



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Reintegrace českého vědce
a vytvoření špičkového týmu v informačních vědách
Fuzzy logic 2

J. Konecny

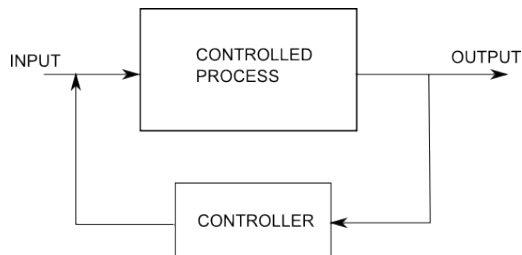
17. listopadu 2011

Student seminar, Fuzzy Rule-Based Systems

Fuzzy Rule-Based Systems

- Particular input/output systems employed in approximate reasoning.
- Approximate reasoning \sim making approximate conclusions from approximate premises.
- Fuzzy rule-based systems are based on so-called IF-THEN rules.
- Main application area: fuzzy control.

General scheme of control system



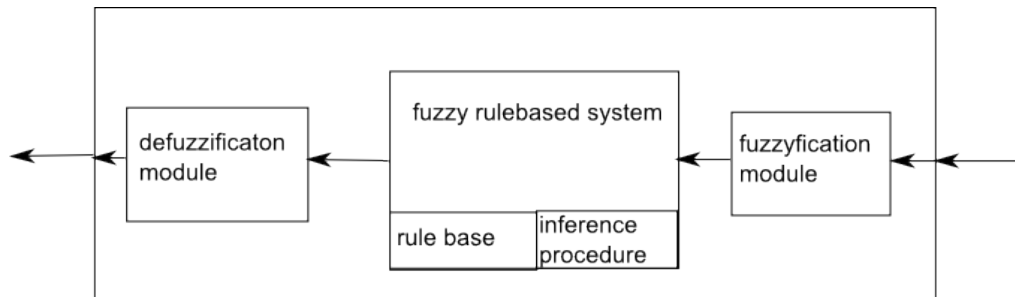
role of controller:

- acts in a way which keeps the output (e.g. temperature) a desired level.
- looks at (measures) output and influences input according to some control strategy

Fuzzy rule-based systems

Fuzzy controllers

- alternative to classical controllers
- general scheme:



Residuated Lattices

Definition

Complete residuated lattice is an algebra $\mathbf{L} = \langle L, \wedge, \vee, \otimes, \rightarrow, 0, 1 \rangle$ where

- $\langle L, \wedge, \vee, 0, 1 \rangle$ is complete lattice with 0 and 1 being the least and the greatest element of L ,
- $\langle L, \otimes, 1 \rangle$ is commutative monoid, i.e. \otimes is associative and commutative and $x \otimes 1 = x$ holds for all $x \in L$,
- the operations \otimes and \rightarrow form an adjoint pair, i.e.

$$a \otimes b \leq c \quad \text{iff} \quad b \leq a \rightarrow c$$

holds for each $a, b, c \in L$.

Remarks:

- Operation \rightarrow is uniquely determined by \otimes .
- Set of truth degrees of classical two-valued logic is special case of residuated lattice.

Examples of Residuated Lattices

The most commonly used residuated lattices are ones with $L = [0, 1]$.

For \wedge and \vee we take minimum and maximum, respectively.

As operations \otimes and \rightarrow we use so called t-norms and corresponding residua.

Three basic continuous t-norms are:

$$a \otimes b = \begin{cases} \max(0, a + b - 1) & \text{\u0179ukasiewicz t-norm} \\ \min(a, b) & \text{G\u00f6del t-norm} \\ a \otimes b = a \cdot b & \text{product t-norm} \end{cases}$$

Their residua are

$$a \rightarrow b = \begin{cases} \min(1 - a + b, 1) & \text{\u0179ukasiewicz} \\ a \rightarrow b = 1 \text{ if } a \leq b \text{ else } b & \text{G\u00f6del.} \\ a \rightarrow b = 1 \text{ if } a \leq b \text{ else } b/a & \text{product} \end{cases}$$

The corresponding algebras are called standard \u0179ukasiewicz algebra, standard G\u00f6del algebra, and standard product algebra.

Definition

Let X be a set (universe of discourse), \mathbf{L} be a complete residuated lattice. An \mathbf{L} -set A in X is a mapping $A : X \rightarrow L$.

Operations with \mathbf{L} -sets are defined componentwise. Therefore, we have

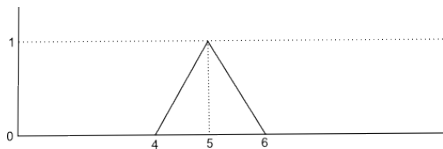
$$(A \cap B)(x) = A(x) \wedge B(x),$$

$$(A \cup B)(x) = A(x) \vee B(x),$$

...

for each $x \in X$.

Graphical representation of fuzzy sets:



An \mathbf{L} -set $A \in \mathbf{L}^X$ is also denoted $\{A(x)/x \mid x \in X\}$. If for all $y \in X$ distinct from x_1, x_2, \dots, x_n we have $A(y) = 0$, we also write

$$\{A(x_1)/x_1, A(x_2)/x_1, \dots, A(x_n)/x_n\}.$$

L-Relations

In classical set theory a (binary) relation is defined as a subset of a cartesian product. In fuzzy case the definition is analogous.

Definition (n -ary fuzzy relation)

An n -ary L -fuzzy relation between sets U_1, \dots, U_n is an L -fuzzy set in $U_1 \times \dots \times U_n$. If $U_1 = \dots = U_n = U$, we speak of an n -ary L -fuzzy relation in U .

- That is, an n -ary L -fuzzy relation R is a mapping $R : U_1 \times \dots \times U_n \rightarrow L$. We assume again that $\mathbf{L} = \langle L, \dots \rangle$ is a complete residuated lattice.
- For $u_i \in U_i$ ($i = 1, \dots, n$), $R(u_1, \dots, u_n)$ is interpreted as a degree to which u_1, \dots, u_n are related.
- Occasionally, we say just fuzzy relation or \mathbf{L} -relation.
- Obviously, if $L = \{0, 1\}$, the concept of an \mathbf{L} -relation coincides with that of (characteristic function of) ordinary relation.

Representation of Binary Fuzzy Relations

Analogously to ordinary binary relations, binary fuzzy relations can be **represented by matrices** (tables). Example: Let $X = \{a, b, c\}$, $Y = \{1, 2, 3, 4\}$. A binary fuzzy relation with $L = [0, 1]$ given by

$$R = \{1/\langle a, 1 \rangle, 0.5/\langle a, 2 \rangle, 0.1/\langle a, 4 \rangle, 0.8/\langle b, 2 \rangle, 1/\langle b, 4 \rangle, 0.8/\langle c, 1 \rangle\}$$

R can be represented by table (left) or an $[0, 1]$ -valued matrix \mathbf{M}_R (right).

R	1	2	3	4
a	1	0.5	0	0.1
b	0	0.8	0	1
c	0.8	0	0	0

$$\mathbf{M}_R = \begin{pmatrix} 1 & 0.5 & 0 & 0.1 \\ 0 & 0.8 & 0 & 1 \\ 0.8 & 0 & 0 & 0 \end{pmatrix}. \text{ Matrix } \mathbf{M}_R \text{ representing}$$

fuzzy relation $R \in L^{\{x_1, \dots, x_m\} \times \{y_1, \dots, y_n\}}$ is an $m \times n$ -matrix with entries m_{ij} defined by

$$m_{ij} = R(x_i, y_j).$$

Graph representation: $R \in L^{\{x_1, \dots, x_m\} \times \{y_1, \dots, y_n\}}$ is represented by an oriented graph, $m + n$ nodes correspond to $x_1, \dots, x_m, y_1, \dots, y_n$, if $R(x_i, y_j) > 0$, we add an arrow from node corresponding to x_i to node corresponding to y_j and attach $R(x_i, y_j)$ as a label to this arrow.

We want to define an analogy of cartesian product of two sets.

Definition

Let $A \in L^X$ and $B \in L^Y$ be \mathbf{L} -sets.

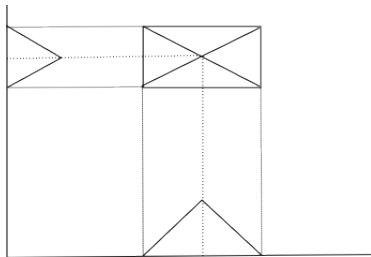
- \otimes -product of A and B is an \mathbf{L} -relation $R \in L^{X \times Y}$ defined by

$$R(x, y) = A(x) \otimes B(y)$$

- \times -product of A and B is an \mathbf{L} -relation $R \in L^{X \times Y}$ defined by

$$R(x, y) = A(x) \wedge B(y)$$

... similarly for n sets.



Rule Base

Rule base is a set $\mathcal{R} = \{R_i \mid i = 1, \dots, m\}$ of IF-THEN rules.

- Rules R_i are linguistic rules describing relationships of input variables $x_j, j = 1, \dots, n$ and output variable y .
- Usually, x_j and y are numeric variables (temperature, change of temperature, r.p.m., speed, ...). X_j is set of values for x_j , Y is set of values for y
- General form of rule R_i :

IF x_1 is \mathcal{A}_{i1} & ... & x_n is \mathcal{A}_{in} THEN y is \mathcal{B}_i ,

where \mathcal{A}_{ij} are linguistic expressions such as MEDIUM, BIG, VERY BIG, ...

Example

$n = 2$ (two inputs): $x_1 \dots$ temperature, $x_2 \dots \Delta$ temperature (change of temperature).

Output $y \dots$ r.p.m. (rotation per minute of a fan/ventilator)

IF x_1 is HIGH & x_2 is MEDIUM THEN y is HIGH

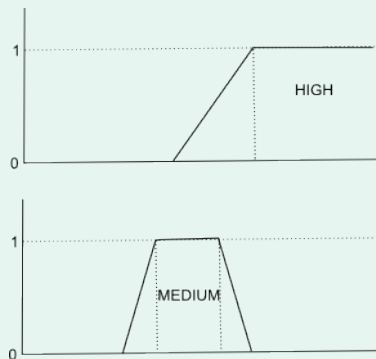
meaning: if temp. is high and Δ temp. is medium then we need to set r.p.m. to high (to cool the system).

How to represent linguistic rulebase mathematically?

- 1 assign L-set A_{ij} in X_j to every \mathcal{A}_{ij} and L-set B_i in Y to every \mathcal{B}_i

Example

HIGH temperature and MEDIUM Δ temperature:



similarly for HIGH r.p.m.

How to represent linguistic rulebase mathematically? (cont.)

2. represent rule \mathcal{R}_i by **L**-relation R_i between $X_1 \times \cdots \times X_n$ (input space) and Y (output space).

Definition of R_i for $x_1 \in X_1, \dots, x_n \in X_n, y \in Y$:

$$R_i(x_1, \dots, x_n, y) = A_{i1}(x_1) \otimes \cdots \otimes A_{in}(x_n) \otimes B_i(y)$$

3. represent \mathcal{R} by fuzzy relation R between $X_1 \times \cdots \times X_n$ and Y

$$R = \bigcup_{i=1}^m R_i \quad \text{that is} \quad R_i(x_1, \dots, x_n, y) = \bigvee_{i=1}^n R_i(x_1, \dots, x_n, y)$$

Compositional Rule of Inference (CRI)

CRI is a technique used in the fuzzy rule-based systems.

Definition (CRI)

Let R be an \mathbf{L} -relation between sets X and Y , A be an \mathbf{L} -set in X . The \mathbf{L} -set B obtained from A and R by compositional rule of inference is defined by

$$B(y) = \bigvee_{x \in X} A(x) \otimes R(x, y),$$

or just $B = A \circ R$, for short.

- In fuzzy logic, CRI is considered an inference rule. Sometimes various wrong claims are being made such as “CRI generalizes the rule of modus ponens”.

- $B(y)$ is a truth degree of “there is $x \in X$ such that x is in A and x is related to y via R ($\langle x, y \rangle$ is in R)”.
- Also, $A \circ R$ can be looked at as matrix multiplication. Let $|X| = m$, $|Y| = n$, let M_A be a $1 \times m$ matrix representing A , M_R be an $m \times n$ matrix representing R . Then $M_A \circ M_R$ is a $1 \times n$ matrix representing $A \circ R$.
- If $X = Y$ and \approx is a fuzzy equivalence relation (similarity) in X , then $(A \circ \approx)(y) = \bigvee_{x \in X} A(x) \otimes (x \approx y)$, i.e. $(A \circ \approx)(y)$ is the truth degree of “there is x in A which is similar to y ”. This interpretation of CRI = applications of fuzzy logic in information retrieval. If A is a collection of “prototypes” a user is interested in, $A \circ \approx$ is the collection of objects similar to some of the prototypes. See next example.

Example

\approx	911	RAV4	Outback	Corolla	Civic	Accord
Porsche 911	1	0.5	0.6	0	0	0
Toyota RAV4	0.5	1	0.9	0	0	0
Subaru Outback	0.6	0.9	1	0	0	0
Toyota Corolla	0	0	0	1	0.9	0.8
Honda Civic	0	0	0	0.9	1	0.8
Honda Accord	0	0	0	0.8	0.8	1

Consider Łukasiewicz operations.

User asks: I am interested in Honda Civic, show me similar cars.

Using CRI, the answer is $A \circ \approx$ with $A = \{^1/\text{Honda Civic}\}$.

In particular, $\{^1/\text{Honda Civic}\} \circ \approx = \{^1/\text{H. Civic}, ^{0.9}/\text{T. Corolla}, ^{0.8}/\text{H. Accord}\}$.

User asks: I am interested in Honda Civic and little bit (to degree 0.4) in Subaru Outback. The answer is:

$\{^1/\text{H. Civic}, ^{0.4}/\text{S. Outback}\} \circ \approx =$

$\{^1/\text{H. Civic}, ^{0.9}/\text{T. Corolla}, ^{0.8}/\text{H. Accord}, ^{0.4}/\text{S. Outback}, ^{0.3}/\text{T. RAV4}\}$.

Inference procedure

Suppose $\mathcal{R} = \{\mathcal{R}_i \mid i = 1, \dots, m\}$ is a system of rules. \mathcal{R} is represented by \mathbf{L} -relation R , \mathcal{R}_i s are represented by \mathbf{L} -relations R_i .

Given input \mathbf{L} -sets A'_1 in X_1, \dots, A'_n in X_n
(actual observation/descriptions of values of x_1, \dots, x_n).

Problem: what is the appropriate output B' – \mathbf{L} -set in Y .

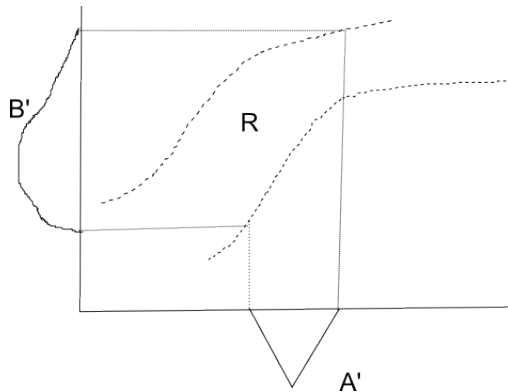
Answer: $B' = A' \circ R$ where A' is a fuzzy set in $X_1 \times \dots \times X_n$ defined by \otimes -product of A'_i . \circ is CRI.

$$A'(x_1, \dots, x_n) = A_1(x_1) \otimes \dots \otimes A_n(x_n)$$

usually \min is used as \otimes .

Geometrical Meaning

First geometric look at $B' = A' \circ R$ ($n = 1$)



Second geometric look at $B' = A' \circ R$

Consider special case where A'_1, \dots, A'_n are singletons

$$A'_1 = \{1/x_1\}, \dots, A'_n = \{1/x_n\}.$$

Note: This is important case. Namely, this is how the inference usually looks in fuzzy controllers (x_i is the actual temperature, etc.)

Then equations get simplified:

$$\begin{aligned} B'(y) &= (A' \circ R)(y) \\ &= \bigvee_{z_1 \in X_1} A'_1(z_1) \otimes \dots \otimes A'_n(z_n) \otimes R(z_1, \dots, z_n, y) \\ &\quad \vdots \\ &\quad z_n \in X_n \\ &= A'_1(x_1) \otimes \dots \otimes A'_n(x_n) \otimes R(x_1, \dots, x_n, y) \\ &= \bigvee_{i=1}^m R_i(x_1, \dots, x_n, y) \\ &= \bigvee_{i=1}^m A_{i1}(x_1) \otimes \dots \otimes A_{in}(x_n) \otimes B_i(y) \end{aligned}$$

$$B'(y) = \bigvee_{i=1}^m A_{i1}(x_1) \otimes \cdots \otimes A_{in}(x_n) \otimes B_i(y)$$

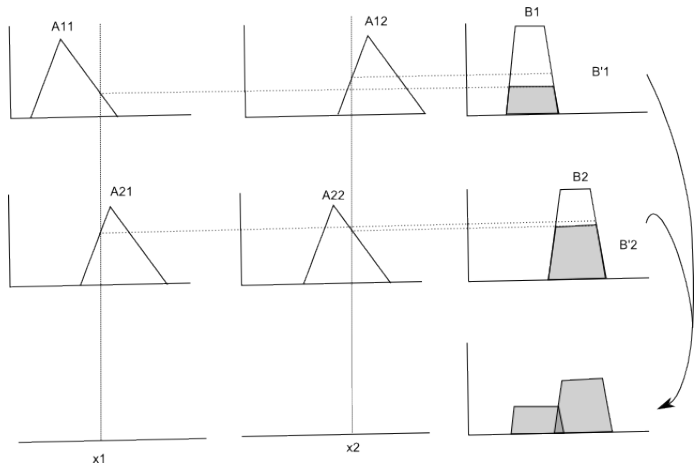
with \otimes being minimum we get

$$B'(y) = \bigvee_{i=1}^m \min(A_{i1}(x_1), \dots, A_{in}(x_n), B_i(y))$$

Thus,

$$B'_i(y) = \min(A_{i1}(x_1), \dots, A_{in}(x_n), B_i(y))$$

$$B = \bigcup_{i=1}^m B_i$$



Fuzzy Rule-Based Systems: Example

Consider $X = Y = \{0, 1, \dots, 10\}$ and a system \mathcal{R} consisting of three rules

IF x is \mathcal{A}_1 THEN y is \mathcal{B}_1 ,

IF x is \mathcal{A}_2 THEN y is \mathcal{B}_2 ,

IF x is \mathcal{A}_3 THEN y is \mathcal{B}_3 .

Let the corresponding fuzzy sets $A_1, B_1, A_2, B_2, A_3, B_3$ be given by:

x/y	0	1	2	3	4	5	6	7	8	9	10
$A_1(x)$	0	0.5	1	0.5	0	0	0	0	0	0	0
$B_1(y)$	0	0.5	0.75	1	1	0.75	0.5	0	0	0	0
$A_2(x)$	0	0	0.25	0.75	1	0.75	0.25	0	0	0	0
$B_2(y)$	0	0	0	0	0	0	0.5	1	0.5	0	0
$A_3(x)$	0	0	0	0	0	0	0.5	1	0.5	0	0
$B_3(y)$	0	0	0	0	0	0	0	0	0.5	1	0.5

Determine the corresponding fuzzy relations R_1, R_2, R_3 , and R .

Recall: $R_i(x, y) = A_i(x) \wedge B_i(y)$, $R = R_1 \cup R_2 \cup R_3$.

R_1

10	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0
6	0	0.5	0.5	0.5	0	0	0	0	0	0	0
5	0	0.5	0.75	0.5	0	0	0	0	0	0	0
4	0	0.5	1	0.5	0	0	0	0	0	0	0
3	0	0.5	1	0.5	0	0	0	0	0	0	0
2	0	0.5	0.75	0.5	0	0	0	0	0	0	0
1	0	0.5	0.5	0.5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
R_1	0	1	2	3	4	5	6	7	8	9	10

R_2

10	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0.25	0.5	0.5	0.5	0.25	0	0	0	0
7	0	0	0.25	0.75	1	0.75	0.25	0	0	0	0
6	0	0	0.25	0.5	0.5	0.5	0.25	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
R_2	0	1	2	3	4	5	6	7	8	9	10

R_3

10	0	0	0	0	0	0	0.5	0.5	0.5	0	0
9	0	0	0	0	0	0	0.5	1	0.5	0	0
8	0	0	0	0	0	0	0.5	0.5	0.5	0	0
7	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
R_3	0	1	2	3	4	5	6	7	8	9	10

R

10	0	0	0	0	0	0	0.5	0.5	0.5	0	0
9	0	0	0	0	0	0	0.5	1	0.5	0	0
8	0	0	0.25	0.5	0.5	0.5	0.5	0.5	0.5	0	0
7	0	0	0.25	0.75	1	0.75	0.25	0	0	0	0
6	0	0.5	0.5	0.5	0.5	0.5	0.25	0	0	0	0
5	0	0.5	0.75	0.5	0	0	0	0	0	0	0
4	0	0.5	1	0.5	0	0	0	0	0	0	0
3	0	0.5	1	0.5	0	0	0	0	0	0	0
2	0	0.5	0.75	0.5	0	0	0	0	0	0	0
1	0	0.5	0.5	0.5	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
R	0	1	2	3	4	5	6	7	8	9	10

Now, make an inference for $A' = \{2, 0.5/3\}$.

Defuzzification = transformation of a fuzzy set (output B' of inference) to a “typical” value y from the output space Y , i.e. a value which characterizes B' well.

Several approaches exist. We assume that the universe $Y \subseteq \mathbb{R}$ is finite, i.e. $Y = \{y_1, \dots, y_k\}$. A defuzzification method can be seen as a function $D : [0, 1]^Y \rightarrow Y$ assigning an element $D(B) \in Y$ to a fuzzy set $B \in [0, 1]^Y$.

Basic defuzzification methods:

- **Center of gravity** (COG, also center of area, centroid):

$$D(B) = \frac{\sum_{i=1}^k B(y_i)y_i}{\sum_{i=1}^k B(y_i)}.$$

If $Y = [a, b] \subseteq \mathbb{R}$ is a real interval, then

$$D(B) = \frac{\int_a^b B(y)ydy}{\int_a^b B(y)dy},$$

which means that $D(B)$ is the y -th coordinate of the center of gravity of the area delineated by B (area under B).

- **Center of maxima** (COM): put $M(B) = \{z \in Y \mid B(z) = h(B)\}$, where $h(B) = \bigvee_{y \in Y} B(y)$ is the height of B . Then

$$D(B) = \frac{\min(M(B)) + \max(M(B))}{2},$$

i.e. $D(B)$ is the average of the least and the greatest value in Y at which B has its maximum.

- **Mean of maxima** (MOM): put $M(B) = \{z \in Y \mid B(z) = h(B)\}$, where $h(B) = \bigvee_{y \in Y} B(y)$ is the height of B . Then

$$D(B) = \frac{\sum_{y \in M} y}{|M|},$$

i.e. $D(B)$ is the average of all values in Y at which B has its maximum.

COG is most popular because it is not sensitive to changes in the defuzzified fuzzy set B .

Example

Let $Y = \{0, 1, \dots, 10\}$. Let

$$B = \{0/0, 0.5/1, 1/2, 1/3, 1/4, 0.5/5, 0/6, 0/7, 0.5/8, 1/9, 0.5/10\}$$

Then

– by COG:

$$D(B) = \frac{0 \cdot 0 + 0.5 \cdot 1 + 1 \cdot 2 + 1 \cdot 3 + 1 \cdot 4 + 0.5 \cdot 5 + 0 \cdot 6 + 0 \cdot 7 + 0.5 \cdot 8 + 1 \cdot 9 + 0.5 \cdot 10}{0 + 0.5 + 1 + 1 + 1 + 0.5 + 0 + 0 + 0.5 + 1 + 0.5} = \frac{30}{6} = 5.$$

– by COM: $M = \{2, 3, 4, 9\}$,

$$D(B) = \frac{2+9}{2} = 5.5.$$

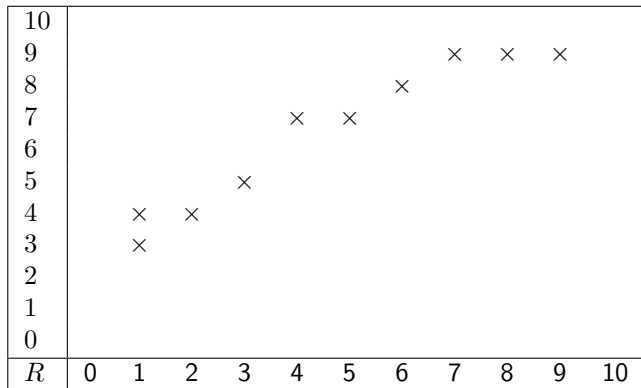
– by MOM: $M = \{2, 3, 4, 9\}$,

$$D(B) = \frac{2+3+4+9}{4} = \frac{18}{4} = 4.5.$$

The function G represented by the rule based system, i.e. $G(x) = D(\{1/x\} \circ R)$:

x	0	1	2	3	4	5	6	7	8	9	10
$D(\{1/x\} \circ R)$	ND	3.5	3.9	4.65	7	7	8.375	9	9	ND	ND

With rounding to closest values in Y :



Universal Mapping Property of Fuzzy Rule-Based Systems

Let X_1, \dots, X_n (input space) and Y (output space) be closed real intervals. Given a rule base \mathcal{R} (including fuzzy sets A_{ij} and B_i which represent meaning of linguistic terms in the rules, with R being the corresponding fuzzy relation), an inference method \circ (such as CRI), a fuzzification method F (such as singleton), and a defuzzification method D (such as COG), there is an associated function $G : X_1 \times \dots \times X_n \rightarrow Y$ defined by

$$G(x_1, \dots, x_n) = D(\langle F(x_1), \dots, F(x_n) \rangle \circ R).$$

That is, for x_1, \dots, x_n , we take their fuzzifications $A'_1 = F(x_1), \dots, A'_n = F(x_n)$ (e.g. singletons), use the rule base and the inference method to compute the output fuzzy set $B' = \langle F(x_1), \dots, F(x_n) \rangle \circ R$ and set $G(x_1, \dots, x_n)$ to be the result of defuzzification of B' .

Question: What types of functions can we represent this way? Namely, function G represents a control strategy. So, are there any limitations to what control strategies can be implemented using fuzzy-rule based systems?

Universal Mapping Property of Fuzzy Rule-Based Systems

For several particular choices of the inference mechanism, fuzzification and defuzzification method, one can prove the following theorem (we omit details):

Theorem (UMP of fuzzy rule-based systems)

Let $f : X_1 \times \dots \times X_n \rightarrow Y$ be a continuous function. For every $\varepsilon > 0$ there exists a fuzzy rule-based system such that for the associated function G and any $x_1 \in X_1, \dots, x_n \in X_n$ we have

$$|f(x_1, \dots, x_n) - G(x_1, \dots, x_n)| \leq \varepsilon.$$

That is, any continuous function can be approximated by a fuzzy rule-based system with arbitrary precision.

This theorem justifies theoretically the universality of fuzzy controllers.

Do you want to know more?

Textbooks

- Dubois, Prade: Fuzzy Sets and Systems: Theory and Applications. Academic Press, 1982.
- Klir, Yuan: Fuzzy Sets and Fuzzy Logic. Theory and Applications. Prentice Hall, 1995.
- Nguyen, Walker: A First Course in Fuzzy Logic. Chapman & Hall/CRC Press, 2000.

Monographs

- Hajek: Metamathematics of Fuzzy Logic. Kluwer, 1998.
- Gottwald: A Treatise on Many-Val. Logics. Res. Stud. Press, 2001.
- Gerla: Fuzzy Logic. Math. Tools for Appr. Reasoning. Kluwer, 2001
- Novak, Perfilieva, Mockor: Mathematical Principles of Fuzzy Logic. Kluwer, 1999.
- Belohlavek: Fuzzy Relational Systems: Foundations and Principles. Kluwer, 2002.