Fuzzy Logic

Material used

- Script *Fuzzy Logic* by Michiel van Lambalgen

1 Introduction

- Aristotle was the first to realize that logic based on True or False alone was not sufficient.
- The mathematics of fuzzy set theory and fuzzy logic.

Proposed in 1965 by L.A. Zadeh
(Fuzzy Sets, Information Control, vol. 8, pp. 338–353)

- generalization of ordinary set theory
- '70 first applications, fuzzy control (Mamdani)
- '80 industrial applications, train operation, pattern recognition
- '90 consumer products, cars, special HW, SW.

General Motivation

- Fuzzy logic handles the concept of partial true, that is true values between “completely true” and “completely false”.
- Janet is 65 years old. Is Janet old? In Boolean logic (True or False). In fuzzy logic (False, True or degree of oldness). Many events or facts have such fuzzy truth values.
- Other Fuzzy Examples
  - How big does a pond have to be to qualify as a lake?
  - How much of an apple do you have to eat for what is left to no longer count as an apple?
  - How broken has a ship to be in order to be called a wreck?
  - What amount of hair loss categorizes you as bald?

Sorites Paradox

$\neg$Bald(1000 000),

$\neg$Bald(n) $\rightarrow$ $\neg$Bald(n-1)

$\therefore$ Bald(0)

or, correspondingly

Bald(0),

Bald(n) $\rightarrow$ Bald(n+1)

$\therefore$ Bald(1 000 000)

Formal aspects of uncertainty and vagueness

(A) Is uncertainty truth functional?

(B) Is vagueness truth-functional?

	Truth functional	Not truth functional
Uncertainty	Possibility measures Certainty factors (e.g. MYCIN)	Probability $P(A \cap B) = f(P(A), P(B))$?
Vagueness	Fuzzy logic Super-valuations	??

Two answers

Question A

Let be Deg(p) my degree of uncertainty of p

Assume Deg(p) = Deg(¬p) for some p (e.g. Susi is pregnant)

If Deg is truth-functional, there is a fixed function F:

Deg(p&p) = F(Deg(p), Deg(p)) = F(Deg(p), Deg(¬p)) = Deg(p&¬p).

Question B

Let be Deg(p) my degree of vagueness of p

Assume Deg(p) = Deg(¬p) for some p (e.g. Peter_{35} is old).

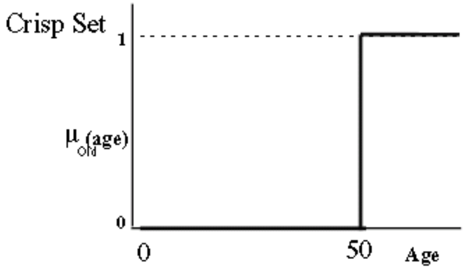
Again: Deg(p&p) = Deg(p&¬p). Better than before??

Four basic questions

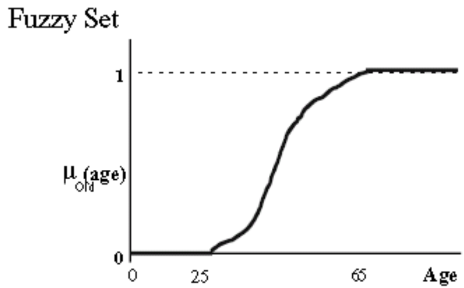
1. How should vagueness of atomic predicates be represented mathematically?
2. How does vagueness of formulas combine under logical operations?
3. How can we determine degrees of vagueness empirically?
4. What can one *do* in practice with degrees of vagueness?

Fuzzy sets and Crisp sets

more than 50 years old
denotes a crisp set
standard set ≡ characteristic function



old denotes a fuzzy set (relative to a certain set A)
fuzzy set ≡ membership function



2 Fuzzy sets: mathematical representation

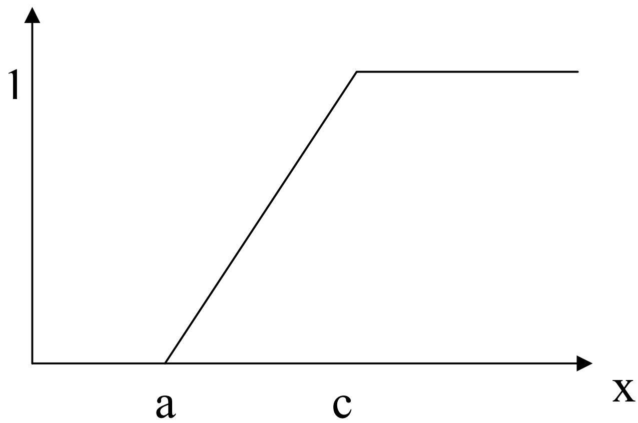
Definition:

Let V be the universe under consideration. A fuzzy set *A* is represented by a function μ_A: V $\rightarrow$ [0, 1].

- μ_A is called the membership function
- μ_A (x) is called the grad of membership of x w.r.t. *A*.
- μ_A (x) is also called the degree of truth of the proposition *that x is an element of A.*
- $\{x \in V: \mu_A (x) > 0\}$ is called the *support* of A

Example 1

$$L(x;a,c) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{c-a} & \text{if } a \le x \le c \\ 1 & \text{if } x > c \end{cases}$$



Example 2: Membership functions for Age

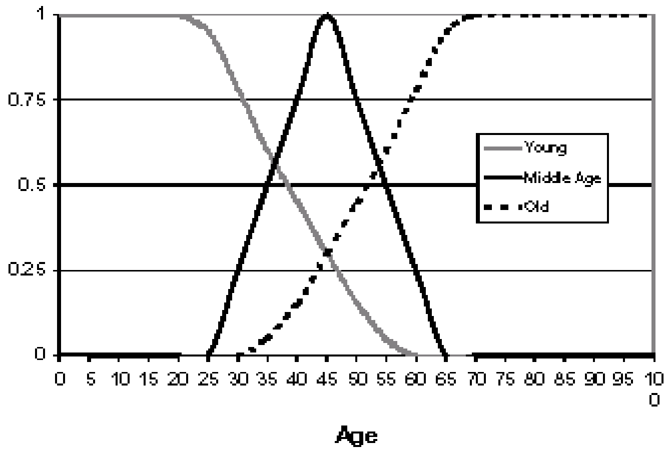
it is simple to provide analytic expressions that give a (step-wise) linear approximation to the three membership functions:

$\mu_{\textit{young man}}(x) = ?$

$\mu_{\textit{old man}}(x) = ?$

$\mu_{\textit{middle age man}}(x) = ?$

$\mu_{\textit{older man}}(x) = ???$



Linguistic variables (Fuzzy variables)

e.g.: OLD {very young, young , middle aged, old, very old}
The values of linguistic variables are fuzzy sets.

(The name derives from the circumstance that the values are often labeled by natural language expressions)

Other example

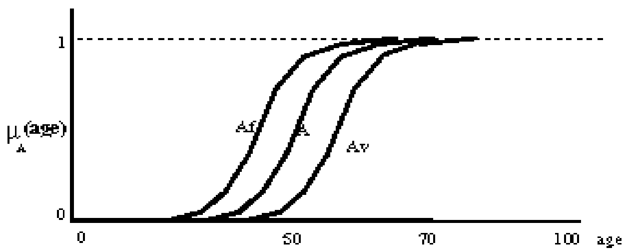
Fairly and very are examples of **fuzzy quantifiers**. These quantifiers can be used to generate linguistic variables.

$$\mu_{\text{very old}}(x) = F_{\text{very}}(\mu_{\text{old}}(x))$$

A: Peter is old

A_f: Peter is fairly old

A_v: Peter is very old



Quantifiers vs. Modifiers

Quantifiers

$\mu_{very\ old}(x) = F_{very}(\mu_{old}(x))$, e.g. $\mu_{very\ old}(x) = (\mu_{old}(x))^2$

However, sometimes $\mu_{old}(x) = 1$ but $\mu_{very\ old}(x) < 1$.

Translation modifiers

$\mu_{mod\ A}(x) = [F_{mod}(\mu_A)](x)$

$\mu_{\tau A}(x) = \mu_A(x-\tau)$

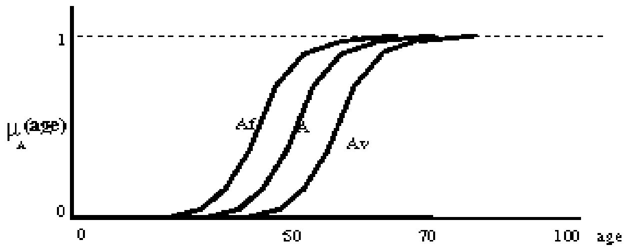
fairly $\cong$ -10 years

very $\cong$ +10 years

A: Peter is old

A_f: Peter is fairly old

A_v: Peter is very old



Where do the numbers come from?

The justification of degrees of truth/membership is a weak point of fuzzy logic.

- Justification of degrees of beliefs in terms of betting behavior (fair bets). However, we cannot bet on fuzzy expressions:
  - I bet you $5 that the patient is older than 30
  - ?? I bet you $5 that the patient is old
- In some sense, fuzzy logic makes a vague expression too precise by insisting on a numerical description.

- For atomic sentences it may be a reasonable strategy to ask a large number of people what they think of a proposition like "this person is old" and take the average.

- However, this cannot work for compound sentences since frequencies do not behave truth-functionally.

- In fuzzy control the problem is different: start with discrete values and *fuzzify* it. E.g. 45 for *age* can be mapped on the set {0, 0.2, 1, 0,2, 0} corresponding to the fuzzy sets {very young, young, middle aged, old, very old}

3 Combining fuzzy sets

Definition

Let A and B be fuzzy sets. The membership functions of A∪B, A∩B, and $\overline{A}$ are defined as follows:

1. $\mu_{A\cap B}(x) = \min(\mu_A(x), \mu_B(x))$
2. $\mu_{A\cup B}(x) = \max(\mu_A(x), \mu_B(x))$
3. $\mu_{\overline{A}}(x) = 1 - \mu_A(x)$

Note: sometimes we will write ¬A instead of $\overline{A}$.

Empirical problems

- The support of $A \cap \neg A$ is empty iff $\forall x(\mu_A(x)=0 \vee \mu_A(x)=1)$. (Exercise: prove it)

  Consider an object that is 40% red (and, consequently 60% *non-red*). Then the same object has a value of 40% for being both *red and non-red*. This is counter-intuitive

- Prototyp semantics and fuzzy sets. The example of a 'stripped apple'.
- Can these and other puzzles by resolved by considering other combination rules? Or is this a principled shortcoming of a compositional approach (as cognitive scientists claim)

Fuzzy sets and possibilities

Classical probability theory start with a set of events *W* and assigns to *all* subsets of *W* a probability μ. Possibility theory is just another approach to assign numbers to subsets of *W*. Instead of the probabilistic axioms P we assume axioms Poss:

P1.	$\mu(\emptyset)=0$	Poss1.	$\mathrm{Poss}(\emptyset)=0$
P2.	$\mu(W)=1$	Poss2.	$\mathrm{Poss}(W)=1$
P3.	$\mu(U \cup V) = \mu(U)+\mu(V)$ if *U* and *V* are disjoint	Poss3.	$\mathrm{Poss}(U \cup V) = \max(\mathrm{Poss}(U),\mathrm{Poss}(V))$ if *U* and *V* are disjoint

For fixed instances *a*, and $X \subseteq V$ the function $\mathrm{Poss}(X) = \mu_X(a)$ is a possibility function.

Principled approach to the choice of semantics

(Paris 1994, Hájek 1998)

(A) For negation

Consider ¬ as a function, ¬: [0, 1] → [0, 1].

N1 $\neg 0 = 1, \ \neg 1 = 0$

N2 ¬ is decreasing

N3 $\neg\neg x = x$ for $x \in [0,1]$

Principles N1-3 fairly constrain ¬; one can show a theorem:

Theorem1

If ¬ satisfies N1–3, then $([0,1], \neg, \leq)$ *is isomorphic to* $([0,1], 1-x, \leq)$.

(B) For conjunction

Consider $\wedge$ as a function, $\wedge: [0, 1] \times [0, 1] \to [0, 1]$.

C1 $0 \wedge 1 = 1 \wedge 0 = 0, \; 1 \wedge 1 = 1$

C2 $\wedge$ is continuous

C3 $\wedge$ is increasing (not necessarily strictly) in each coordinate

C4 $\wedge$ is associative.

A conjunction satisfying C1-C4 is called a t-Norm.

Examples for t-norms

- min(x, y) is a t-norm (Gödel t-norm)
- max(0, x+y-1) is a t-norm (Lukasiewicz t-norm)
- xy is a t-norm (Product t-norm)

Exercise: Show that these are t-norms

Theorem 2

Suppose $\wedge$ satisfies C1-4.

1. *If for all* $x \in [0,1]$, $x \wedge x = x$, *then* $\wedge = min$.

2. *If for some* a, b, c *such that* $0 \leq a < c < b \leq 1$: $c \wedge c = a$,

   then $([a,b], \wedge, \leq)$ *is isomorphic to*

   $([0,1], max(0, x + y - 1), \leq)$

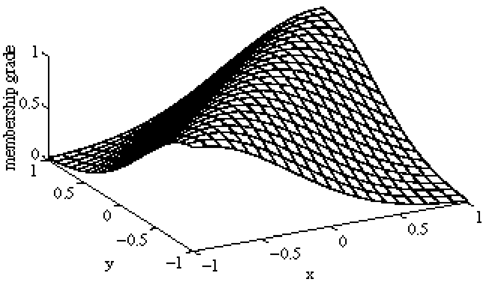
3. *Otherwise,* $([a,b], \wedge, \leq)$ *is isomorphic to* $([0,1], \times, \leq)$.

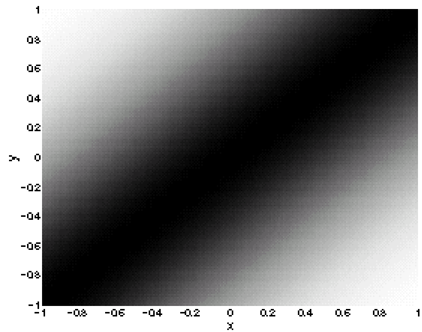
A good reference w.r.t. the mathematical details is Petr Hájek: Metamathematics of Fuzzy Logic. Kluwer 1998.

Example

Example: $R: \ x \approx y$ ("x is approximately equal to y")

$$\mu_R(x, y) = e^{-(x-y)^2}$$





Relational Composition 1

Let A be a a fuzzy set on the domain X and B a fuzzy relation on the domain X×X.

$A \circ B$, *the* composition *of* A *and* B *(w.r.t.* X*) is given by:*

$$\mu_{A \circ B}(y) = \sup_{x \in X}(\min(\mu_A(x), \mu_B(\langle x, y \rangle)))$$

Relational Composition 2

Let B and C fuzzy relation on the domain Y×Y.

$B \circ C$ *(w.r.t.* Y*) is given by:*

$$\mu_{B \circ C}(\langle x, z \rangle) = \sup_{y \in Y}(\min(\mu_B(\langle x, y \rangle), \mu_C(\langle y, z \rangle)))$$

Definition 3

Fuzzy propositional logic has the syntax of classical propositional logic and semantics given by real valuations ν (i.e. valuations assigning real numbers from the interval [0, 1]) satisfying

$$
\begin{aligned}
v(\neg A) &= 1 - v(A) \\
v(A \wedge B) &= \min(v(A), v(B)) \\
v(A \vee B) &= \max(v(A), v(B)) \\
v(A \rightarrow B) &= \min(1, 1 - v(A) + v(B))
\end{aligned}
$$

Definition 4

- φ is a valid formula iff $\nu(\varphi) = 1$ for all real valuations ν.
- ϕ is a valid fuzzy consequence of $\varphi_1 \ldots \varphi_n$, written $(\varphi_1, \ldots, \varphi_n) \models_F \phi$, if for all real valuations ν:
  $\nu(\varphi_1 \wedge \ldots \wedge \varphi_n) \leq \nu(\phi)$.

Example (deduction theorem)

Show that $A \rightarrow B$ is valid iff $A \models_F B$

Definition 5

Fuzzy predicate logic has the syntax of classical predicate logic and semantics given by first order models M = [D, ν] with valuations ν such that $\nu(A) : D \rightarrow [0, 1]$; $\nu(R) : DxD \rightarrow [0, 1]$. Further, ν has to satisfy the above rules for the propositional connectives and in addition these rules:

$$\nu(A(t)) = [\nu(A)]\ \nu(t)$$
$$\nu(R(t, u)) = [\nu(R)]\ (\nu(t),\ \nu(u))$$
$$\nu(\exists x\varphi(x)) = \sup\ (\{\nu(\varphi(t)) \text{: t is a term}\})$$
$$\nu(\forall x\varphi(x)) = \inf\ (\{\nu(\varphi(t)) \text{: t is a term}\})$$

Further, we assume that for each object $x \in D$ there exists a unique name $\underline{x}$.

The connection to fuzzy set theory hen is simply given by the following principle:

$$\nu(\underline{A}(\underline{x})) = \mu_A(x)$$

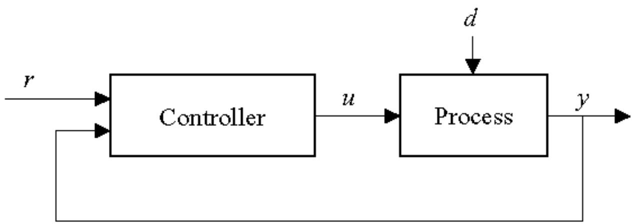
This principle allows translating valid statements in fuzzy set theory into valid logical formulas of fuzzy set logic.

Definition 6

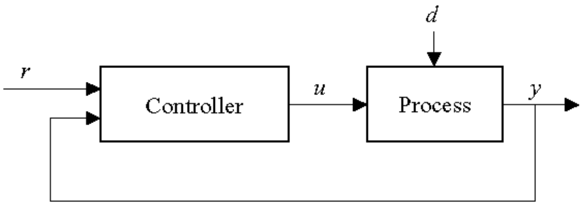
ϕ is a valid fuzzy consequence of $\varphi_1 \ldots \varphi_n$, written
$(\varphi_1, \ldots, \varphi_n) \models_F \phi$, if for all first order models $M = [D, \nu]$:
$\nu(\varphi_1 \wedge \ldots \wedge \varphi_n) \leq \nu(\phi)$.

5 Fuzzy control

The general situation is as follows: We have a process S, yielding output y(t), where t is the time variable. d is a disturbance (input that cannot be influenced). The output signal is compared to the desired r(t). If y(t) differs significantly from r(t), then a corrective signal u(y,r,t) is supplied to S with the purpose of bringing y(t) in line with r(t). *The task is to determine the most appropriate u*.

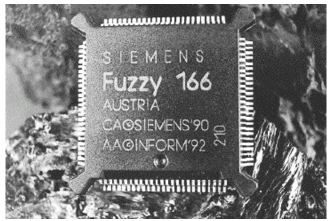


Mathematically



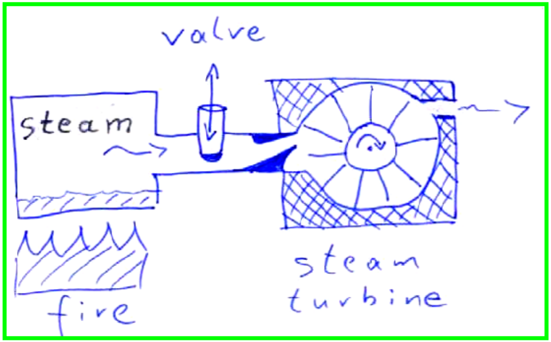
- $dy/dt = h(y,t,u)$, $y(0) = c$
- $\int_0^T \sigma(y(t) - u(t))dt \rightarrow$ minimize and determine u!

Difficult to solve ! Calculus of variation, dynamic programming.
Fuzzy logic has been applied to easy the computational difficulties.



Example

Steam engine driving a steam turbine (Mamdami & Asilian 1975)



u [VC]

y [DR,CDR]

- VC, the required change in the valve opening
- DR, the deviation in the number of revolutions
- CDR, the change in DR with respect to the last measurement

IF …THEN rules are all of the form: IF *DR is A* AND *CDR is B*, THEN *VC is C*

e.g. IF *DR is LN* AND *CDR is SP* THEN *VC is LP*

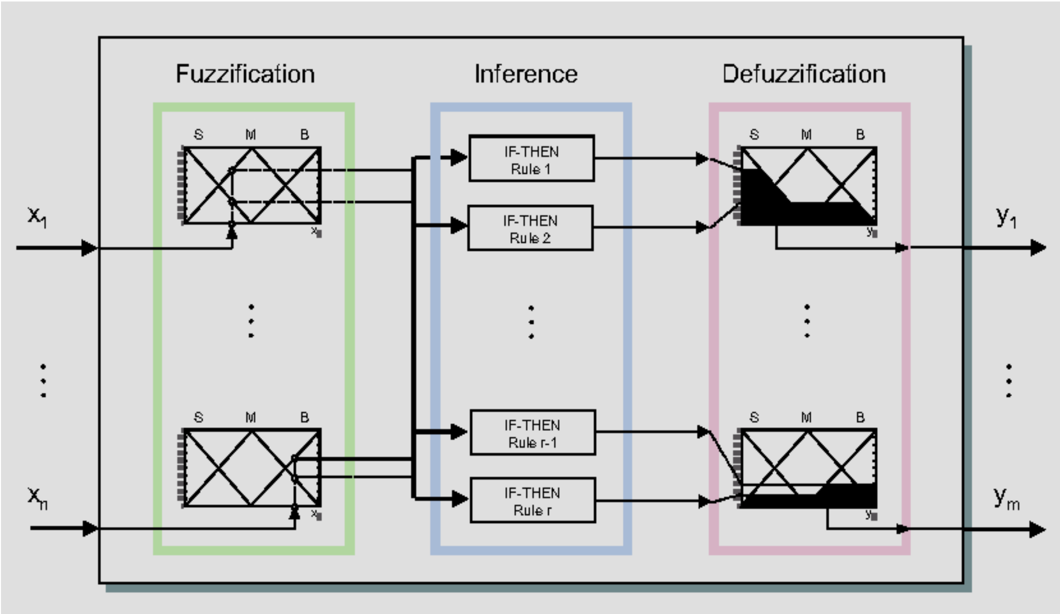
Values of A, B, C:

LP	*large positive*
MP	*medium positive*
SP	*small positive*
0	*zero*
SN	*small negative*
MN	*medium negative*
LN	*large negative*

Rule	DR	CDR	VC
1	LN	$\neg(LN \vee MN)$	LP
2	MN	$LP \vee MP \vee SP$	SP
3	SN	$LP \vee MP$	SP
4	0^-	LP	SP
5	$0^- \vee 0^+$	$SP \vee SN \vee 0$	0
6	0^+	LP	SN
7	SP	$LP \vee MP$	SN
8	MP	$LP \vee MP \vee SP$	SN
9	LP	$\neg(LN \vee MN)$	LN

Such rules are intuitively plausible. They give a qualitative analysis of the corresponding differential equation.

Fuzzy system components



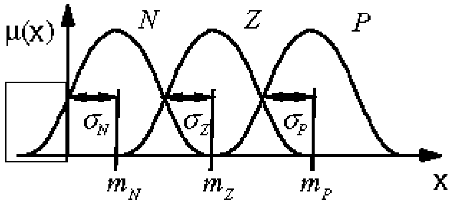
(A) Fuzzification

It is fairly straightforward in the case under discussion.

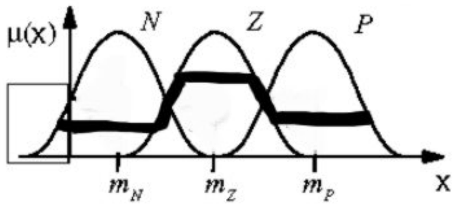
- Measure the precise outcome x for a linguistic variable X

- Determine the number $\mu_A(x)$ for each fuzzy value A of X

  e.g. $\mu_N(x) = 0.3$, $\mu_Z(x) = 0.8$, $\mu_P(x) = 0.4$



- Combine these numbers into a new fuzzy set

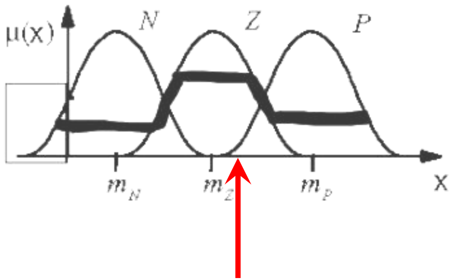


(B) De-Fuzzification

Various strategies have been formulated, all of them apparently rather arbitrary.

Center of gravity method

$$y_0 = \frac{\sum_{j=1}^{F} \mu_{B'}(y_j) y_j}{\sum_{j=1}^{F} \mu_{B'}(y_j)}$$



Fuzzy implications: **IF X is A THEN Y is B**

X, Y linguistic variables; A,B fuzzy sets

Ideas

1. Represent each if–then rule as a fuzzy relation.
2. Aggregate these relations in one relation representative for the entire rule base.
3. Given an input, use *relational composition* to derive the corresponding output.

Idea 1

- Fuzzy implications correspond to fuzzy relations:

[IF X is A THEN Y is B](x,y) = I(μ_A(x), (μ_B(y)) = min(1, 1-μ_A(x)+μ_B(y))

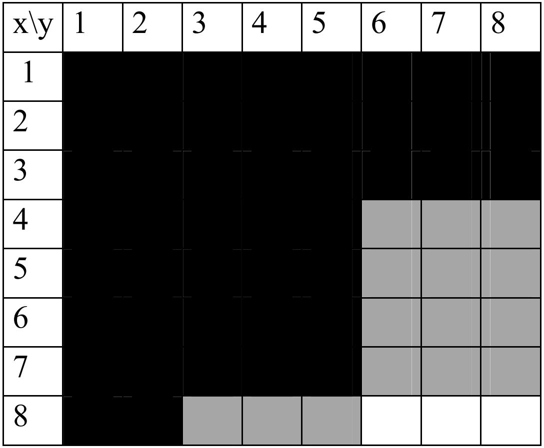
	1	2	3	4	5	6	7	8
μ_{small}	1	1	.5	.5	.5	0	0	0
μ_{medium}	0	.5	.5	.5	1	1	.5	.5
μ_{large}	0	0	0	.5	.5	.5	.5	1

e.g. If X is large then Y is small: I(μ_{large}(x), (μ_{small}(y))

If X is small then Y is large: I(μ_{small}(x), (μ_{large}(y))

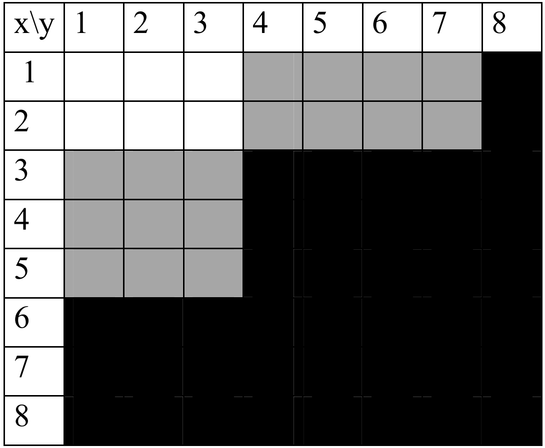
If X is large then Y is small

$I(\mu_{large}(x), (\mu_{small}(y))$



If X is small then Y is large

$I(\mu_{small}(x), (\mu_{large}(y))$

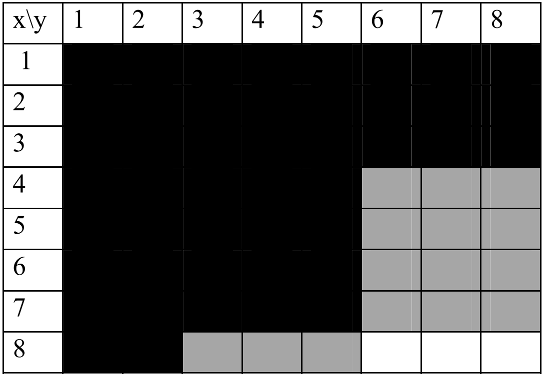


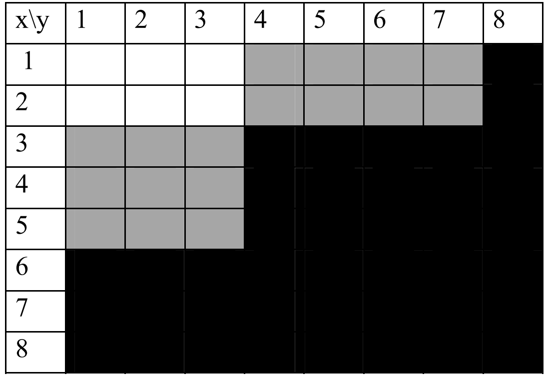
Idea 2

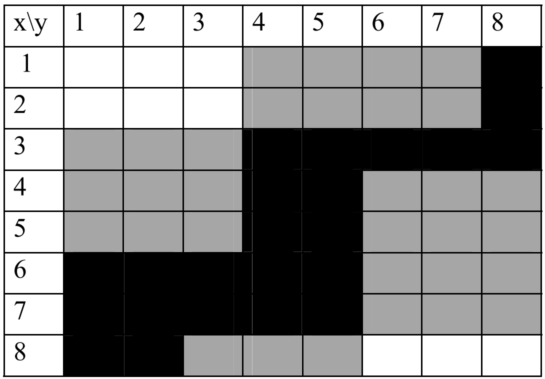
- Using *aggregation* to construct one relation representative for the entire rule base.

$$\mu_R(x,y) = \mathrm{aggr}(\mu_{A_i}(x))$$

The *aggr* operator is the minimum for implications (notice that the *maximum* in case conjunction has been used)


x\y
1 2 3 4 5 6 7 8
1 2 3 4 5 6 7 8


x\y
1 2 3 4 5 6 7 8
1 2 3 4 5 6 7 8


x\y
1 2 3 4 5 6 7 8
1 2 3 4 5 6 7 8

Idea 3

- Using relational fuzzy composition we can derive the fuzzy set A* ○ R. This fuzzy set represents the consequence that can be drawn from the fact that **X is an A*** and the set of the **collected Rules**.

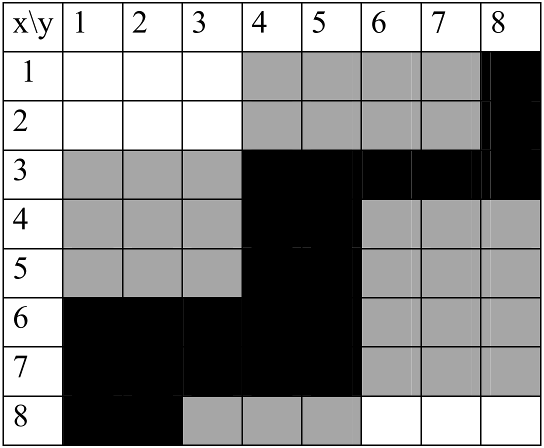
For example, let A* be the crisp set {6}, then

The operation of composition, repeated here,

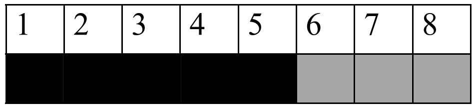
$A \circ B$, *the* composition *of* A *and* B *(w.r.t.* X*) is given by:*

$$\mu_{A \circ B}(y) = \sup_{x \in X}(\min(\mu_A(x), \mu_B(\langle x, y \rangle)))$$

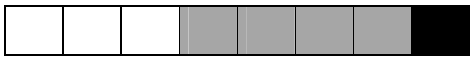
gives us the following fuzzy set:



For A* = {6} $\mu_B =$



For A* = {1} $\mu_B =$



For A* ={8} $\mu_B =$