Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems II

SSIE 617 – Fall 2008

Radim BELOHLAVEK

Dept. Systems Sci. & Industrial Eng.
Watson School of Eng. and Applied Sci.
Binghamton University – SUNY

Fuzzy Relations

Definition (n-ary fuzzy relation)

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

- That is, an n-ary L-fuzzy relation R is a mapping $R : U_1 \times \cdots \times U_n \to L$. We assume again that $\mathbf{L} = \langle L, \cdots \rangle$ is a complete residuated lattice.
- For $u_i \in U_i$ $(i = 1, \ldots, n)$, $R(u_1, \ldots, u_n)$ is interpreted as a degree to which $u_1, \ldots, u_n$ 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.

J. A. Goguen (1967)

...the importance of fuzzy relations is almost self-evident. Science is, in a sense, the discovery of relations between observables ... Difficulties arise in so-called "soft" sciences because the relations involved do not appear to be "hard" as they are, say, in classical physics ...

Example (fuzzy relations)

- Similarity: Let U be a set of objects. Let $\approx: U \times U \to [0,1]$ assign to every $u, v \in U$ a degree $(u \approx v) \in [0,1]$ to which u and v are similar. $\approx$ is a binary fuzzy relation. Many measures of similarity for various kinds of objects have been proposed in the literature.
- Betweennes: Let U be a set of points in a plane. Let $R(u, v, w) \in L$ denote degree to which v lies approximately on the line connecting u and w. R is a fuzzy relation. Explicit rule:
  $R(u, v, w) = max(0, 1 - d(v, uw))$, where $d(v, uw)$ is the distance between point u and the line connecting u and w.

Example (fuzzy relations)

Relational databases: Let $D \subseteq U_1 \times \cdots \times U_n$ be an n-ary relation. D can be seen as a database (table) with columns corresponding to $U_1, \ldots, U_n$ (U_i are called domains in relational databases). $\langle u_1, \ldots, u_n \rangle \in D$ means that $\langle u_1, \ldots, u_n \rangle$ is stored in the database (i.e. $\langle u_1, \ldots, u_n \rangle$ is one row of the table).

name	age	education
Adams A.	50	Business
Brown B.	31	Computer Science
Cross C.	30	Law
Dobb D.	25	Computer Engineering
Edwards E.	32	Systems Science
...	...	...

Let Q denote an approximate query such as "show all employees with age around 30" or "show all candidates with education similar to Computer Engineering" or "show all pictures similar to My Photograph". Let for $\langle u_1, \ldots, u_n \rangle \in D$ denote by $R_Q(u_1, \ldots, u_n)$ a degree to which $\langle u_1, \ldots, u_n \rangle \in D$ satisfies query Q.

Example (cntd.)

Then R_Q is a fuzzy relation. It can be depicted in a table with the first column of row $\langle u_1, \ldots, u_n \rangle$ containing the degree $R_Q(u_1, \ldots, u_n)$.
Examples: For Q being "show all employees with age around 30", the table describing R_Q (provided $R_Q(u_1, u_2, u_3) = \max(0, 1 - 0.25 \cdot |30 - u_2|)$) is:

R_Q	name	age	education
0	Adams A.	50	Business
0.75	Brown B.	31	Computer Science
1.0	Cross C.	30	Law
0	Dobb D.	25	Computer Engineering
0.5	Edwards E.	32	Systems Science
	...	...	...

exercise

- Think of various vague relationships such as "being close", "being similar", "being a prerequisite for", etc. Try to formalize the relationships by means of fuzzy relations. Hint: In case of similarity of objects, try to find a suitable representation of objects and define the similarity degree based on the representation. For instance, if object x represented by a set $A_x \subseteq F$ of features in F (such as $A_{\mathrm{John}} = \{$male, married, MS$\}$), we may define

$$x \approx y = \frac{|A_x \cap A_y|}{|A_x \cup A_y|}$$

  if $A_x \cap A_y \neq \emptyset$ and $x \approx y = 1$ if $A_x \cap A_y = \emptyset$.
- Go back to the databases example. Think of various examples of databases. Formulate various approximate queries and show the corresponding fuzzy relations.

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

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

Matrix $\mathbf{M}_R$ 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.

Operations with Fuzzy Relations

- Fuzzy relations are fuzzy sets. Therefore, any general operation with fuzzy sets (intersections, unions, a-cuts, ...) or relationship between fuzzy sets (inclusion, equality) can be applied to fuzzy relations as well.
- As an example, think of the database table above (or other). If R_{Q_1} and R_{Q_2} are the fuzzy relations corresponding to the results of approximate queries Q_1 ("show candidates with age around 30") and Q_2 ("show candidates with education similar to Computer Science") then intersection $R_{Q_1} \otimes R_{Q_2}$ corresponds to the result of approximate conjunctive query Q_1 AND Q_2.
- In addition to that, we introduce new operations such as compositions of fuzzy relations, inverse fuzzy relations, extensions of fuzzy relations, etc. This is on what we focus on the subsequent slides.

Definition

An inverse relation of a fuzzy relation R between U and V is a fuzzy relation R^{-1} between V and U defined for $u \in U, v \in V$ by

$$R^{-1}(v,u) = R(u,v).$$

Example

Let $X = \{a,b,c\}$, $Y = \{1,2,3,4\}$, $L = [0,1]$, and

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

Then

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

Matrices of R and R^{-1}:

$$\mathbf{M}_R = \begin{pmatrix} 1 & 0.5 & 0 & 0.1 \\ 0 & 0.8 & 0 & 1 \\ 0.8 & 0 & 0 & 0 \end{pmatrix}. \qquad \mathbf{M}_{R^{-1}} = \begin{pmatrix} 1 & 0 & 0.8 \\ 0.5 & 0.8 & 0 \\ 0 & 0 & 0 \\ 0.1 & 1 & 0 \end{pmatrix}.$$

Inverse Fuzzy Relations

The following lemma shows obvious properties of inverse relations. Try to prove it.

Lemma

For $R, R_1, R_2 \in L^{X \times Y}$ *we have*

$$R = (R^{-1})^{-1},$$
$$(\bigcap_i R_i)^{-1} = \bigcap_i R_i^{-1},$$
$$(\bigcup_i R_i)^{-1} = \bigcup_i R_i^{-1},$$
$$(R_1 \otimes R_2)^{-1} = R_1^{-1} \otimes R_2^{-1},$$
$$(R_1 \rightarrow R_2)^{-1} = R_1^{-1} \rightarrow R_2^{-1},$$
$$S(R_1, R_2) = S(R_1^{-1}, R_2^{-1}),$$
$$(R_1 \approx R_2) = (R_1^{-1} \approx R_2^{-1}).$$

Compositions of Fuzzy Relations

David Hume (An Enquiry Concerning Human Understanding, 1758)

From causes which appear similar, we expect similar effects. This is the sum total of all our experimental conclusions.

The basic situation is this: Given an **L**-relation R between X and Y and an **L**-relation S between Y and Z, it might be desirable to obtain from R and S a binary **L**-relation $R * S$ between X and Z. $R * S$ will be called the composition of R and S. In this general setting, a composition of binary **L**-relations between X and Y, and Y and Z is therefore a mapping $* : L^{X \times Y} \times L^{Y \times Z} \to L^{X \times Z}$, i.e. a mapping assigning to any $R \in L^{X \times Y}$ and $S \in L^{Y \times Z}$ their composition $R * S \in L^{X \times Z}$. We are interested in cases where $*$ can be defined by logical formulas, i.e. in cases where the composition can be described verbally.

Compositions of Fuzzy Relations

Let X be a set of patients, Y be a set of symptoms (of diseases), and Z be a set of diseases. Let R be a fuzzy relation between X and Y, S be a fuzzy relation between Y and Z. R may represent results of a medical examination, i.e. for a patient $x \in X$ and a symptom $y \in Y$ (e.g. a headache), $R(x, y)$ is a degree to which x has a headache (notice that typically, R is a noncrisp fuzzy relation). Similarly, S may represent expert knowledge (can be found in medical literature), i.e. for a symptom $y \in Y$ and a disease $z \in Z$, $S(y, z)$ is the truth value to which y is a symptom of z (again, S is a typical fuzzy relation).

Now, can we find out which patients have which diseases? The fuzzy relation in question between X (patients) and Z (diseases) results as a certain composition $R * S$ of R and S, using appropriate composition operation $*$.

Definition (composition of fuzzy relations)

Let R and S be **L**-fuzzy relations between X and Y and between Y and Z. Fuzzy relations $(R \circ S)$, $(R \triangleleft S)$, $(R \triangleright S)$, and $(R \square S)$ between X and Z are defined by

$$\begin{aligned}
(R \circ S)(x,z) &= \bigvee_{y \in Y}(R(x,y) \otimes S(y,z)),\\
(R \triangleleft S)(x,z) &= \bigwedge_{y \in Y}(R(x,y) \rightarrow S(y,z)),\\
(R \triangleright S)(x,z) &= \bigwedge_{y \in Y}(S(y,z) \rightarrow R(x,y)),\\
(R \square S)(x,z) &= \bigwedge_{y \in Y}(R(x,y) \leftrightarrow S(y,z)),
\end{aligned}$$

for all $x \in X$, $z \in Z$.

- $(R \circ S)$, $(R \triangleleft S)$, $(R \triangleright S)$, and $(R \square S)$ are called the $\circ$-, $\triangleleft$-, $\triangleright$-, $\square$-compositions (products) of R and S.

- For $L = \{0,1\}$, $(R \circ S)$ is just the (characteristic function) of the usual composition of R and S. (Show in detail!)
- Describe the meaning of $(R \lhd S)$, $(R \rhd S)$, and $(R \square S)$. (Cf. next point.)
- meaning:
  - $(R \circ S)(x,z)$ = truth degree of "there exists $y \in Y$ such that $\langle x,y \rangle$ is in R and $\langle y,z \rangle$ is in S."
  - $(R \lhd S)(x,z)$ = truth degree of "for each $y \in Y$: if $\langle x,y \rangle$ is in R then $\langle y,z \rangle$ is in S."
  - $(R \rhd S)(x,z)$ = truth degree of "for each $y \in Y$: if $\langle y,z \rangle$ is in S then $\langle x,y \rangle$ is in R."
  - $(R \square S)(x,z)$ = truth degree of "for each $y \in Y$: $\langle x,y \rangle$ is in R iff $\langle y,z \rangle$ is in S."
- In terms of the patient-symptom-disease example, $(R \circ S)(x,z)$ is the degree to which x has a symptom which is characteristic of z, $(R \lhd S)(x,z)$ is the degree to which every symptom of x is characteristic of z, etc.
- One could define other compositions, e.g. $R * S(x,z)$ could express a degree to which most of the symptoms of x are characteristic for z.

Example

Consider crisp relations R between X and Y and S between Y and Z given by the following matrices:

$$\mathbf{M}_R = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \qquad \mathbf{M}_S = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}$$

For matrices $\mathbf{M}_{R \circ S}$, $\mathbf{M}_{R \lhd S}$, $\mathbf{M}_{R \rhd S}$, and $\mathbf{M}_{R \square S}$ representing $R \circ S$, $R \lhd S$, $R \rhd S$, and $R \square S$ we have:

$$\mathbf{M}_{R \circ S} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}, \mathbf{M}_{R \lhd S} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}, \mathbf{M}_{R \rhd S} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}, \mathbf{M}_{R \square S} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Example

Consider fuzzy relations R between X and Y and S between Y and Z given by the following matrices:

$$\mathbf{M}_R = \begin{pmatrix} 0 & 1 & 0.4 \\ 0.4 & 0.2 & 0.2 \\ 0.2 & 0.4 & 0.8 \\ 1 & 1 & 0 \end{pmatrix}, \qquad \mathbf{M}_S = \begin{pmatrix} 1 & 0.1 \\ 1 & 0.9 \\ 0 & 0.2 \end{pmatrix}$$

Consider Gödel (minimum) operations on $[0,1]$. For matrices $\mathbf{M}_{R \circ S}$, $\mathbf{M}_{R \triangleleft S}$, $\mathbf{M}_{R \triangleright S}$, and $\mathbf{M}_{R \square S}$ representing $R \circ S$, $R \triangleleft S$, $R \triangleright S$, and $R \square S$ we have:

$$\mathbf{M}_{R \circ S} = \begin{pmatrix} 1 & 0.9 \\ 0.4 & 0.2 \\ 0.4 & 0.4 \\ 1 & 0.9 \end{pmatrix}, \mathbf{M}_{R \triangleleft S} = \begin{pmatrix} 0 & 0.2 \\ 0 & 0.1 \\ 0 & 0.1 \\ 1 & 0.1 \end{pmatrix}, \mathbf{M}_{R \triangleright S} = \begin{pmatrix} 0 & 0 \\ 0.2 & 0.2 \\ 0.2 & 0.4 \\ 1 & 0 \end{pmatrix},$$

$$\mathbf{M}_{R \square S} = \begin{pmatrix} 0 & 0 \\ 0 & 0.1 \\ 0 & 0.1 \\ 1 & 0 \end{pmatrix}.$$

exercise

- Determine $\mathbf{M}_{R \circ S}$, $\mathbf{M}_{R \triangleleft S}$, $\mathbf{M}_{R \triangleright S}$, and $\mathbf{M}_{R \,\square\, S}$ form the previous example with Łukasiewicz and product structures.
- Notice the similarity in calculation of $\mathbf{M}_{R \circ S}$, $\mathbf{M}_{R \triangleleft S}$, $\mathbf{M}_{R \triangleright S}$, and $\mathbf{M}_{R \,\square\, S}$ compared to
  the matrix multiplication from linear algebra. In linear algebra, we have
  $$(\mathbf{M}_{R \cdot S})_{ij} = \sum_{l \in Y} (\mathbf{M}_R)_{il} \cdot (\mathbf{M}_S)_{lj},$$
  i.e. we compute "row of $\mathbf{M}_R$ times column of $\mathbf{M}_S$". The same happens for $\mathbf{M}_{R \circ S}$, $\mathbf{M}_{R \triangleleft S}$, $\mathbf{M}_{R \triangleright S}$, and $\mathbf{M}_{R \,\square\, S}$, but the operations involved are different. For example, in $\mathbf{M}_{R \circ S}$, $\sum$ is replaced by $\bigvee$ and $\cdot$ is replaced by $\otimes$.
- Define fuzzy relations R between patients and symptoms, and S between symptoms and diseases. Determine $R \circ S$, $R \triangleleft S$, $R \triangleright S$, and $R \,\square\, S$

Properties of Compositions of Fuzzy Relations

We list selected properties. Proofs and more properties can be found in Chapter 6 of Belohlavek R.: Fuzzy Relational Systems: Foundations And Principles. Kluwer, New York, 2002. [available at `http://bingweb.binghamton.edu/~rbelohla/Chap6-Belohlavek-FuzzyRelationalSystems.pdf`]

Theorem (products and similarity)

$$
\begin{aligned}
(R_1 \approx R_2) \otimes (S_1 \approx S_2) &\leq (R_1 \circ S_1) \approx (R_2 \circ S_2),\\
(R_1 \approx R_2) \otimes (S_1 \approx S_2) &\leq (R_1 \triangleleft S_1) \approx (R_2 \triangleleft S_2),\\
(R_1 \approx R_2) \otimes (S_1 \approx S_2) &\leq (R_1 \triangleright S_1) \approx (R_2 \triangleright S_2),\\
(R_1 \approx R_2) \otimes (S_1 \approx S_2) &\leq (R_1 \,\square\, S_1) \approx (R_2 \,\square\, S_2).
\end{aligned}
$$

Theorem (products and inverse relations)

$$\begin{aligned}(R \circ S)^{-1} &= S^{-1} \circ R^{-1},\\ (R \triangleleft S)^{-1} &= S^{-1} \triangleright R^{-1},\\ (R \triangleright S)^{-1} &= S^{-1} \triangleleft R^{-1},\\ (R \square S)^{-1} &= S^{-1} \square R^{-1}.\end{aligned}$$

Theorem (products and associativity)

$$R \circ (S \circ T) = (R \circ S) \circ T,$$
$$R \triangleleft (S \triangleright T) = (R \triangleleft S) \triangleright T,$$
$$R \triangleleft (S \triangleleft T) = (R \circ S) \triangleleft T,$$
$$(R \triangleright S) \triangleright T = R \triangleright (S \circ T).$$

Theorem (products and distributivity)

$$(\bigcap_i R_i) \circ S \subseteq \bigcap_i (R_i \circ S), \quad R \circ (\bigcap_i S_i) \subseteq \bigcap_i (R \circ S_i),$$

$$(\bigcup_i R_i) \circ S = \bigcup_i (R_i \circ S), \quad R \circ (\bigcup_i S_i) = \bigcup_i (R \circ S_i),$$

$$\bigcup_i (R_i \lhd S) \subseteq (\bigcap_i R_i) \lhd S, \quad R \lhd (\bigcap_i S_i) = \bigcap_i (R \lhd S_i),$$

$$(\bigcup_i R_i) \lhd S = \bigcap_i (R_i \lhd S), \quad \bigcup_i (R \lhd S_i) \subseteq R \lhd (\bigcup_i S_i),$$

$$(\bigcap_i R_i) \rhd S = \bigcup_i (R_i \rhd S), \quad \bigcup_i (R \rhd S_i) \subseteq R \rhd (\bigcap_i S_i),$$

$$\bigcap_i (R_i \rhd S) \subseteq (\bigcup_i R_i) \rhd S, \quad R \rhd (\bigcup_i S_i) = \bigcap_i (R \rhd S_i).$$

Binary Fuzzy Relations on a Set

Pierre Duhem (The Aim and Structure of Physical Theory, 1954)

It is impossible to describe a practical fact without attenuating by the use of the word "approximately" or "nearly"; on the other hand, all the elements constituting the theoretical fact are defined with rigorous exactness.

Bertrand Russell (The Philosophy of Logical Atomism, 1918)

Everything is vague to a degree you do not realize till you have tried to make it precise.

We introduce reflexivity, symmetry, transitivity, and assymmetry of fuzzy relations only (most frequently used).

Definition (basic properties)

A binary fuzzy relation R on a set U is called reflexive, symmetric, transitive (w.r.t. $\otimes$), antisymmetric (w.r.t. fuzzy equality $x \approx^U y$ on U, see later), if it satisfies

$$R(x,x) = 1,$$
$$R(x,y) = R(y,x),$$
$$R(x,y) \otimes R(y,z) \leq R(x,z),$$
$$R(x,y) \wedge R(y,x) \leq (x \approx^X y),$$

respectively.

Note: Transitivity depends on what truth function $\otimes$ of conjunction we use.

- One could consider the properties in degrees. For example, a degree $\mathrm{Ref}(R)$ to which R is reflexive can be defined by $\mathrm{Ref}(R) = \bigwedge_{x \in U} R(x, x)$. Then, R is reflexive iff $\mathrm{Ref}(R) = 1$ (verify).
- The definition presents a generalization of the ordinary properties of relations. That is, for $L = \{0, 1\}$, the properties defined above coincide with the ordinary properties. (Verify.)
- More information on fuzzy relations in a set can be found in the literature, e.g. Klir, Yuan: Fuzzy Sets and Fuzzy Logic. Prentice Hall, 1995, or Belohlavek: Fuzzy Relational Systems. Kluwer, 2002.

Example

Consider fuzzy relations R and S on $U = \{u, v, w, x\}$ given by matrices

$$\mathbf{M}_R = \begin{pmatrix} 1 & 1 & 0.4 & 0 \\ 1 & 1 & 0.2 & 0 \\ 0.4 & 0.2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{M}_S = \begin{pmatrix} 1 & 0.8 & 0.4 & 0 \\ 0.8 & 1 & 0.2 & 0 \\ 0.4 & 0.2 & 1 & 0.2 \\ 0 & 0 & 0.1 & 0.9 \end{pmatrix}.$$

Determine whether R and S are reflexive, symmetric, transitive (w.r.t. Łukasiewicz t-norm).

R: reflexivity yes (1s on main diagonal), symmetry yes (symmetric along main diagonal), transitivity no:
$R(v, u) \otimes R(u, w) = 1 \otimes 0.4 = 0.4 \not\leq 0.2 = R(v, w)$.

S: reflexivity no ($R(x, x) = 0.9$), symmetry no ($R(w, x) = 0.2 \neq 0.1 = R(x, w)$), transitivity yes (check). But S is "almost reflexive" and "almost symmetric".

Definition (fuzzy equivalence)

A binary relation $\approx$ on a set U is called a fuzzy equivalence if it is reflexive, symmetric, transitive.
If, moreover, $(u \approx v) = 1$ implies $u = v$, $\approx$ is called a fuzzy equality.

More information about fuzzy equivalences canbe found in pp. 37–42 of [CH1] (Chapter 1 of Belohlavek, Vychodil: Fuzzy Equational Logic. Springer, 2005), available at
`http://bingweb.binghamton.edu/~rbelohla/ssie517f2007.html`.
In particular,

- Definition 1.69 and Remark 1.70, Example 1.71,
- Theorem 1.72,
- Fuzzy equivalences and partitions: Definition of an L-equivalence class, Definition 1.73, Theorem 1.75,
- Inducing a fuzzy equivalence $\theta_{\mathcal{S}}$ in U from a system $\mathcal{S}$ of fuzzy sets in U, Theorem 1.76, Remark 1.77.

Example (fuzzy equivalence and partition)

θ	u_1	u_2	u_3	u_4
u_1	1	0.9	0.8	1
u_2	0.9	1	0.7	0.9
u_3	0.8	0.7	1	0.8
u_4	1	0.9	0.8	1

describes an **L**-equivalence on $U = \{u_1, \dots, u_4\}$ with $L = [0, 1]$ and Łukasiewicz operations on L (verify).
Classes of θ: $[u_1]_\theta = [u_1]_\theta = \{^1/u_1, ^{0.9}/u_2, ^{0.8}/u_3, ^1/u_4\}$, $[u_2]_\theta = \{^{0.9}/u_1, ^1/u_2, ^{0.7}/u_3, ^{0.9}/u_4\}$, $[u_3]_\theta = \{^{0.8}/u_1, ^{0.7}/u_2, ^1/u_3, ^{0.8}/u_4\}$.
Therefore, $\{\{^1/u_1, ^{0.9}/u_2, ^{0.8}/u_3, ^1/u_4\}, \{^{0.9}/u_1, ^1/u_2, ^{0.7}/u_3, ^{0.9}/u_4\}, \{^{0.8}/u_1, ^{0.7}/u_2, ^1/u_3, ^{0.8}/u_4\}\}$ is an **L**-partition Π_θ on U induced by θ.
Note that an **L**-partition on U represents a system of similarity-based clusters (fuzzy sets) of elements of U.

Example (Leibniz equivalence)

Consider set U consisting of Porsche 911, ..., Honda Accord, and a set $\mathcal{S}$ of fuzzy sets (attributes) AWD, expensive, and high MPG. Let

car	AWD	expensive	high MPG
Porsche 911	1	1	0.2
Toyota RAV4	1	0.7	0.7
Toyota Corolla	0	0.4	1
Subaru Outback	1	0.8	0.6
Honda Civic	0	0.3	1
Honda Accord	0	0.5	0.8

Then with Łukasiewicz operations, the Leibniz similarity $\approx$ is given by

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

Example (Leibniz equivalence)

Rearranging the table yields:

	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

The concept of fuzzy equivalence illustrates the two sides of fuzzy logic: symbolic (qualitative) one and numerical (quantitative) one. For example, transitivity of fuzzy equivalence $\approx$ says:

- "If u and v are similar and v and w are similar then u and w are similar" which is the formula behind transitivity (symbolic, describes meaning in natural lagnuage).
- $(u \approx v) \otimes (v \approx w) \leq (u \approx w)$ (numerical, makes the symbolic/verbal description more precise).

Extension Principle for Fuzzy Systems

- Proposed by Zadeh.
- Represents one way of "'fuzzification" of systems.
- Typical scenario: A given system performs a function $f : X \to Y$. Given input $x \in X$, $f(x) \in Y$ is the corresponding output. What if inputs are not known precisely? What if, instead of $x \in X$, we only know that the input is "approximately x" etc.? Such inputs can be described by fuzzy sets in X. How do we apply f to inputs which are fuzzy sets?
  Two particular cases:
  - Interval computations (e.g., Vladik Kreinovich): When there is imprecision in measurement, it is more realistic to represent the measured quantities by intervals rather than by numbers. One advantage is better control of how error spreads, i.e. more robust computations. Note: intervals are particular cases of fuzzy sets (crisp fuzzy sets).
  - Fuzzy neural networks.

Definition (extension principle)

Let $f : X \to Y$ be a function. By extension principle, f induces a function $\overline{f} : L^X \to L^Y$ defined for $A \in L^X$ by

$$(\overline{f}(A))(y) = \bigvee \{A(x) \,|\, x \in X, f(x) = y\}.$$

- For simplicity, one uses just $\overline{f}(A)$ instead of $f(A)$.
- Input A and output $\overline{f}(A)$ are fuzzy sets. $f(A)(y)$ can be seen as a degree to which there exists x in A which is mapped to y by f.

Definition (extension principle, multiple inputs)

Let $f : X_1 \times \cdots \times X_n \to Y$ be a function. By extension principle, f induces a function $\overline{f} : L^{X_1} \times \cdots \times L^{X_n} \to L^Y$ defined for $A_1 \in L^{X_1}, \ldots, A_n \in L^{X_n}$ by

$$(\overline{f}(A_1, \ldots, A_n))(y) = \bigvee \{A_1(x_1) \wedge \cdots \wedge A_n(x_n) \,|\, x_1 \in X_1, \ldots, x_n \in X_n, f(x_1, \ldots, x_n) = y\}.$$

Example (extension principle)

(1) Let $X = \{1, 2, 3, 4, 5\}$, $Y = \{a, b, c, d\}$, f be defined by $f(1) = a$, $f(2) = b$, $f(3) = b$, $f(4) = d$, $f(5) = d$. Determine $f(A)$.

- $A = \{{}^{0.5}/1, {}^{1}/2, {}^{0.5}/3\}$. $f(A) = \{{}^{0.5}/a, {}^{1}/b\}$.
- $A = \{{}^{0.5}/1, {}^{0.4}/4, {}^{0.6}/5\}$. $f(A) = \{{}^{0.5}/a, {}^{0.6}/d\}$.
- $A = \emptyset$. $f(A) = \emptyset$.
- $A = X$. $f(A) = \{{}^{1}/a, {}^{1}/b, {}^{1}/d\}$.

(2) Let $X = Y = \mathbb{R}$, $f(x) = 3x$,
$A = \{{}^{0.25}/1, {}^{0.5}/2, {}^{0.75}/3, {}^{1}/4, {}^{0.75}/5, {}^{0.5}/6, {}^{0.25}/7\}$.
$f(A) = \{{}^{0.25}/3, {}^{0.5}/6, {}^{0.75}/9, {}^{1}/12, {}^{0.75}/15, {}^{0.5}/18, {}^{0.25}/21\}$.
(3) Let $X = Y = \mathbb{R}$, $f(x) = 3x$, A be a triangular fuzzy set given by parameters a, b, c. Then $f(A)$ is a triangular fuzzy set given by parameters $3a, 3b, 3c$.

For $f : X_1 \times \cdots \times X_n \to Y$, computing $\overline{f}(A_1, \dots, A_n)$ for $A_1 \in L^{X_1}, \cdots, A_n \in L^{X_n}$ is demanding (see definition).

For the particular case when $X_i = \mathbb{R}$, A_i are intervals (note that we can consider intervals as particular crisp fuzzy sets), and f are basic arithmetic operations such as addition, subtraction, multiplication, extension principle yields well-known formulas of interval arithmetic (verify them using the definition of EP):

- $[a, b] + [c, d] = [a + c, b + d]$,
- $[a, b] - [c, d] = [a - d, b - c]$,
- $[a, b] \cdot [c, d] =$
$$\begin{cases} [ac, bd] & \text{for } a \geq 0, c \geq 0, \\ [bd, ac] & \text{for } b < 0, d < 0, \\ [\min\{ad, bc\}, \max\{ad, bc\}] & \text{for } ab \geq 0, cd \geq 0, ac < 0, \\ [\min\{ad, bc\}, \max\{ac, bd\}] & \text{for } ab < 0, cd < 0. \end{cases}$$

Link to EP: Let f be addition of real numbers, $A_{[a,b]}$ be a crisp fuzzy set corresponding to interval $[a, b]$, i.e. $A_{[a,b]}(x) = 1$ if $x \in [a, b]$, $A_{[a,b]}(x) = 0$ if $x \notin [a, b]$; same for $A_{[c,d]}$. Then one can see that using EP, $A_{[a,b]} \overline{+} A_{[c,d]} = A_{[a+c,b+d]}$. The same holds true for subtraction and multiplication.

For division, the situation is technically bit complicated. One needs to distinguish cases taking care of possible division by 0.

Question: Consider $L = [0, 1]$, $X_1 = \cdots = X_n = Y = \mathbb{R}$. Can $\overline{f}(A_1, \ldots, A_n)$ be computed efficiently?

Yes, in special cases: If f is continuous and if a-cuts of $A_1, \ldots, A_n$ are compact subsets of $\mathbb{R}$ for $a > 0$.
Note: Compactness is a topological notion. Particular cases (the ones important for applications) of compact subsets of $\mathbb{R}$ are closed intervals or, more generally, unions of closed intervals.

Theorem (reducing extension principle to interval computation)

Let $f : \mathbb{R}^n \to \mathbb{R}$ *be a continuous function, let* $A_1, \ldots, A_n \in [0,1]^{\mathbb{R}}$ *be such that for every* $b \in (0,1]$, ${}^bA_i$ *is a compact subset of* $\mathbb{R}$ *(e.g. a closed interval). Then for every* $a \in (0,1]$*:*

$$ {}^af(A_1, \ldots, A_n) = f({}^aA_1, \ldots, {}^aA_n). $$

- Note that, as an example, if A_i has a continuous membership function (such as triangular or trapezoidal fuzzy sets), then ${}^bA_i$ is compact.
- Therefore, under conditions of the theorem, we can compute a-cuts of the output from a-cuts of the input.
- Examples (assume A_i are trapezoidal fuzzy sets):
  - ${}^a(A_1 + A_2) = {}^aA_1 + {}^aA_2$, where on the right hand side, we use interval arithmetic,
  - ${}^a(A_1 - A_2) = {}^aA_1 - {}^aA_2$,
  - ${}^a(A_1 \cdot A_2) = {}^aA_1 \cdot {}^aA_2$,
  - ${}^a(c \cdot A_2) = c \cdot {}^aA_1$,
  - ${}^a(e^{A_2}) = e^{{}^aA_1}$.

Example

Determine a-cuts of fuzzy set C, for $a = 1, 0.5, 0.1$.

- $C = A + B$, where $A = \mathrm{Tri}(b_1, c_1, d_1)$ and $B = \mathrm{Tri}(b_2, c_2, d_2)$ are triangular fuzzy sets given by parameters b_1, c_1, d_1 and b_2, c_2, d_2:
  - $A = \mathrm{Tri}(1, 2, 3)$ and $B = \mathrm{Tri}(2, 4, 6)$.
  - $A = \mathrm{Tri}(0, 2, 4)$ and $B = \mathrm{Tri}(-7, -6, -5)$.
  - $A = \mathrm{Tri}(0, 2, 4)$ and $B = 6$, i.e. $B = \mathrm{Tri}(6, 6, 6)$.
  - $A = 5$, i.e. $A = \mathrm{Tri}(5, 5, 5)$, and $B = 6$, i.e. $B = \mathrm{Tri}(6, 6, 6)$.
- $C = A - B$, where $A = \mathrm{Tri}(b_1, c_1, d_1)$ and $B = \mathrm{Tri}(b_2, c_2, d_2)$ are triangular fuzzy sets given by parameters b_1, c_1, d_1 and b_2, c_2, d_2:
  - $A = \mathrm{Tri}(1, 2, 3)$ and $B = \mathrm{Tri}(2, 4, 6)$.
  - $A = \mathrm{Tri}(0, 2, 4)$ and $B = \mathrm{Tri}(-7, -6, -5)$.
  - $A = \mathrm{Tri}(0, 2, 4)$ and $B = 6$, i.e. $B = \mathrm{Tri}(6, 6, 6)$.
  - $A = 5$, i.e. $A = \mathrm{Tri}(5, 5, 5)$, and $B = 6$, i.e. $B = \mathrm{Tri}(6, 6, 6)$.
- $C = A \cdot B$, where $A = \mathrm{Tri}(b_1, c_1, d_1)$ and $B = \mathrm{Tri}(b_2, c_2, d_2)$ are triangular fuzzy sets given by parameters b_1, c_1, d_1 and b_2, c_2, d_2:
  - $A = \mathrm{Tri}(1, 2, 3)$ and $B = \mathrm{Tri}(2, 4, 6)$.
  - $A = \mathrm{Tri}(0, 2, 4)$ and $B = \mathrm{Tri}(-7, -6, -5)$.
  - $A = \mathrm{Tri}(0, 2, 4)$ and $B = 6$, i.e. $B = \mathrm{Tri}(6, 6, 6)$.

Example (results for previous example)

- $C = A + B$:
  - ${}^{1}(C) = [6, 6] = \{6\}$, ${}^{0.5}(C) = [4.5, 7.5]$, ${}^{0.1}(C) = [3.3, 8.7]$.
  - ${}^{1}(C) = [-4, -4]$, ${}^{0.5}(C) = [-5.5, -1.5]$, ${}^{0.1}(C) = [-6.7, -1.3]$.
  - ${}^{1}(C) = [8, 8]$, ${}^{0.5}(C) = [7, 9]$, ${}^{0.1}(C) = [6.2, 9.8]$.
  - ${}^{1}(C) = [11, 11] = \{11\}$, ${}^{0.5}(C) = [11, 11] = \{11\}$, ${}^{0.1}(C) = [11, 11] = \{11\}$.
- $C = A - B$:
  - ${}^{1}(C) = [-2, -2]$, ${}^{0.5}(C) = [-2.5, -1.5]$, ${}^{0.1}(C) = [-2.9, -1.1]$.
  - ${}^{1}(C) = [8, 8]$, ${}^{0.5}(C) = [6.5, 9.5]$, ${}^{0.1}(C) = [5.3, 10.7]$.
  - ${}^{1}(C) = [-4, -4]$, ${}^{0.5}(C) = [-5, -3]$, ${}^{0.1}(C) = [-5.8, -2.2]$.
  - ${}^{1}(C) = [-1, -1]$, ${}^{0.5}(C) = [-1, -1]$, ${}^{0.1}(C) = [-1, -1]$.
- $C = A \cdot B$:
  - ${}^{1}(C) = [8, 8]$, ${}^{0.5}(C) = [4.5, 12.5]$, ${}^{0.1}(C) = [1.1 \cdot 2.2, 2.9 \cdot 5.8]$.
  - ${}^{1}(C) = [-12, -12]$, ${}^{0.5}(C) = [-19.5, -5.5]$, ${}^{0.1}(C) = [3.8 \cdot -6.9, 0.2 \cdot -5.1]$.
  - ${}^{1}(C) = [12, 12]$, ${}^{0.5}(C) = [6, 18]$, ${}^{0.1}(C) = [0.2 \cdot 6, 3.8 \cdot 6]$.

Example (neural networks with imprecise inputs)

Consider function

$$f(x_1, x_2) = \frac{1}{1 + e^{-(w_1 \cdot x_1 + w_2 \cdot x_2 - \theta)}}.$$

Such function occurs in multi-layer neural networks. $x_1, x_2 \in \mathbb{R}$ are inputs to a neuron, $w_1, w_2 \in \mathbb{R}$ are weights of connections from the inputs to the neuron, $\theta \in \mathbb{R}$ is a threshold, $\frac{1}{1+e^{-z}}$ is a transfer function of a neuron (sigmoid function). $f(x_1, x_2)$ is then the neuron output given x_1, x_2 as the inputs. Neurons are organized and connected in layers, outputs of neurons serve as inputs of other neurons. This is, basically, the architecture of back-propagation type neural networks.

Determine the output $f(x_1, x_2)$ of a neuron given that $w_1 = 3$, $w_2 = 2$, $\theta = -2$, $x_1 = A_1$ and $x_2 = A_2$ are triangular fuzzy sets given by parameters $1, 2, 3$ (for A_1) and $1, 3, 5$ (for A_2).

Example (neural networks with imprecise inputs, cntd.)

For function

$$f(x_1, x_2) = \frac{1}{1 + e^{-(w_1 \cdot x_1 + w_2 \cdot x_2 - \theta)}},$$

determine $f(A_1, A_2)$ given that $w_1 = 3$, $w_2 = 2$, $\theta = -2$, A_1 and A_2 are triangular fuzzy sets given by parameters $1, 2, 3$ (for A_1) and $1, 3, 5$ (for A_2).

Fuzzy set $f(A_1, A_2)$ is determined by its a-cuts ${}^af(A_1, A_2)$. The usual way to compute ${}^af(A_1, A_2)$ is to consider $f(x_1, x_2)$ a composite function and to apply interval arithmetic to the functions from which f is composed.

Example for $a = 0.5$:

1. $w_1 \cdot {}^{0.5}A_1 = 3 \cdot [1.5, 2.5] = [4.5, 7.5]$, $w_2 \cdot {}^{0.5}A_2 = 2 \cdot [2, 4] = [4, 8]$,
2. $w_1 \cdot {}^{0.5}A_1 + w_2 \cdot {}^{0.5}A_2 - \theta = -[8.5, 15.5] + 2 = [10.5, 17.5]$.
3. $\frac{1}{1+e^{-(w_1 \cdot {}^{0.5}A_1 + w_2 \cdot {}^{0.5}A_1 - \theta)}} = \frac{1}{1+e^{-[10.5,17.5]}} = [0.99997246, 0.99999997489]$.

Further Applications of Extension Principle

– Extension principle is used in various applications of fuzzy sets in which computation with numbers is replaced by computation with fuzzy sets which represent imprecisely known numbers.
– Such imprecisely known numbers are called fuzzy quantities, or fuzzy numbers or fuzzy intervals.
– Fuzzy arithmetic is implemented e.g. in Mathematica.
– Particular areas and literature:
  – Linear programming with inexact data: Fiedler M.: Linear Optimization Problems with Inexact Data, Springer, 2006.
  – Engineering applications are covered in Hanss M.: Applied Fuzzy Arithmetic, An Introduction with Engineering Applications.
  – Decision making under uncertainty, many papers and books exist.

Distance of Fuzzy Sets in $\mathbb{R}$

When using EP for to extend functions f defined for real numbers to functions $\overline{f}$ defined for fuzzy sets in $\mathbb{R}$, one often needs to assess the distance of two fuzzy sets A and B in $\mathbb{R}$.
The following function, called the generalized Hausdorff pseudo-metric, is often used:

$$d_\infty(A, B) = \sup_{a \in K} d_H({}^aA, {}^aB)$$

where d_H is the well-known Hausdorff pseudometric defined by

$$d_H({}^aA, {}^aB) = \max\{\sup_{x \in {}^aA} \inf_{y \in {}^aB} |x - y|, \sup_{y \in {}^aB} \inf_{x \in {}^aA} |x - y|.\}$$

Usually, K is some subset of $(0, 1]$ such as $K = (0, 1]$, $K = (0.1, 0.2, \dots, 0.9, 1]$.

Distance of Fuzzy Sets in $\mathbb{R}$

Example

Determine $d_\infty(A, B)$ for a triangular fuzzy set A given by parameters $1, 3, 5$ and a trapezoidal fuzzy set B given by parameters $2, 4, 5, 6$. Consider $K = 0.5, 1$. We have

$d_H({}^1A, {}^1B) = d_H([3, 3], [4, 5]) = \max\{1, 2\} = 2,$

$d_H({}^{0.5}A, {}^{0.5}B) = d_H([2, 4], [3, 5.5]) = \max\{1, 1.5\} = 1.5,$

hence $d_\infty(A, B) = \sup\{2, 1.5\} = 2.$

Compositional Rule of Inference (CRI)

CRI is a technique used in so-called fuzzy rule-based systems, the basic components of fuzzy controllers.

Definition (CRI)

Let R be an **L**-relation between sets X and Y, A be an **L**-set in X. The **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)".
- In fact, $A \circ R$ can be seen as a $\circ$-composition of fuzzy relations: Let $W = \{1\}$ and define a fuzzy relation Q between W and X by $Q(1, x) = A(x)$ for every $x \in X$. Then, for $B = A \circ R$ we have
$$B(y) = (Q \circ R)(1, y)$$
for every $y \in Y$.
- 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 to 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:

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

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

Fuzzy Rule-Based Systems and Fuzzy Control

Material:

Handouts (available at `http://bingweb.binghamton.edu/~rbelohla/FuzzyRuleBasedSystems.pdf`).

Chapter 11.4 (up to the end of p. 319), Chapters 12.1, 12.2, 12.3 of Klir, Yuan: Fuzzy Sets and Fuzzy Logic. Theory and Applications. Prentice Hall, 1995 (available at `http://bingweb.binghamton.edu/~rbelohla/KlirYuan1995.PDF`).

Fuzzy Rule-Based Systems and Fuzzy Control

Fuzzification and defuzzification

Chapter 12 of Klir, Yuan: Fuzzy Sets and Fuzzy Logic. Theory and Applications. Prentice Hall, 1995 (available at `http://bingweb.binghamton.edu/~rbelohla/KlirYuan1995.PDF`).

- Provide interface to fuzzy rule-based systems.
- Inputs, particularly in fuzzy control, are numbers in most cases (temperature from sensors, etc.).
- Outputs in fuzzy control need to be numbers (action, e.g. r.p.m. of a ventilator).

Fuzzification = transformation of inputs which are elements of the input space to inputs which are fuzzy sets in the input space. Note: Inputs to fuzzy rule-based systems need to be fuzzy sets in the input space.

Case of one input variable x_1:

- If the value of x_1 is a_1, the input fuzzy set A'_1 is set to
  - singleton $A'_1 = \{^1/a_1\}$, or
  - "narrow fuzzy set" with core $\{a_1\}$, e.g. a triangular fuzzy set with parameters $a_1 - \varepsilon, a_1, a_1 + \varepsilon$, for small ε

Case of n input variables $x_1, \dots, x_n$:

- If the values of $x_1, \dots, x_n$ are $a_1, \dots, a_n$, the input fuzzy sets $A'_1, \dots, A'_n$ are set to
  - singleton $A'_1 = \{^1/a_1\}, \dots, A'_n = \{^1/a_n\}$ or
  - "narrow fuzzy sets" with cores $\{a_i\}$, e.g. triangular fuzzy sets with parameters $a_i - \varepsilon, a_i, a_i + \varepsilon$.

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. $Y = \{y_1, \ldots, y_k\}$. A defuzzification method can be seen as a function $D : [0,1]^Y \to Y$ 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
$$D(B) = \frac{\int_a^b B(y) y dy}{\int_a^b B(y) dy},$$
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 $Y = \{0, 1, \ldots, 10\}$. 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: $M = \{2, 3, 4, 9\}$,
$$D(B) = \frac{2+3+4+9}{4} = \frac{18}{4} = 4.5.$$

Fuzzy Rule-Based Systems: Example

Consider $X = Y = \{0, 1, \ldots, 10\}$ 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: $R_i(x, y) = A_i(x) \wedge B_i(y)$, $R = R_1 \cup R_2 \cup R_3$.

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

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

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

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

Determine $A \circ R$ for $A_1 = \{{}^{1}/5\}$, $A_2 = \{{}^{0.5}/3, {}^{1}/4\}$.
$B_1' = A_1' \circ R = \{{}^{0.5}/6, {}^{0.75}/7, {}^{0.5}/8\}$
$B_2' = A_2' \circ R = \{{}^{0.5}/1, {}^{0.5}/2, {}^{0.5}/3, {}^{0.5}/4, {}^{0.5}/5, {}^{0.5}/6, {}^{0.1}/7, {}^{0.5}/8\}$
What are the corresponding defuzzified value for A_1' (using COG)?
$D(B_1') = \frac{0.5 \cdot 6 + 0.75 \cdot 7 + 0.5 \cdot 8}{0.5 + 0.75 + 0.5} = \frac{12.25}{1.75} = 7.$

Determine the output fuzzy sets and their defuzzified values (with COG) for all singleton inputs from X. That is, for $x \in X$, determine $B' = A' \circ R$ for $A' = \{^1/x\}$ and determine $D(B')$ (defuzzification of B').

inp. x	output fuzzy set $\{^1/x\} \circ R$	defuzzified value $D(\{^1/x\} \circ R)$
10	$\emptyset$ (empty fuzzy set)	Not Defined
9	$\emptyset$ (empty fuzzy set)	Not Defined
8	$\{^{0.5}/8, {}^{0.5}/9, {}^{0.5}/10\}$	9
7	$\{^{0.5}/8, {}^{1}/9, {}^{0.5}/10\}$	9
6	$\{^{0.25}/6, {}^{0.25}/7, {}^{0.5}/8, {}^{0.5}/9, {}^{0.5}/10\}$	8.375
5	$\{^{0.5}/6, {}^{0.75}/7, {}^{0.5}/8\}$	7
4	$\{^{0.5}/6, {}^{1}/7, {}^{0.5}/8\}$	7
3	$\{^{0.5}/1, {}^{0.5}/2, {}^{0.5}/3, {}^{0.5}/4, {}^{0.5}/5, {}^{0.5}/6, {}^{0.75}/7, {}^{0.5}/8\}$	4.65
2	$\{^{0.5}/1, {}^{0.75}/2, {}^{1}/3, {}^{1}/4, {}^{0.75}/5, {}^{0.5}/6, {}^{0.25}/7, {}^{0.25}/8\}$	3.9
1	$\{^{0.5}/1, {}^{0.5}/2, {}^{0.5}/3, {}^{0.5}/4, {}^{0.5}/5, {}^{0.5}/6\}$	3.5
0	$\emptyset$ (empty fuzzy set)	Not Defined

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:

10											
9								×	×	×	
8							×				
7					×	×					
6											
5				×							
4		×	×								
3		×									
2											
1											
0											
R	0	1	2	3	4	5	6	7	8	9	10

Computing the output fuzzy sets for singleton inputs from X using the "geometric method", i.e. determining degrees to which rule fire, cutting the output fuzzy sets and making union of the cut fuzzy sets.

That is, for input $x \in X$:

- determine $a_i = A_i(x)$ (degree to which i-th rule fires),
- determine $B'_i = a_i \wedge B_i$ (i.e., $B'_i(y) = a_i \wedge B_i(y)$),
- determine the output: $B' = B'_1 \cup B'_2 \cup B'_3$.

Example: input $x = 3$:
$a_1 = A_1(3) = 0.5$, $a_2 = A_2(3) = 0.75$, $a_3 = A_3(3) = 0$,
$B'_1 = a_1 \wedge B_1 = \{{}^{0.5}/1, {}^{0.5}/2, {}^{0.5}/3, {}^{0.5}/4, {}^{0.5}/5, {}^{0.5}/6\}$,
$B'_2 = a_2 \wedge B_2 = \{{}^{0.5}/6, {}^{0.75}/7, {}^{0.5}/8\}$,
$B'_3 = a_3 \wedge B_2 = \emptyset$,
$B' = B'_1 \cup B'_2 \cup B'_3 = \{{}^{0.5}/1, {}^{0.5}/2, {}^{0.5}/3, {}^{0.5}/4, {}^{0.5}/5, {}^{0.5}/6\}, {}^{0.75}/7, {}^{0.5}/8\}$,
which is the same fuzzy set as the one computed by projection of ${}^{1}/x$ via R (see the table with singleton inputs and fuzzy set outputs).

Universal Mapping Property of Fuzzy Rule-Based Systems

Let $X_1, \ldots, X_n$ (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 $\circ$ (such as COG), there is an associated function $G : X_1 \times \cdots \times X_n \to Y$ defined by

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

That is, for $x_1, \ldots, x_n$, we take their fuzzifications $A'_1 = F(x_1), \ldots,$ $A'_n = F(x_n)$ (e.g. singletons), use the rule base and the inference method to compute the output fuzzy set $B' = \langle F(x_1), \ldots, F(x_n)\rangle \circ R$ and set $G(x_1, \ldots, x_n)$ 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 an prove the following theorem (we omit details):

Theorem (UMP of fuzzy rule-based systems)

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

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

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.

Fuzzy Logic as Logic

- overview of logical aspects of fuzzy logic
- fuzzy logic in broad sense vs. in narrow sense
  FL in broad sense: principles and methods developed in fuzzy set theory
  FL in narrow sense: mathematical logic (studies notions such as axioms, provability, entailment) which allows a partially ordered scale of truth degrees
- start with overview of formal treatment of classical propositional logic
  `http://bingweb.binghamton.edu/~rbelohla/logic.pdf`
- then overview of two approaches to fuzzy logic in narrow sense (slides 18–24 from `http://bingweb.binghamton.edu/~rbelohla/CLA2006tutorialBelohlavekI.pdf`)
- case study in detail: attribute implications in fuzzy setting (slides 1–20 from `http://bingweb.binghamton.edu/~rbelohla/CLA2006tutorialBelohlavekIII.pdf`)