Review 2nd Midterm

- Hedge
  - Dilation: $\boldsymbol{HA}(x) = (\boldsymbol{A}(x))^{1/2}$ → more or less of $\boldsymbol{A}$
  - Concentration: $\boldsymbol{HA}(x) = (\boldsymbol{A}(x))^{2}$ → very $\boldsymbol{A}$
  - Plus: $\boldsymbol{HA}(x) = (\boldsymbol{A}(x))^{1.25}$
  - Intensification:

$$HA(x) = \begin{cases} 2A(x)^2 & \text{if } A(x) \in \left[0, \frac{1}{2}\right] \\ 1 - 2(1 - A(x))^2 & \text{otherwise} \end{cases}$$

→ slightly

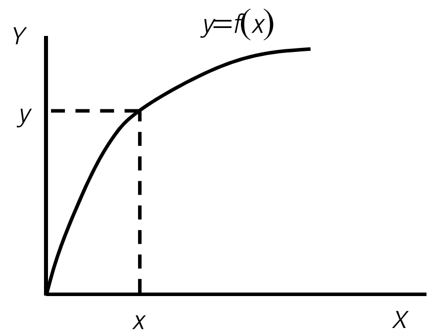
  - Vague

$$\boldsymbol{HA}(x) = \begin{cases} \sqrt{\dfrac{\boldsymbol{A}(x)}{2}} & \text{if } \boldsymbol{A}(x) \le 0.5 \\ 1 - \sqrt{\dfrac{1 - \boldsymbol{A}(x)}{2}} & \text{if } \boldsymbol{A}(x) > 0.5 \end{cases}$$

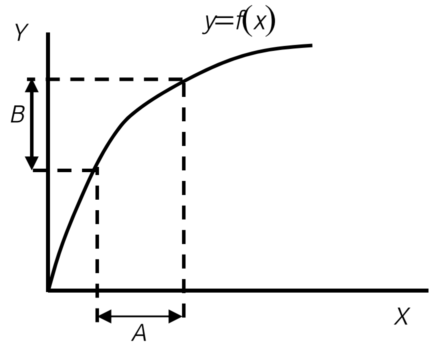
→ seldom

- Generalized modus ponen

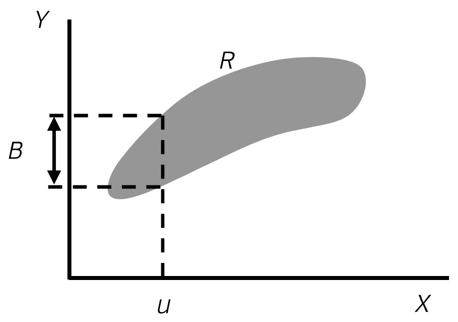
implication:	If χ is $\boldsymbol{A}$, then Y is $\boldsymbol{B}$
premise:	χ is $\boldsymbol{A'}$
conclusion:	Y is $\boldsymbol{B'}$

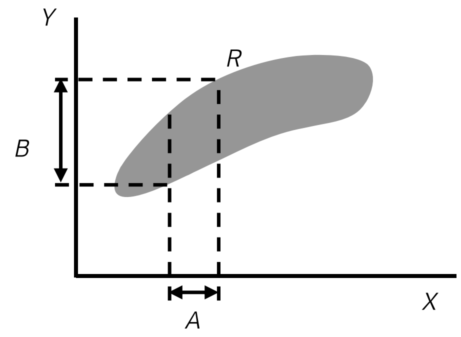


$\chi = x$ we get $Y = y = f(x)$



χ is in A we get Y in $B = \{\, y \in Y \mid y = f(x),\ x \in A \}$

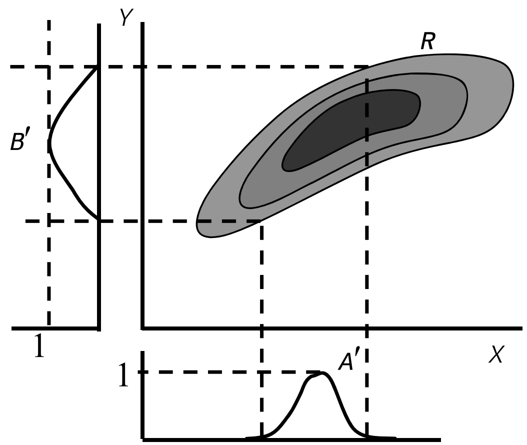




Relation R on $X \times Y$ if $\chi = u$ and R we get $Y \in B$
where $B = \left\{ y \in Y \,\middle|\, \langle x, y \rangle \in R \right\}$ (left image)

And if $\chi \in A$ we get $Y \in B$ where
$B = \left\{ y \in Y \,\middle|\, \langle x, y \rangle \in R,\ x \in A \right\}$ (right image)

$$\mathcal{X}_B(y) = \sup_{x \in X} \min\left[\mathcal{X}_A(x), \mathcal{X}_R(x, y)\right]$$



If R is a fuzzy relation on $X \times Y$ where A and A' are fuzzy sets on X, B and B' are fuzzy sets on Y

If R (relation between A and B) and A' are given, we can find B' using

$$B'(y) = \sup_{x \in X} \min\left[A'(x), R(x, y)\right]$$

Or $B' = A' \bullet R$ operator $\bullet$ is a composition operator

Hence, this is called compositional rule of inference

- Example

implication: if x and y are approximately equal

<u>premise: x is little</u>

conclusion: ?

Suppos little $\rightarrow$ $\boldsymbol{A} = \{(1,1), (2,0.6), (3,0.2), (4,0)\}$

Suppose

$$R = \begin{bmatrix} 1.0 & 0.5 & 0.0 & 0.0 \\ 0.5 & 1.0 & 0.5 & 0.0 \\ 0.0 & 0.5 & 1.0 & 0.5 \\ 0.0 & 0.0 & 0.5 & 1.0 \end{bmatrix}$$

from $\boldsymbol{B} = \boldsymbol{A} \bullet \boldsymbol{R}$

$$B = \begin{bmatrix} 1.0 & 0.6 & 0.2 & 0.0 \end{bmatrix} \bullet \begin{bmatrix} 1.0 & 0.5 & 0.0 & 0.0 \\ 0.5 & 1.0 & 0.5 & 0.0 \\ 0.0 & 0.5 & 1.0 & 0.5 \\ 0.0 & 0.0 & 0.5 & 1.0 \end{bmatrix}$$

$\rightarrow$ $B = \begin{bmatrix} 1.0 & 0.6 & 0.5 & 0.2 \end{bmatrix}$

Max-min composition

- p: If χ is $\boldsymbol{A}$, then Y is $\boldsymbol{B}$ for $x \in X$ and $y \in Y$, relation will be

$$\boldsymbol{R}(x,\mathrm{y}) = \mathrm{I}(\boldsymbol{A}(x), \boldsymbol{B}(y))$$

Where I is fuzzy implication

  - Lukasiewicz: $\mathrm{I}(\boldsymbol{A}(x),\boldsymbol{B}(y)) = \min[1, 1-\boldsymbol{A}(x)+\boldsymbol{B}(y)]$
  - Zadeh: $\mathrm{I}(\boldsymbol{A}(x),\boldsymbol{B}(y)) = \max((1-\boldsymbol{A}(x)), \min(\boldsymbol{A}(x),\boldsymbol{B}(y)))$
  - Kleen-Dienes: $\mathrm{I}(\boldsymbol{A}(x),\boldsymbol{B}(y)) = \max((1-\boldsymbol{A}(x)),\boldsymbol{B}(y))$
  - Mamdani (correlation-min): $\mathrm{I}(\boldsymbol{A}(x),\boldsymbol{B}(y)) = \min(\boldsymbol{A}(x),\boldsymbol{B}(y))$
  - Correlation-product: $\mathrm{I}(\boldsymbol{A}(x),\boldsymbol{B}(y)) = \boldsymbol{A}(x)\times\boldsymbol{B}(y)$
  - Goedel:

$$I(A(x),B(y)) = \begin{cases} 1 & \text{if} \quad A(x) \le B(y) \\ B(y) & \text{if} \quad A(x) > B(y) \end{cases}$$

- If χ is $\boldsymbol{A}$, then Y is $\boldsymbol{B}$ → If χ is $\boldsymbol{A}$, then Y is $\boldsymbol{B}$ else V is ***UNKNOWN*** where χ, Y and V are variables in X, Y and V and $\boldsymbol{A}$, $\boldsymbol{B}$ and ***UNKNOWN*** are fuzzy sets on X, Y and V

$$\boldsymbol{R} = \boldsymbol{A} \times \boldsymbol{B} + \overline{A} \times \boldsymbol{UNKNOWN}$$

× → minimum, + → maximum

- If χ is $\boldsymbol{A}$, then Y is $\boldsymbol{B}$ else Z is $\boldsymbol{C}$ →

$$\boldsymbol{R} = \boldsymbol{A} \times \boldsymbol{B} + \overline{A} \times \boldsymbol{C}$$

- Example

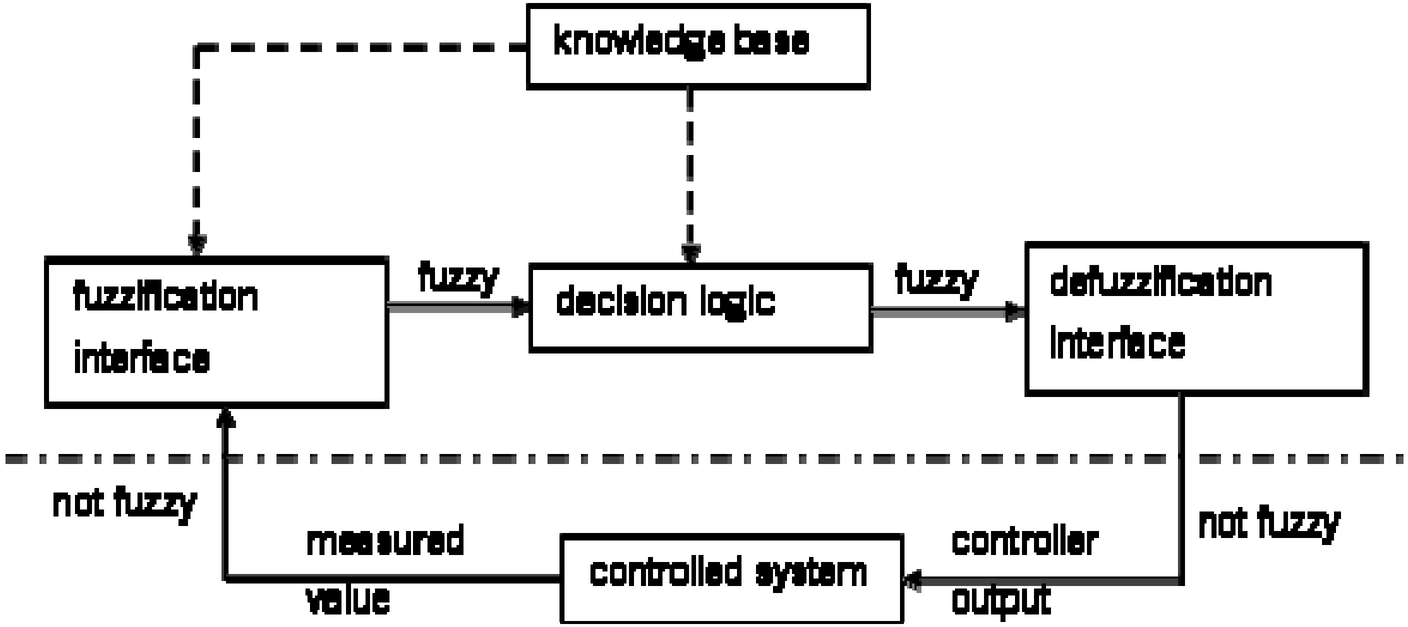
$\boldsymbol{A} = 1/1 + 0.4/2$, $\boldsymbol{B} = 0.4/2 + 1/3$ and $\boldsymbol{C} = 1/1 + 0.6/2$

If χ is $\boldsymbol{A}$, then Y is $\boldsymbol{B}$ else Z is $\boldsymbol{C}$ →

$$R = \begin{bmatrix} 1.0 \\ 0.4 \\ 0.0 \end{bmatrix} \begin{bmatrix} 0.0 & 0.4 & 1.0 \end{bmatrix} + \begin{bmatrix} 0.0 \\ 0.6 \\ 1.0 \end{bmatrix} \begin{bmatrix} 1.0 & 0.6 & 0.0 \end{bmatrix} \quad R = \begin{bmatrix} 0.0 & 0.4 & 1.0 \\ 0.0 & 0.4 & 0.4 \\ 0.0 & 0.0 & 0.0 \end{bmatrix} + \begin{bmatrix} 0.0 & 0.0 & 0.0 \\ 0.6 & 0.6 & 0.0 \\ 1.0 & 0.6 & 0.0 \end{bmatrix} \quad R = \begin{bmatrix} 0.0 & 0.4 & 1.0 \\ 0.6 & 0.6 & 0.4 \\ 1.0 & 0.6 & 0.0 \end{bmatrix}$$

If χ is $\boldsymbol{A}$, then Y is $\boldsymbol{B}$ →

$$R = \begin{bmatrix} 1.0 \\ 0.4 \\ 0.0 \end{bmatrix} \begin{bmatrix} 0.0 & 0.4 & 1.0 \end{bmatrix} + \begin{bmatrix} 0.0 \\ 0.6 \\ 1.0 \end{bmatrix} \begin{bmatrix} 1.0 & 1.0 & 1.0 \end{bmatrix} \quad R = \begin{bmatrix} 0.0 & 0.4 & 1.0 \\ 0.0 & 0.4 & 0.4 \\ 0.0 & 0.0 & 0.0 \end{bmatrix} + \begin{bmatrix} 0.0 & 0.0 & 0.0 \\ 0.6 & 0.6 & 0.6 \\ 1.0 & 1.0 & 1.0 \end{bmatrix} \quad R = \begin{bmatrix} 0.0 & 0.4 & 1.0 \\ 0.6 & 0.6 & 0.6 \\ 1.0 & 1.0 & 1.0 \end{bmatrix}$$

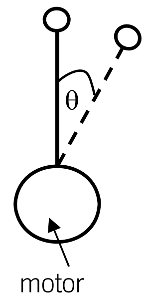

knowledge base
fuzzification interface
fuzzy
decision logic
fuzzy
defuzzification interface
not fuzzy
measured value
controlled system
controller output
not fuzzy

- Example: Inverted pendulum

  input: angle θ and angular velocity Δθ

  output: current v_t

  membership functions → same name for θ , Δθ and v_t



negative large (***NL***), negative medium (***NM***), negative small (***NS***), zero (***ZE***), positive small (***PS***), positive medium (***PM***), and positive large (***PL***)

rule j: if θ is $\boldsymbol{A}_j$ and Δθ is $\boldsymbol{B}_j$ then v_t is $\boldsymbol{C}_j$

e.g. if θ is ***NL*** and Δθ is ***NL*** then v_t is ***PL*** → (***NL***,***NL***;***PL***)

or if θ is ***ZE*** and Δθ is ***ZE*** then v_t is ***ZE*** → (***ZE***,***ZE***;***ZE***)

$\Delta\theta$ \ θ	*NL*	*NM*	*NS*	*ZE*	*PS*	*PM*	*PL*
NL				***PL***			
NM				***PM***			
NS				***PS***			
ZE	***PL***	***PM***	***PS***	***ZE***	***NS***	***NM***	***NL***
PS				***NS***			
PM				***NM***			
PL				***NL***			

Total number of rules → $7 \times 7 \times 7 = 343$ but some of rules do not make sense. Hence, the number of usable rules will be less than that

- Fuzzy associative memory (FAM)

rule 1: If χ is $\boldsymbol{A}_1$, then Y is $\boldsymbol{B}_1 \rightarrow (\boldsymbol{A}_1, \boldsymbol{B}_1) \rightarrow \boldsymbol{R}_1$

rule 2: If χ is $\boldsymbol{A}_2$, then Y is $\boldsymbol{B}_2 \rightarrow (\boldsymbol{A}_2, \boldsymbol{B}_2) \rightarrow \boldsymbol{R}_2$

$\vdots$

rule n: If χ is $\boldsymbol{A}_n$, then Y is $\boldsymbol{B}_n \rightarrow (\boldsymbol{A}_n, \boldsymbol{B}_n) \rightarrow \boldsymbol{R}_n$

premise: χ is $\boldsymbol{A}'$

---

conclusion: Y is $\boldsymbol{B}'$

$$\boldsymbol{B}' = w_1\boldsymbol{B}_1' + w_2\boldsymbol{B}_2' + \ldots + w_n\boldsymbol{B}_n' \text{ where } \boldsymbol{B}_i' = \boldsymbol{A}' \bullet \boldsymbol{R}_i$$

Supposed: $\boldsymbol{A}_i = a_1/x_1 + a_2/x_2 + ... + a_m/x_m \rightarrow \boldsymbol{A}_i = [a_1, a_2, ..., a_\mathrm{m}]$

Correlation-min $\rightarrow$

$$\boldsymbol{R}_i = \boldsymbol{A}_i^T \circ \boldsymbol{B}_i \quad \rightarrow \quad \boldsymbol{R}_i = \begin{bmatrix} a_1 \wedge \boldsymbol{B}_i \\ a_2 \wedge \boldsymbol{B}_i \\ \vdots \\ a_m \wedge \boldsymbol{B}_i \end{bmatrix}$$

$$\boldsymbol{R}_i = \begin{bmatrix} b_1 \wedge \boldsymbol{A}_i^\mathrm{T} & b_2 \wedge \boldsymbol{A}_i^\mathrm{T} \cdots b_p \wedge \boldsymbol{A}_i^\mathrm{T} \end{bmatrix}$$

Correlation-product $\rightarrow$

$$\boldsymbol{R}_i = \boldsymbol{A}_i^T \circ \boldsymbol{B}_i \quad \rightarrow \quad \boldsymbol{R}_i = \begin{bmatrix} a_1 \boldsymbol{B}_i \\ a_2 \boldsymbol{B}_i \\ \vdots \\ a_m \boldsymbol{B}_i \end{bmatrix}$$

$$\boldsymbol{R}_i = \begin{bmatrix} b_1 \boldsymbol{A}_i^\mathrm{T} & b_2 \boldsymbol{A}_i^\mathrm{T} \cdots b_p \boldsymbol{A}_i^\mathrm{T} \end{bmatrix}$$

Conclude→ $\boldsymbol{B}' = \boldsymbol{A}' \bullet \boldsymbol{R}$

$$= \boldsymbol{A}' \bullet \bigcup_{j \in \mathrm{N}_n} \boldsymbol{R}_j$$

$$= \boldsymbol{A}' \bullet \sup_{j \in \mathrm{N}_n} \boldsymbol{R}_j$$

$$= \sup_{j \in \mathrm{N}_n} \left(A' \bullet R_j \right)$$

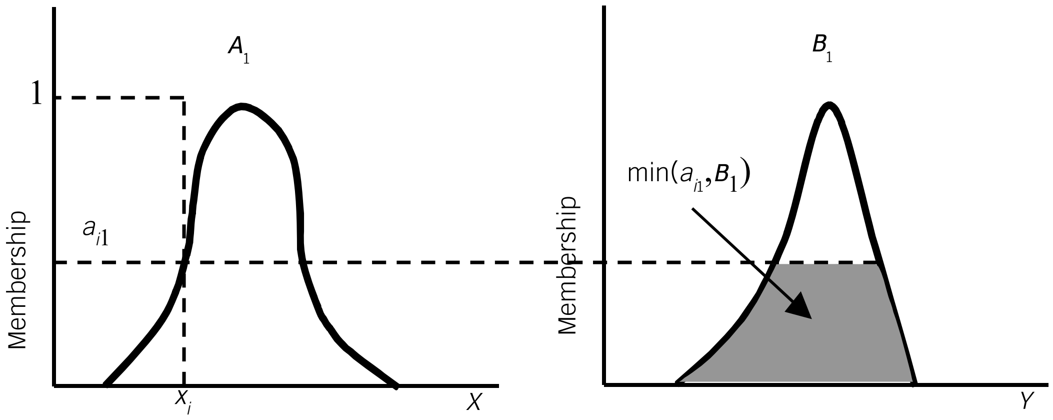
$$= \sup_{j \in \mathrm{N}_n} \left(A' \bullet A_j^{\mathrm{T}} \circ B_j \right)$$

$$= \sup_{j \in \mathrm{N}_n} \left(\left(A' \bullet A_j^{\mathrm{T}} \right) \circ B_j \right)$$

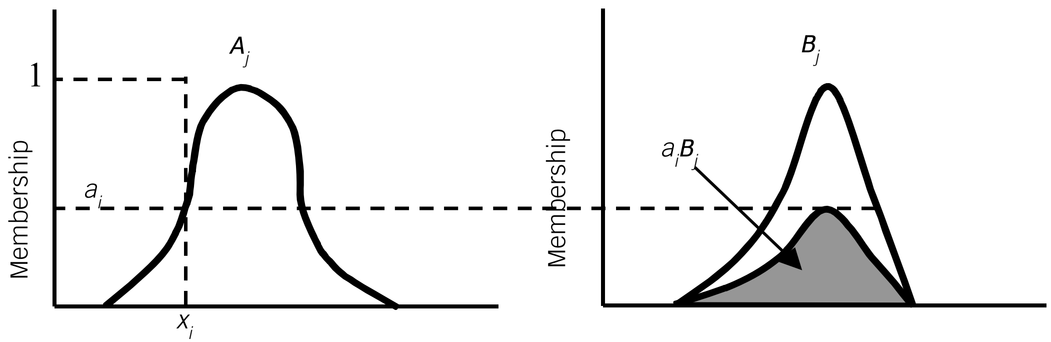
$$B = \sup \left[\left(A' \bullet A_1^{\mathrm{T}} \right) \circ B_1, \left(A' \bullet A_2^{\mathrm{T}} \right) \circ B_2, \ldots, \left(A' \bullet A_i^{\mathrm{T}} \right) \circ B_i, \ldots, \left(A' \bullet A_n^{\mathrm{T}} \right) \circ B_n \right]$$

- $\boldsymbol{A'} = 0/x_1 + 0/x_2 + \ldots + 1/x_i + 0/x_{i+1} + \ldots + 0/x_{\mathrm{m}}$

Correlation-min→ $(\boldsymbol{A'} \bullet \boldsymbol{A}_j^{\mathrm{T}}) \circ \boldsymbol{B}_j = \min(\boldsymbol{A}_i(x_i), \boldsymbol{B}_j)$



Correlation-product→ $(\boldsymbol{A'} \bullet \boldsymbol{A}_j^{\mathrm{T}}) \circ \boldsymbol{B}_j = \boldsymbol{A}_i(x_i)\boldsymbol{B}_j$



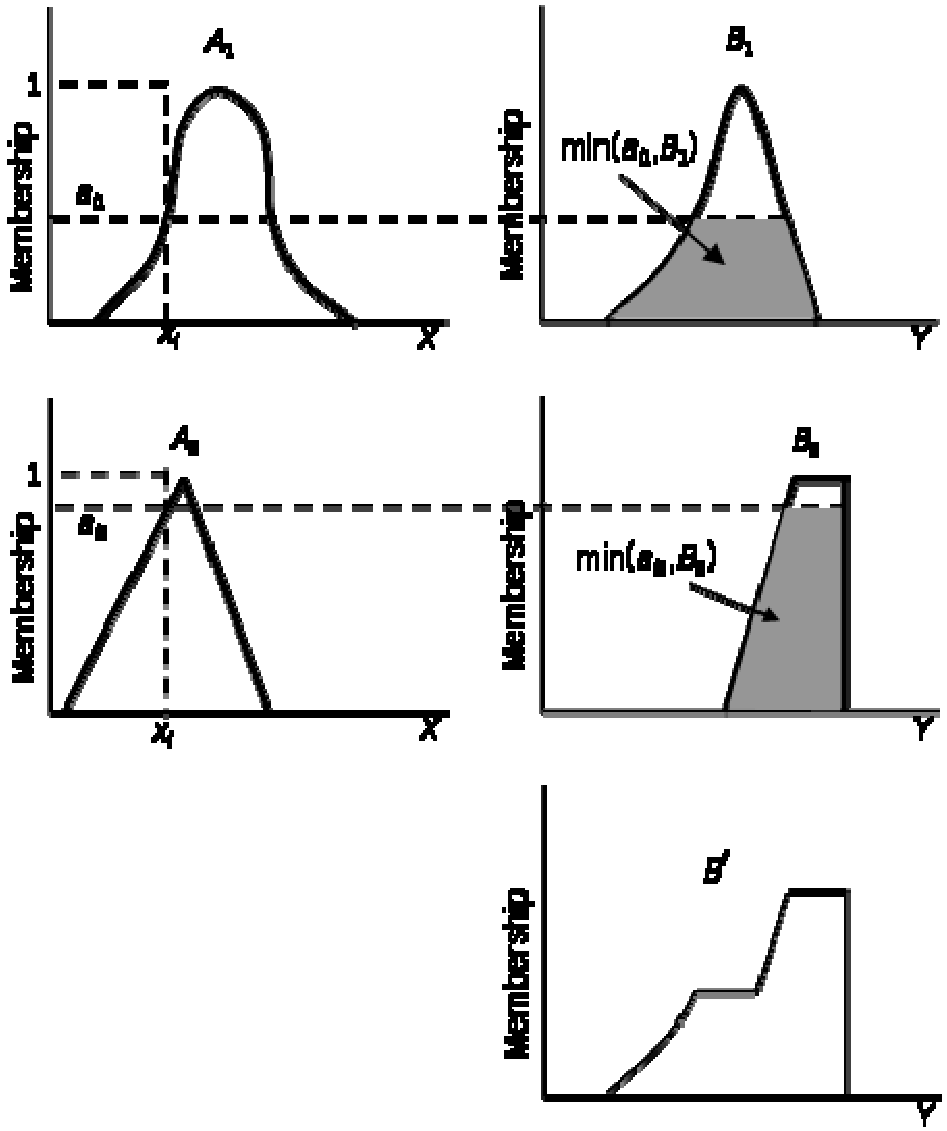
- Example

$\boldsymbol{A}' = 0/x_1 + 0/x_2 + ... + 1/x_i + 0/x_{i+1} + ... + 0/x_m$

Supposed that there are 2 rules $(\boldsymbol{A}_1;\boldsymbol{B}_1)$ and $(\boldsymbol{A}_2;\boldsymbol{B}_2)$

$\boldsymbol{B}' = \sup\,[(\boldsymbol{A}'\bullet\boldsymbol{A}_1^{\mathrm{T}}) \circ \boldsymbol{B}_1, (\boldsymbol{A}'\bullet\boldsymbol{A}_2^{\mathrm{T}}) \circ \boldsymbol{B}_2]$

where $(\boldsymbol{A}'\bullet\boldsymbol{A}_1^{\mathrm{T}}) = a_{i1}$ and $(\boldsymbol{A}'\bullet\boldsymbol{A}_2^{\mathrm{T}}) = a_{i2}$



- Example

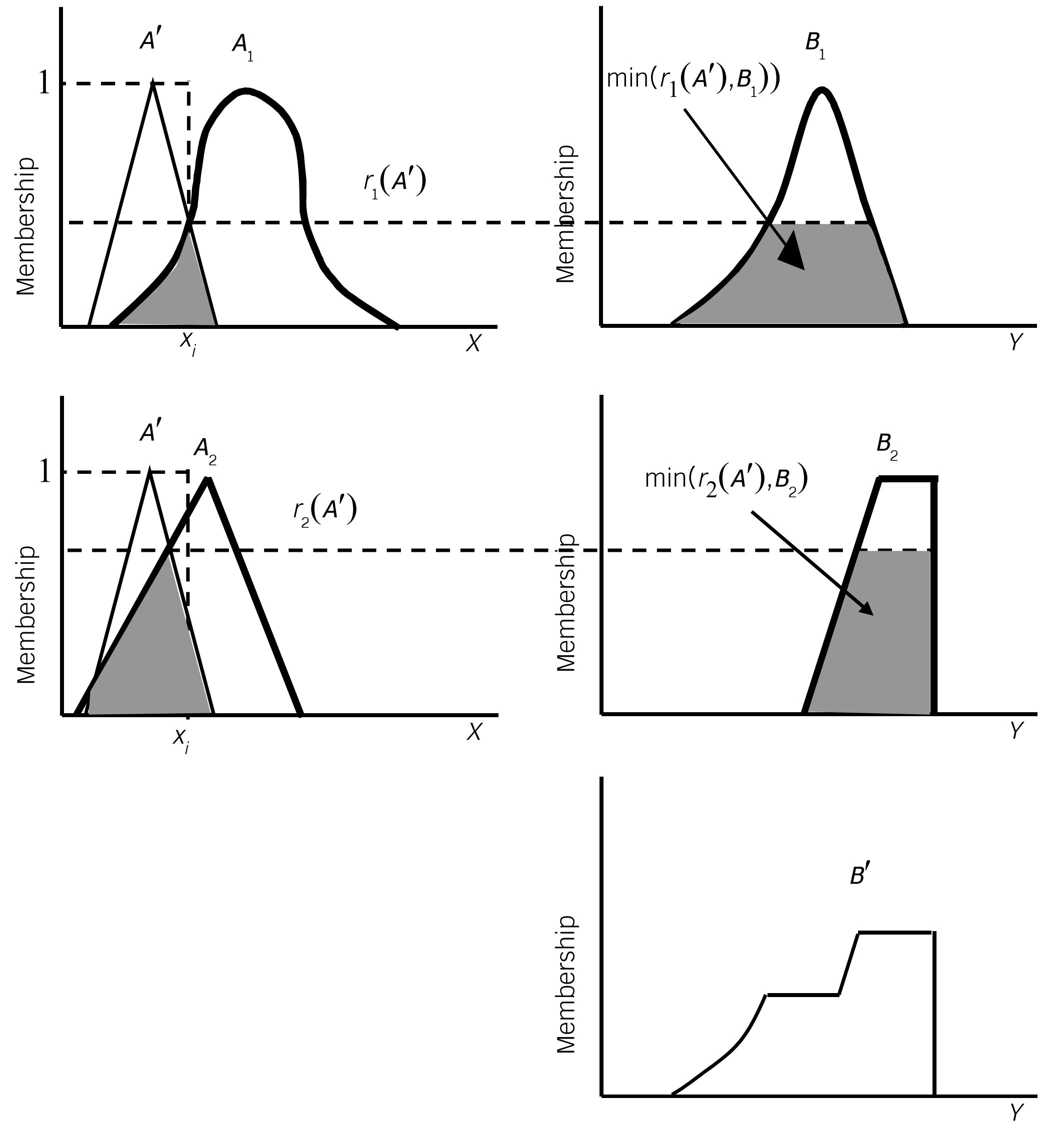
$\boldsymbol{A'}$ is a fuzzy set not crisp set

Supposed that there are 2 rules($\boldsymbol{A}_1$;$\boldsymbol{B}_1$) and ($\boldsymbol{A}_2$; $\boldsymbol{B}_2$)

$\boldsymbol{B'}$ = sup [($\boldsymbol{A'}\bullet\boldsymbol{A}_1^{\mathrm{T}}$) ∘ $\boldsymbol{B}_1$, ($\boldsymbol{A'}\bullet\boldsymbol{A}_2^{\mathrm{T}}$) ∘ $\boldsymbol{B}_2$]

where

$$r_j\left(A'\right)=\left(A'\bullet A_j^T\right)=\sup_{x\in X}\min\left[A'(x),A_j(x)\right]$$



- If there are more than 1 input→ (***A***,***B***;***C***)

$\boldsymbol{A} = a_1/x_1 + a_2/x_2 + ... + a_m/x_m$ and $\boldsymbol{B} = b_1/y_1 + b_2/y_2 + ... + b_m/y_m$

split into (***A***;***C***) และ (***B***;***C***) →

$$M_{AC} = A^T \circ C \quad \text{and} \quad M_{BC} = B^T \circ C$$

fact: $\boldsymbol{A}' = \mathrm{I}_x^{\,i} = 0/x_1 + 0/x_2 + ... + 1/x_i + 0/x_{i+1} + ... + 0/x_m$ and

$\boldsymbol{B}' = \mathrm{I}_y^{\,j} = 0/y_1 + 0/y_2 + ... + 1/y_j + 0/j_{i+1} + ... + 0/y_m$

compute: $\mathrm{F}(\boldsymbol{A}', \boldsymbol{B}') = \boldsymbol{C}' = [\boldsymbol{A}' \bullet \boldsymbol{M_{AC}}] \cap [\boldsymbol{B}' \bullet \boldsymbol{M_{BC}}]$

where $\boldsymbol{A}' \bullet \boldsymbol{M_{AC}} = \mathrm{I}_x^{\,i} \bullet \boldsymbol{M_{AC}}$

$$= \mathrm{I}_x^{\,i} \bullet \begin{bmatrix} a_1 \wedge C \\ a_2 \wedge C \\ \vdots \\ a_m \wedge C \end{bmatrix} \qquad = a_i \wedge \boldsymbol{C}$$

and $\boldsymbol{B'} \bullet \boldsymbol{M_{BC}} = \mathrm{I}_y^j \bullet \boldsymbol{M_{BC}}$

$$= \mathrm{I}_y^j \bullet \begin{bmatrix} b_1 \wedge C \\ b_2 \wedge C \\ \vdots \\ b_m \wedge C \end{bmatrix} = b_j \wedge \boldsymbol{C}$$

Hence, the conclusion will be $\boldsymbol{C'} = (a_i \wedge \boldsymbol{C}) \cap (b_j \wedge \boldsymbol{C}) = (\min(a_i, b_j)) \wedge \boldsymbol{C}$

If use correlation-product

$$\boldsymbol{C'} = (a_i \boldsymbol{C}) \cap (b_j \boldsymbol{C}) = (\min(a_i, b_j)) \boldsymbol{C}$$

- Mamdani model

If ξ_1 is $\boldsymbol{A}^{(1)}$, and ξ_2 is $\boldsymbol{A}^{(2)}$ and ... and ξ_n is $\boldsymbol{A}^{(n)}$ then η is $\boldsymbol{B}$

where $\boldsymbol{A}^{(1)}$, $\boldsymbol{A}^{(2}$ and ... and $\boldsymbol{A}^{(n)}$ are linguistic terms in $\boldsymbol{T}(x_i)$ for $1 \le i \le n$ and $\boldsymbol{B}$ is linguistic term in $\boldsymbol{T}(y)$ and suppose there are more than 1 rule

Let $\boldsymbol{T}(x_1) \rightarrow \boldsymbol{A}_1{}^{(1)}, \boldsymbol{A}_2{}^{(1)}, \ldots, \boldsymbol{A}_{N1}{}^{(1)}$

$\boldsymbol{T}(x_2) \rightarrow \boldsymbol{A}_1{}^{(2)}, \boldsymbol{A}_2{}^{(2)}, \ldots, \boldsymbol{A}_{N2}{}^{(2)}$

$\vdots$

$\boldsymbol{T}(x_n) \rightarrow \boldsymbol{A}_1{}^{(n)}, \boldsymbol{A}_2{}^{(n)}, \ldots, \boldsymbol{A}_{Nn}{}^{(n)}$

and $\boldsymbol{T}(y) \rightarrow \boldsymbol{B}_1, \boldsymbol{B}_2, \ldots, \boldsymbol{B}_{N0}$

rule j: If ξ_1 is $\boldsymbol{A}_{i1,j}{}^{(1)}$, and ξ_2 is $\boldsymbol{A}_{i2,j}{}^{(2)}$ and ... and ξ_n is $\boldsymbol{A}_{in,j}{}^{(n)}$ then η is $\boldsymbol{B}_{i,j}$

where $i1 \in \{1,2,\ldots,N1\}$, $i2 \in \{1,2,\ldots,N2\}, \ldots,$ $in \in \{1,2,\ldots,Nn\}$ and

$i \in \{1,2,\ldots,N0\}$

Firing strength of rule *j* will be

$$\alpha_j = \min\{\boldsymbol{A}_{i1,j}{}^{(1)}(x_1), \boldsymbol{A}_{i2,j}{}^{(2)}(x_2), \ldots, \boldsymbol{A}_{in,j}{}^{(n)}(x_n)\}$$

Output of rule *j* will be

$$OUT^{(j)}_{x_1,x_2,\ldots,x_n}(y) = \min\left[A^{(1)}_{i1,j}(x_1), A^{(2)}_{i2,j}(x_2), \ldots, A^{(n)}_{in,j}(x_n), B_{i,j}(y)\right]$$

Overall output

$$OUT_{x_1,x_2,\ldots,x_n}(y) = \max_{j\in\{1,2,\ldots,k\}} \min\left[A^{(1)}_{i1,j}(x_1), A^{(2)}_{i2,j}(x_2), \ldots, A^{(n)}_{in,j}(x_n), B_{i,j}(y)\right]$$

- Defuzzification
  - Max membership → select de_y where $\boldsymbol{B'}(de\_y) \geq \boldsymbol{B'}(de\_y')$ for all $de\_y' \in Y$
    if $M = \{y \in [y_1, y_2] \mid \boldsymbol{B'}(y) = h(\boldsymbol{B'})\}$ where $h(\boldsymbol{B'})$ is a height of $\boldsymbol{B'}$

$$de\_y = \frac{\sum_{y_k \in M} y_k}{|M|}$$

  - Center of area (centroid)

$$de\_y = \frac{\sum_k B'(y_k)\, y_k}{\sum_k B'(y_k)}$$

  - Simplified centroid

$$de\_y = \frac{\sum_{j \in \mathrm{N}_n} c_j\, y^0_{B_j}}{\sum_{j \in \mathrm{N}_n} c_j}$$

where c_j is firing strength (a_{ij} or $r_j(\boldsymbol{A'})$) and $y^0{}_{\boldsymbol{B}j}$ is the mode of output of rule j

- Example

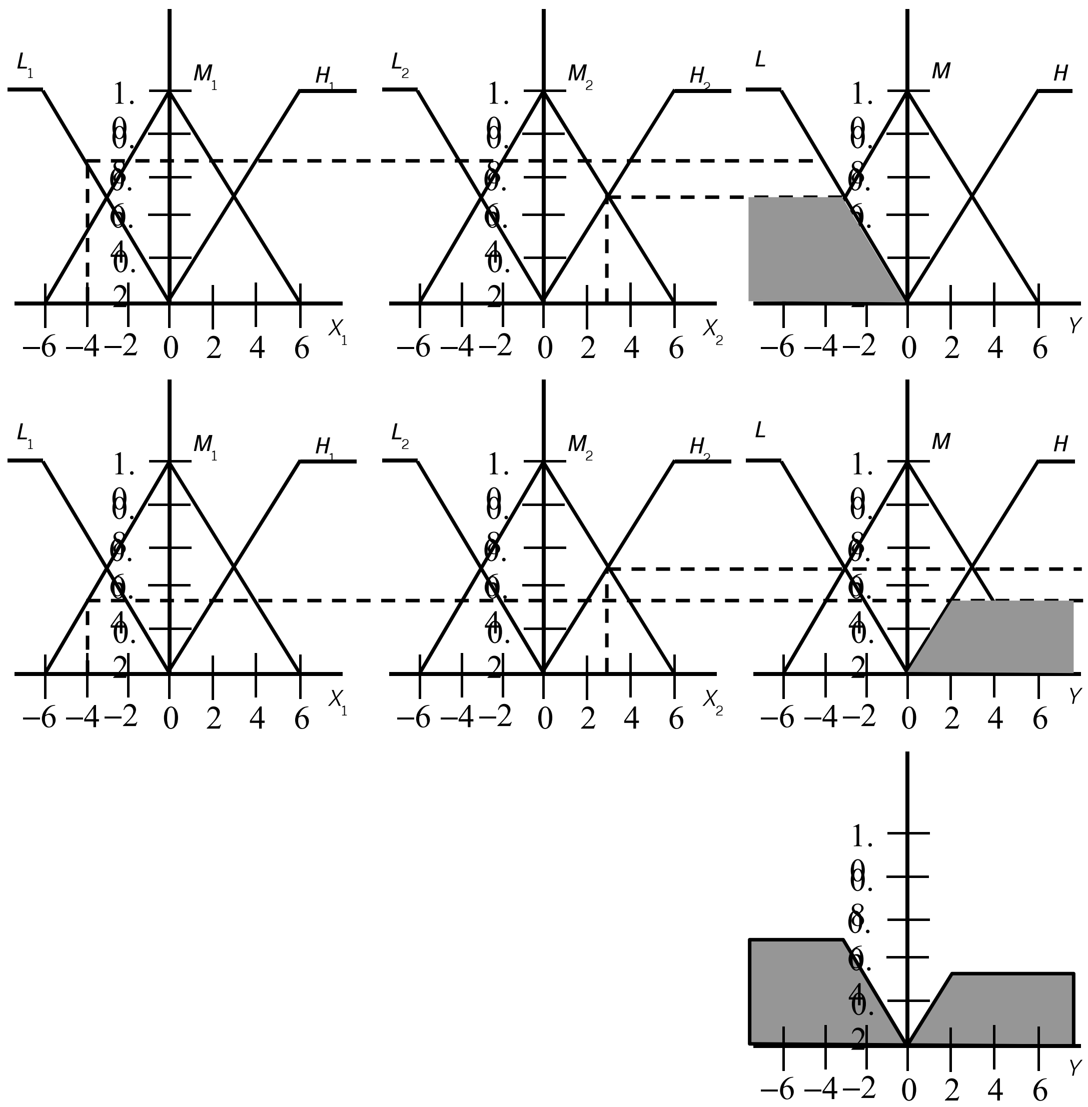
  rule 1: If x_1 is $\boldsymbol{L}_1$ และ x_2 is $\boldsymbol{H}_2$, then y is $\boldsymbol{L}$

  rule 2: If x_1 is $\boldsymbol{M}_1$ และ x_2 is $\boldsymbol{M}_2$, then y is $\boldsymbol{H}$

  input (-4,3)

$\alpha_1 = \min(\boldsymbol{L}_1(\text{-}4), \boldsymbol{H}_2(3)) = \min(0.67, 0.5) = 0.5$

$\alpha_2 = \min(\boldsymbol{M}_1(\text{-}4), \boldsymbol{M}_2(3)) = \min(0.33, 0.5) = 0.5$



- Takagi-Sugeno model

If ξ_1 is $\boldsymbol{A}^{(1)}$, and ξ_2 is $\boldsymbol{A}^{(2)}$ and ... and ξ_n is $\boldsymbol{A}^{(n)}$ then η is $\boldsymbol{B}$

where $\boldsymbol{A}^{(1)}$, $\boldsymbol{A}^{(2}$ and ... and $\boldsymbol{A}^{(n)}$ are linguistic terms in $\boldsymbol{T}(x_i)$ for $1 \le i \le n$ and $\boldsymbol{B}$ is linguistic term in $\boldsymbol{T}(y)$ and suppose there are more than 1 rule

Let $\boldsymbol{T}(x_1) \rightarrow \boldsymbol{A}_1^{(1)}, \boldsymbol{A}_2^{(1)}, \ldots, \boldsymbol{A}_{N1}^{(1)}$

$\boldsymbol{T}(x_2) \rightarrow \boldsymbol{A}_1^{(2)}, \boldsymbol{A}_2^{(2)}, \ldots, \boldsymbol{A}_{N2}^{(2)}$

$\vdots$

$\boldsymbol{T}(x_n) \rightarrow \boldsymbol{A}_1^{(n)}, \boldsymbol{A}_2^{(n)}, \ldots, \boldsymbol{A}_{Nn}^{(n)}$

rule j: If ξ_1 is $\boldsymbol{A}_{i1,j}^{(1)}$, and ξ_2 is $\boldsymbol{A}_{i2,j}^{(2)}$ and ... and ξ_n is $\boldsymbol{A}_{in,j}^{(n)}$ then

$$\eta_j = f_j(\xi_1, \xi_2, \ldots, \xi_n)$$

where $i1 \in \{1,2,\ldots,N1\}$, $i2 \in \{1,2,\ldots,N2\}, \ldots, in \in \{1,2,\ldots,Nn\}$ and

$$f_j(x_1, x_2, \ldots, x_n) = a_1^j x_1 + a_2^j x_2 + \ldots + a_n^j x_n + a_0^j$$

overall output $\rightarrow$ $\eta = \dfrac{\sum_{j=1}^{k} \alpha_j f_j(x_1, x_2, \ldots x_n)}{\sum_{j=1}^{k} \alpha_j}$

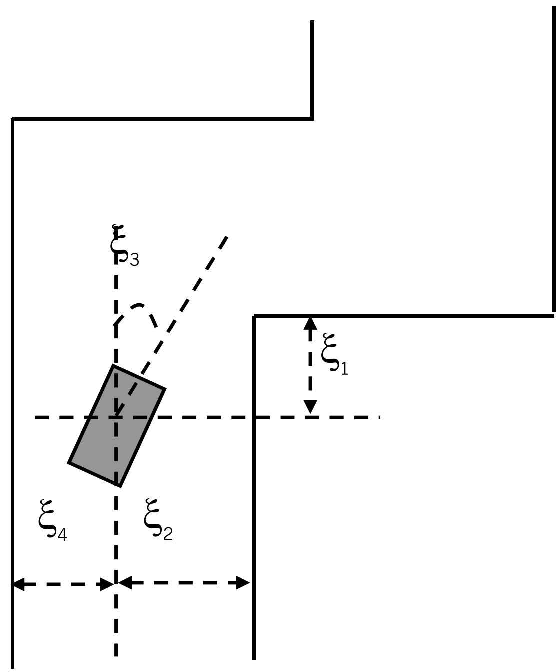
- Example → turning steering wheel control system

input: ξ_1, ξ_2, ξ_3 and ξ_4

universal set: X_1=[0, 150]cm, X_2=[0, 150] cm,

X_3=[−90, 90] degree and X_4=[0, 150] cm

output: rotation speed (η) of the steering wheel



Rule j: If ξ_1 is $\boldsymbol{A}_{i1,j}^{(1)}$, และ ξ_2 is $\boldsymbol{A}_{i2,j}^{(2)}$ และ ξ_3 is $\boldsymbol{A}_{j3,j}^{(3)}$ และ ξ_4 is $\boldsymbol{A}_{j4,j}$ then $\eta = p_0 + p_1\xi_1 + p_2\xi_2 + p_3\xi_3 + p_4\xi_4$

rule	ξ_1	ξ_2	ξ_3	ξ_4	p_0	p_1	p_2	p_3	p_4
1	–	–	outwards	small	3.000	0.000	0.000	-0.045	-0.004
2	–	–	forward	small	3.000	0.000	0.000	-0.030	-0.090
3	small	small	outwards	–	3.000	-0.041	0.004	0.000	0.000
4	small	small	forward	–	0.303	-0.026	0.061	-0.050	0.000
5	small	small	inwards	–	0.000	-0.025	0.070	-0.075	0.000
6	small	big	outwards	–	3.000	-0.066	0.000	-0.034	0.000
7	small	big	forward	–	2.990	-0.017	0.000	-0.021	0.000
8	small	big	inwards	–	1.500	0.025	0.000	-0.050	0.000
9	medium	small	outwards	–	3.000	-0.017	0.005	-0.036	0.000
10	medium	small	forward	–	0.053	-0.038	0.080	-0.034	0.000
11	medium	small	inwards	–	-1.220	-0.016	0.047	-0.018	0.000
12	medium	big	outwards	–	3.000	-0.027	0.000	-0.044	0.000
13	medium	big	forward	–	7.000	-0.049	0.000	-0.041	0.000
14	medium	big	inwards	–	4.000	-0.025	0.000	-0.100	0.000
15	big	small	outwards	–	0.370	0.000	0.000	-0.007	0.000
16	big	small	forward	–	-0.900	0.000	0.034	-0.030	0.000
17	big	small	inwards	–	-1.500	0.000	0.005	-0.100	0.000
18	big	big	outwards	–	1.000	0.000	0.000	-0.013	0.000
19	big	big	forward	–	0.000	0.000	0.000	-0.006	0.000
20	big	big	inwards	–	0.000	0.000	0.000	-0.010	0.000

Input

$\xi_1 = 10$ cm, $\xi_2 = 30$ cm, $\xi_3 = 0$ degree (straight forward) and $\xi_4 = 50$ cm

Only rule 4 and 7→ fire

$\alpha_4 = 0.25$ and $\eta_4 = 0.303 - 0.026(10) + 0.061(30) - 0.050(0) + 0.000(50) = 1.873$

$\alpha_7 = 0.167$ and $\eta_7 = 2.990 - 0.017(10) + 0.000(30) - 0.021(0) + 0.000(50) = 2.820$

Output of the system will be

$$\eta = \frac{0.25(1.873) + 0.167(2.820)}{0.25 + 0.167} = 2.252$$

- Fuzzy measure

Address the ambiguity axis of uncertainty $\rightarrow$ how likely can the answer to a question be found in various subsets of the sources of information

$X = \{x_1, x_2, \dots x_n\}$, $g: 2^X \rightarrow [0,1]$

Properties

1. $g(\phi) = 0$
2. $g(A) \leq g(B)$ if $A \subseteq B$

$g^i = g\{x_i\}$ $\rightarrow$ fuzzy density function $\rightarrow$ importance of the single information source x_i in determine the answer to a particular question

- Sugeno Lambda measure→satisfy 2 properties with additional one

  3. For all A,B $\subseteq$X with A$\cap$B = ϕ

$$g_\lambda(A \cap B) = g_\lambda(A) + g_\lambda(B) + \lambda g_\lambda(A) g_\lambda(B) \qquad \exists \lambda > -1$$

From $X = \bigcup_{i=1}^{n} x_i$ and $g_\lambda(X) = 1$ →

$$(1+\lambda) = \prod_{i=1}^{n} \left(1 + \lambda g^i\right)$$

$$= 1 + \lambda \sum_{j=1}^{n} g^j + \lambda^2 \sum_{j=1}^{n-1} \left(\sum_{k=j+1}^{n} g^j g^k \right) + \lambda^3 \sum_{j=1}^{n-2} \left(\sum_{k=1}^{n-1} \sum_{i=1}^{n} g^j g^k g^i \right) + \ldots + \lambda^n g^1 g^2 \ldots g^n$$

- Example

X=$\{x_1, x_2, x_3\}$→ fuzzy densities g^1=0.2, g^2=0.3, g^3=0.1

Subset	g_λ
ϕ	0→ from property
$\{x_1\}$	0.2
$\{x_2\}$	0.3
$\{x_3\}$	0.1
$\{x_1, x_2\}$	0.69
$\{x_1, x_3\}$	0.36
$\{x_2, x_3\}$	0.5
$\{x_1, x_2, x_3\}$	1 → from property

find λ using $(1+\lambda)=(1+0.2\lambda)(1+0.3\lambda)(1+0.1\lambda)$

$$0.006\lambda^2+0.11\lambda-0.4=0$$

$$\lambda=\frac{-0.11\pm\sqrt{0.11^2-4(0.006)(-0.4)}}{2(0.006)}=3.2,-21.44$$

Hence λ =3.2

$g(\{x_1, x_2\})$ = 0.2+0.3+3.2(0.2)(0.3) = 0.69

$g(\{x_1, x_3\})$ = 0.2+0.1+3.2(0.2)(0.1) = 0.36

$g(\{x_2, x_3\})$ = 0.3+0.1+3.2(0.3)(0.1) = 0.5

$g(\{x_1, x_2, x_3\})$ = 0.69+0.1+3.2(0.69)(0.1) = 1.01≈1

- Fuzzy integral
  - Sugeno fuzzy integral

Contiuous case: $\int h(x) \circ g = \sup_{\alpha \in [0,1]} \min[\alpha, g(A_\alpha)]$ where $A_\alpha = \{x | h(x) \geq \alpha\}$

Finite case: X={x_1, x_2,…, x_n }→ reorder $X = \{x_{(\ )}, x_{(2)}, \ldots x_{(n)}\}$ so that $h(x_{(1)}) \geq h(x_{(2)}) \geq \ldots \geq h(x_{(n)})$

Hence, sugeno fuzzy integral → $S_g(h) = \max_{i=1}^{n} \min[h(x_{(i)}), g(A_{(i)})]$ where $A_{(i)} = \{x_{(1)}, x_{(2)}, \ldots x_{(i)}\}$

- Example

X={x_1, x_2, x_3 }→$h(x_1) = 0.7$, $h(x_2) = 0.9$, $h(x_3) = 0.2$ and assume that g are the same as previous example

Reorder $h(x_2) > h(x_1) > h(x_3)$

S_g=(0.9∧g ({x_2 })) ∨ (0.7 ∧g ({ x_1, x_2 })) ∨ (0.2 ∧g ({ x_1, x_2, x_3 })) =0.69

- Choquet fuzzy integral

Contiuous case: $\int_X h(x) \circ g = \int_0^1 g(A_\alpha) d\alpha$ where $A_\alpha = \{x | h(x) \geq \alpha\}$

Finite case: $X=\{x_1, x_2, \ldots, x_n\}$ → reorder $X = \{x_{(\ )}, x_{(2)}, \ldots x_{(n)}\}$ so that $h(x_{(1)}) \geq h(x_{(2)}) \geq \ldots \geq h(x_{(n)})$

Hence, choquet fuzzy integral → $C_g(h) = \sum_{i=1}^{n} \left[h(x_{(i)}) - h(x_{(i+1)})\right] g(A_{(i)})$ where $A_{(i)} = \{x_{(1)}, x_{(2)}, \ldots x_{(i)}\}$ and $h(x_{(n+1)}) = 0$

Let $\delta_i(g) = g(A_i) - g(A_{i-1})$ →

$$
\begin{aligned}
C_g(h) &= \sum_{i=1}^{n} \left[h(x_{(i)}) - h(x_{(i+1)})\right] g(A_{(i)}) \\
&= \left[h(x_{(1)}) - h(x_{(2)})\right] g(A_{(1)}) + \left[h(x_{(2)}) - h(x_{(3)})\right] g(A_{(2)}) + \ldots + \left[h(x_{(n)}) - h(x_{(n+1)})\right] g(A_{(n)}) \\
&= h(x_{(1)}) g(A_{(1)}) + h(x_{(2)}) \left[g(A_{(2)}) - g(A_{(1)})\right] + \ldots + h(x_{(n)}) \left[g(A_{(n)}) - g(A_{(n-1)})\right] - h(x_{(n+1)}) g(A_{(n)})
\end{aligned}
$$

Hence $C_g(h) = \sum_{i=1}^{n} \delta_i(g) h(x_{(i)})$ where $\delta_i(g) = \left[g(A_{(i)}) - g(A_{(i-1)})\right]$

We can have linear order statistic

$$LOS_{w_k}(h(x)) = \sum_{x_k \in w_k} w_k h(x_{(k)})$$

where $\{w_1, w_2, \ldots, w_n\}$ satisfy $w_i \in [0,1]$ and $\sum_{k=1}^{n} w_k = 1$

- Example

X=$\{x_1, x_2, x_3\} \rightarrow h(x_1) = 0.7$, $h(x_2) = 0.9$, $h(x_3) = 0.2$ and assume that g are the same as previous example

Reorder $h(x_2) > h(x_1) > h(x_3)$

C_g=(0.9-0.7)(0.3)+(0.7-0.2)(0.69+(0.2-0.1)(1) = 0.605

- Optimization training

Let training set → T={(o_j, α_j)|j=1,2,…,m } where o_j=j^{th} object and α_j= desired output of j^{th} object, and there are *m* training samples

Suppose→ there are *n* sensors/algorithms

Output from choquet integral of j^{th} object will be

$$C_g\left(h(o_j)\right)=\sum_{i=1}^{n}\left[h\left(o_j;x_{(i)}\right)-h\left(o_j;x_{(i+1)}\right)\right]g\left(A_{(i)}\right)$$

Want to find fuzzy measure

$$\vec{g}=\begin{bmatrix} g_1 \\ g_2 \\ \cdots \\ g_n \\ g_{12} \\ g_{13} \\ \cdots \\ g_X \end{bmatrix}$$

such that $C_g(h(o_j))$ is close to α_j → $E^2=\sum_{j=1}^{m}\left[C_g\left(h\left(o_j\right)\right)-\alpha_j\right]^2$

No	$h(x_1)$	$h(x_2)$	…	$h(x_n)$	α
1					
2					
…					
m					

Training set with *m* samples each has desired output (α). There are *n* sensors/algorithms

Let $$v_{j,A} = \begin{cases} h(o_j; x_{(i)}) - h(o_j; x_{(i+1)}) & \text{if } A = A_i \ \exists i \\ 0 & \text{else} \end{cases}$$

We have
$$\vec{v}_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ h(o_j; x_{(1)}) - h(o_j; x_{(2)}) \\ \vdots \\ h(o_j; x_{(n-1)}) - h(o_j; x_{(n)}) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\vec{v}_j \in R^{2n-2}$ with only $n-1$ possible nonzero

Then
$$C_g\big(h(o_j)\big) = \vec{v}_j^{\,T} \vec{g} + h(o_j; x_{(n)})$$

Objective function

$$
\begin{aligned}
E^2 &= \frac{1}{2}\sum_{j=1}^{m}\left[C_g\left(h\left(o_j\right)\right)-\alpha_j\right]^2 \\
&= \frac{1}{2}\sum_{j=1}^{m}\left[\vec{v}_j{}^T\vec{g}+h\left(o_j;x_{(n)}\right)-\alpha_j\right]^T\left[\vec{v}_j{}^T\vec{g}+h\left(o_j;x_{(n)}\right)-\alpha_j\right] \\
&= \frac{1}{2}\sum_{j=1}^{m}\left[\vec{g}^T\vec{v}_j\vec{v}_j{}^T\vec{g}+2\left[h\left(o_j;x_{(n)}\right)-\alpha_j\right]\vec{v}_j{}^T\vec{g}+\left[h\left(o_j;x_{(n)}\right)-\alpha_j\right]^2\right] \\
&= \frac{1}{2}\vec{g}^T\left[\sum_{j=1}^{m}\vec{v}_j\vec{v}_j{}^T\right]\vec{g}+\left[\sum_{j=1}^{m}\left[h\left(o_j;x_{(n)}\right)-\alpha_j\right]\vec{v}_j{}^T\right]\vec{g}+\frac{1}{2}\sum_{j=1}^{m}\left[h\left(o_j;x_{(n)}\right)-\alpha_j\right]^2
\end{aligned}
$$

Hence

$$
E^2 = \frac{1}{2}\vec{g}^T\mathbf{D}\vec{g}+\vec{v}^T\vec{g}+\beta^2 \quad \text{where } \mathbf{D}=\sum_{j=1}^{m}\vec{v}_j\vec{v}_j{}^T, \vec{v}=\sum_{j=1}^{m}\left[h\left(o_j;x_{(n)}\right)-\alpha_j\right]\vec{v}_j{}^T \text{ and } \beta^2=\frac{1}{2}\sum_{j=1}^{m}\left[h\left(o_j;x_{(n)}\right)-\alpha_j\right]^2
$$

Similar to

$$
\underset{\vec{g}}{\text{minimize}}\ \frac{1}{2}\vec{g}^T\mathbf{D}\vec{g}+\vec{v}^T\vec{g}
$$

$$
\text{subject to } \mathbf{A}\,\vec{g}\le\vec{b} \quad \text{and } \vec{0}\le\vec{g}\le\vec{1}
$$

From

$g_1 - g_{12} \leq 0$

$g_1 - g_{13} \leq 0$

$\cdots$

$g_1 - g_{12\ldots n} \leq 0$

$\cdots$

$g_{123\ldots n\text{-}1} - g_{123\ldots n} \leq 0$

$\cdots$

$g_{23\ldots n} - g_{123\ldots n} \leq 0$

$$\rightarrow \quad \mathbf{A} = \begin{bmatrix} 1 & 0\cdots 0 & -1 & 0\cdots & 0\cdots & & 0 \\ 1 & 0\cdots 0 & 0 & -1 & 0\cdots & & 0 \\ \vdots & & & & & & \\ 0 & 0\cdots 0 & \cdots & & 0 & 1\,0\cdots & 0 \\ \vdots & & & & & & \\ 0 & 0\cdots 0 & \cdots & & 0 & 0\,0\cdots & 1 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

- Example

Let

No	$h(x_1)$	$h(x_2)$	$h(x_3)$	α
1	0.68	0.53	0.81	0.743
2	0.74	0.99	0.86	0.926
3	0.45	0.07	0.08	0.301

find $\vec{g} = \begin{bmatrix} g_1 \\ g_2 \\ g_3 \\ g_{12} \\ g_{13} \\ g_{23} \end{bmatrix}$ such that E^2 is minimized

O_1: $C_g(O_1) = (0.81 - 0.68)\, g_3 + (0.68 - 0.53)\, g_{13} + (0.53) = 0.13 g_3 + 0.15\, g_{13} + 0.53$

O_2: $C_g(O_2) = (0.99 - 0.86)\, g_2 + (0.86 - 0.74)\, g_{23} + (0.74) = 0.13 g_2 + 0.12\, g_{23} + 0.74$

O_3: $C_g(O_3) = (0.45 - 0.08)\, g_1 + (0.08 - 0.07)\, g_{13} + (0.07) = 0.37 g_3 + 0.01\, g_{13} + 0.07$

Hence

$$\vec{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 0.13 \\ 0 \\ 0.15 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ 0.13 \\ 0 \\ 0 \\ 0 \\ 0.12 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0.37 \\ 0 \\ 0 \\ 0 \\ 0.01 \\ 0 \end{bmatrix}$$

And

$$\mathbf{D} = \sum_{j=1}^{3} \vec{v}_j \vec{v}_j^T = \begin{bmatrix} 0.14 & 0 & 0 & 0 & 0.004 & 0 \\ 0 & 0.017 & 0 & 0 & 0 & 0.016 \\ 0 & 0 & 0.017 & 0 & 0.02 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0.004 & 0 & 0.02 & 0 & 0.023 & 0 \\ 0 & 0.016 & 0 & 0 & 0 & 0.014 \end{bmatrix}$$

$$\vec{v} = \sum_{j=1}^{3} \left[h\left(o_j ; x_{(n)}\right) - \alpha_j \right] \vec{v}_j^{\,T} = \begin{bmatrix} -0.086 \\ -0.024 \\ -0.028 \\ 0 \\ -0.034 \\ -0.022 \end{bmatrix}$$

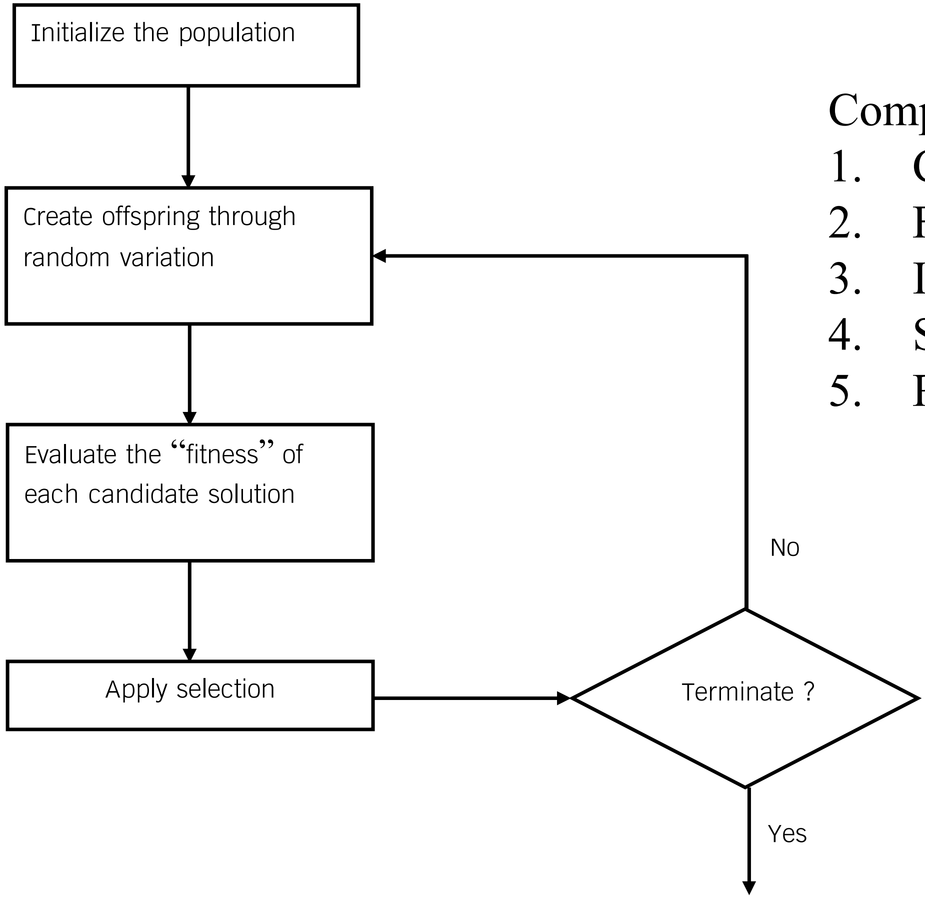
$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Now we can compute

$$\underset{\vec{g}}{\text{minimize}} \frac{1}{2} \vec{g}^T \mathbf{D} \vec{g} + \vec{v}^T \vec{g}$$

subject to $\mathbf{A}\,\vec{g} \le \vec{b}$ and $\vec{0} \le \vec{g} \le \vec{1}$

Evolutionary computing



Component

1. Chromosome encoding
2. Fitness value or survival strength of individual
3. Initial population
4. Selection operator
5. Reproduction operator

Without mutation→population tends to converge to a homogeneous state where individuals vary only slightly from each other

- Genetic Algorithm

  Algo

  1. Set t=0
  2. Initial population $P(t)$
  3. Calculate each individual fitness value
  4. While not converge do
     1. Select individual to P^1 (intermediate population)→Mating Pool (MP)
     2. Select from MP to mate → P^2 → mutate chromosome in P^2 → P^3
     3. Select chromosome in P^3 and $P(t)$ for replacement→ $P(t+1)$
     4. Set t= t+1
  5. End while

If $(|P^3|)=N$ then P^3 become $P(t+1)$ else if $(|P^3|)<N$ select q missing chromosome from $P(t)$ or P^1

• Example

Chromosome *A*, *B*, *C*, *D* → each with 8 genes

Fitness function: number of 1 in the string

chromosome label	chromosome string	fitness value
A	00000110	2
B	11101110	6
C	00100000	1
D	00110100	3
Average fitness value		12/4

P(*t*)

Crossover
B and *D*
at crossing site: 1 get *E* and *F*
And *B* and *C* are copied

chromosome label	chromosome string	fitness value
B	11101110	6
C	00100000	1
E	10110100	4
F	01101110	5

B is mutated at 1st bit
E is mutated at 6th bit

P(*t*+1)

chromosome label	chromosome string	fitness value
B′	01101110	5
C	00100000	1
E′	10110000	3
F	01101110	5
Average fitness value		14/4

- Initial population
  - random gene values of each gene in each chromosome→uniform representation of the entire search space
  - Small population → small part of search space, time complexity is low, need more generations to converge→EA force to explore a larger search space by increasing the rate of mutation
  - Large population→ large area of search space, less generation to converge, time complexity is increased
- Selection operation
  - Selection techniques exist
    - Explicit fitness remapping→fitness values is mapped into a new range, e.g., normalization to [0,1]
    - Implicit fitness remapping→use actual fitness values for selection several selection operators

- Random selection

random→no reference to fitness→ each individual has an equal chance of being selected

- Proportional selection

$$P(t)=\left\{x^1,x^2,\ldots,x^N\right\}$$ total fitness $$F=\sum_{i=1}^{N} f\left(x^i\right)$$ where $f(x^i)$→fitness

value of chromosome x^i →selection probability of x^i

$$p_i=\frac{f\left(x^i\right)}{F} \quad \text{for } i=1,2,\ldots,N$$

Expected number of copied of x^i $\quad n_i = Np_i \quad \text{for } i=1,2,\ldots,N$

Or

$$n_i = N\frac{f\left(x^i\right)}{F}=\frac{f\left(x^i\right)}{\overline{f}} \quad \text{for } i=1,2,\ldots,N$$

- Algo

For i=1 to N

Calculate $q_i = \sum_{k=1}^{i} p_k \quad \text{for } i = 1, 2, \ldots, N$

End for

For i=1 to N

random $\xi \in [0,1]$ → uniform distribution

If $0 \le \xi \le q_1$, chromosome x^1 is selected

If $q_{i-1} \le \xi \le q_i$ for i=1,2,…N then chromosome x^i *is* selected

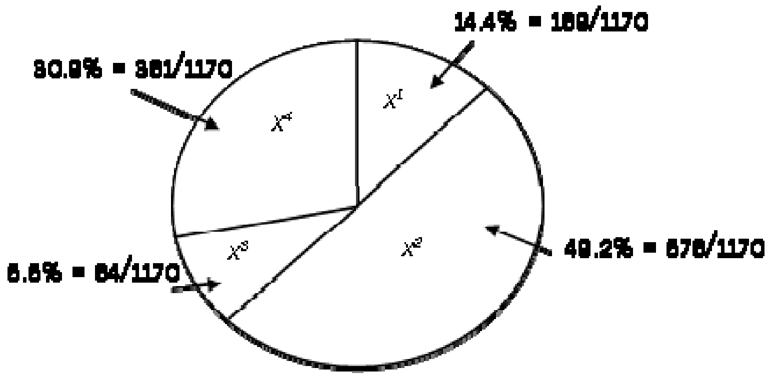
End for

- Example

$$\max_x f(x) = \max_x x^2$$ with x is varied between 0-31,

assume random ξ 4 times with the value of 0.8, 0.5, 0.1 and 0.6

No	Chromosome	x (in real number)	f(x)	p_i	q_i	N_i	Number of copies
1	01101	13	169	0.14	0.14	0.58	1
2	11000	24	576	0.49	0.63	1.97	2
3	01000	8	64	0.06	0.69	0.22	0
4	10011	19	361	0.31	1.00	1.23	1
		Total	1170	1.00			
		Average	293	0.25			
		maximum	576	0.49			



Crossover → x^1 and x^2 at crossover site 4

x^3 and x^4 at crossover site 2

Mutation probability is 0.001→ there are 20 bits→ expected number of bit undergoes mutation is 20(0.001) = 0.02 bits → suppose no mutation

Hence, New population will be

No	Chromosome (x)	mate	Crossover site	New population	X (in real number)	f(x)
1	01101	2	4	01100	12	144
2	11000	1	4	11001	25	625
3	11000	4	2	11011	27	729
4	10011	3	2	10000	16	256
					Total	1754
					Average	439
					maximum	729

- Disadvantage
  - Premature convergence
  - Slow convergence