R}\{\min\{\mu_a(x), \mu_b(y)\} \mid x+y=z\}$$

what about other operators, for example multiplication?

Operations on Fuzzy Numbers

- Multiplication of two Fuzzy Numbers:



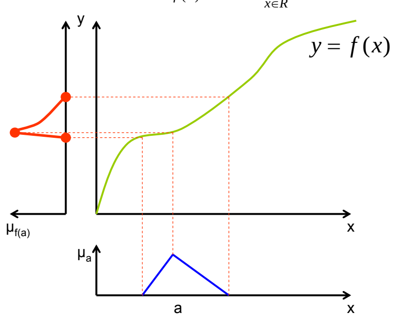
$$\mu_{a*b}(z) = \max_{x,y\in R}\{\min\{\mu_a(x), \mu_b(y)\} \mid x*y=z\}$$

what about arbitrary functions?

Operations on Fuzzy Numbers

- Crisp Functions applied to Fuzzy Numbers:

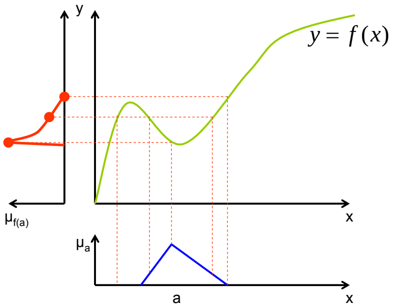
$$\mu_{f(a)}(y) = \max_{x\in R}\{\mu_a(x) \mid y = f(x)\}$$



Operations on Fuzzy Numbers

- Crisp (non monotone) function applied to Fuzzy Numbers:

$$\mu_{f(a)}(y) = \max_{x\in R}\{\mu_a(x) \mid y = f(x)\}$$



Operations on Fuzzy Numbers

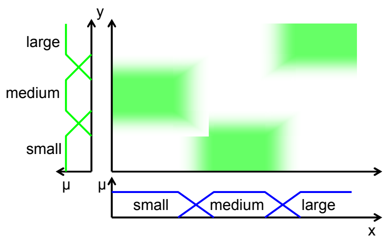
- Based on Extension Principle:

$$\mu_{f(a)}(y) = \max_{x_1,\dots,x_n\in R}\{\min\{\mu_{a_1}(x_1),\dots,\mu_{a_n}(x_n)\} \mid y = f(x_1,\dots,x_n)\}$$

- Computable in Practice:
  - via series of α-cuts
  - via polynomial representation of left and right side of fuzzy numbers (closed under addition and multiplication a.o.)
  - via granulation on variables (grid based)
  - …

Imprecise Functions (Fuzzy Graphs)

- Real-valued Function: infinite set of crisp points
- Fuzzy Graph: finite set of imprecise points
  - IF x IS small THEN y IS medium
  - IF x IS medium THEN y IS small
  - IF x IS large THEN y IS large



Imprecise Functions (Fuzzy Graphs)

- Each fuzzy point describes an imprecise relation:

$$(\vec{x}, y)\ \text{IS}\ \mathbf{A}\times B$$

- A fuzzy graph is a collection of fuzzy points:

$$(\vec{x}, y)\ \text{IS}\ \mathbf{A}_1\times B_1\ \text{OR}\ \dots\ \text{OR}\ \mathbf{A}_r\times B_r$$

- Reasoning using fuzzy graphs:

$$\vec{x}\ \text{IS}\ \mathbf{A}$$

$$\tilde{f}\ \text{IS}\ \bigcup_{j=1}^{r}\mathbf{A}_j\times B_j$$

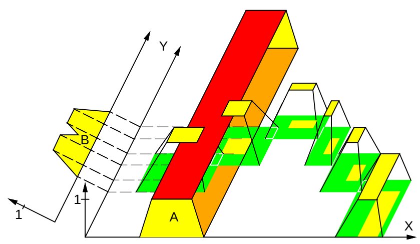
$$\overline{}$$

$$y\ \text{IS}\ B$$

Fuzzy Graphs

- Approximation through cylindrical extension:

$$B = \text{proj}_y\{(\mathbf{A}\times I)\cap \tilde{f}\}$$



Fuzzy Graphs

- Learning of Fuzzy Graphs from Data: Demo
  - Influence of granulation of target variable
  - Influence of outliers: large number of fuzzy points
    ⇒ avoidance via outlier model

Wrap-Up / Conclusions

- Fuzzy Methods
  - model imprecise knowledge (fuzzy rules)
  - draw imprecise conclusions
- Main Concepts:
  - Fuzzy Set and Degree of Membership
  - Fuzzy Set Operators
  - Extension Principle: allow to extend classical functions to fuzzy numbers
- Algorithms (mostly heuristic) to
  - find fuzzy rules in data
  - build fuzzy approximators (fuzzy graphs/rule systems)
- Fuzzy Controllers in wide spread use