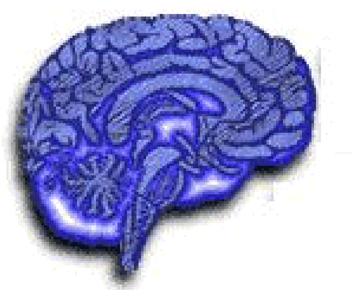
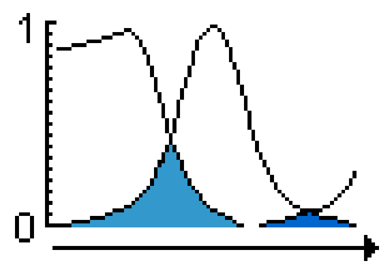
# Introdução à LÓGICA DIFUSA






**Prof. Doutor Paulo Salgado**

2019 @ UTAD

psal@utad.pt

## Tópicos do Curso

- **Introdução à LÓGICA DIFUSA (*Fuzzy Logic*).**
- **Operações Lógicas e Inferência difusa**
- **Técnicas automáticas de geração de regras difusas.**
- **Exemplo prático de aplicação: Controlo do Movimento de um Robô Móvel em ambientes complexos.**

# LÓGICA

*Lógica* é o estudo do método da “razão” em todas as suas formas possíveis.

*Lógica proposicional* (área da lógica) usa combinações de variáveis (lógicas) de acordo com proposições arbitrárias.

Uma simples proposição pode ser expressa em geral, na forma canónica:

***x*** is *P*

em que *x* designa o *sujeito* e *P* designa o *predicado*.

A lógica clássica utiliza *proposições* que assumem apenas dois valores: *verdadeiro* ou *falso*.

***P*** (o predicado) desempenha o papel de uma função em ***X***.

Esta função é geralmente chamada de *função predicado* e é representada por $P(x)$, assumindo apenas dois valores, *verdadeiro* ou *falso*, quando um dado *sujeito* em ***X*** é substituído em $x$.

**Exemplo:**

**"coffee *is Hot*"**

***Hot*(coffee)=**

*Verdadeiro*

ou

*falso*.

# Teória dos Conjuntos Bivalente

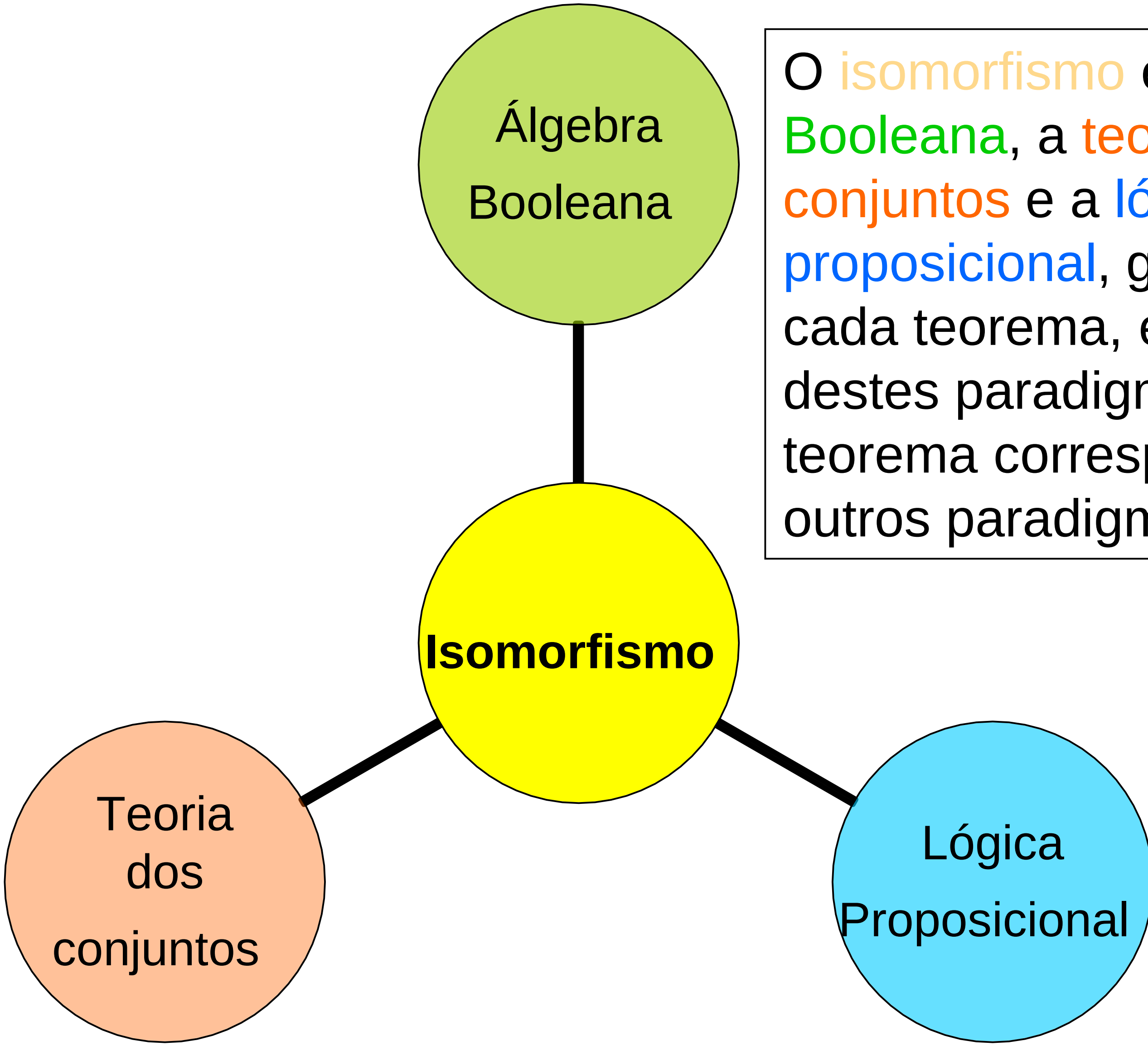
Um conjunto é definido de tal forma que dicotomiza um ponto individual de um dado universo em dois grupos:

- os membros, que são aqueles que seguramente lhe pertencem, (grau de pertença unitário, 1)
- os não-membros, que são aqueles que de certeza não lhe pertencem (grau de pertença nulo, 0).

Exemplo:

Seja $\boldsymbol{X}$ = { A, B, C, …, Z, W }, o conjunto das letras do abecedário e o conjunto das vogais: **Vogais** = { A, E, I, O, U }

$$A \in \textbf{Vogais} \Leftrightarrow \mu_{\textbf{Vogais}}(A) = 1$$




O isomorfismo entre a álgebra Booleana, a teoria dos conjuntos e a lógica proposicional, garante que para cada teorema, em qualquer um destes paradigmas, existe um teorema correspondente nos outros paradigmas.

## CONJUNTOS DIFUSOS

O conceito de conjunto “difuso” designa um conjunto em que a cada membro está associado um grau de pertença; o grau de pertença é uma variável contínua, em que valores elevados evidenciam um alto grau de pertença ao conjunto, enquanto valores baixos denotam um diminuto grau de pertença.

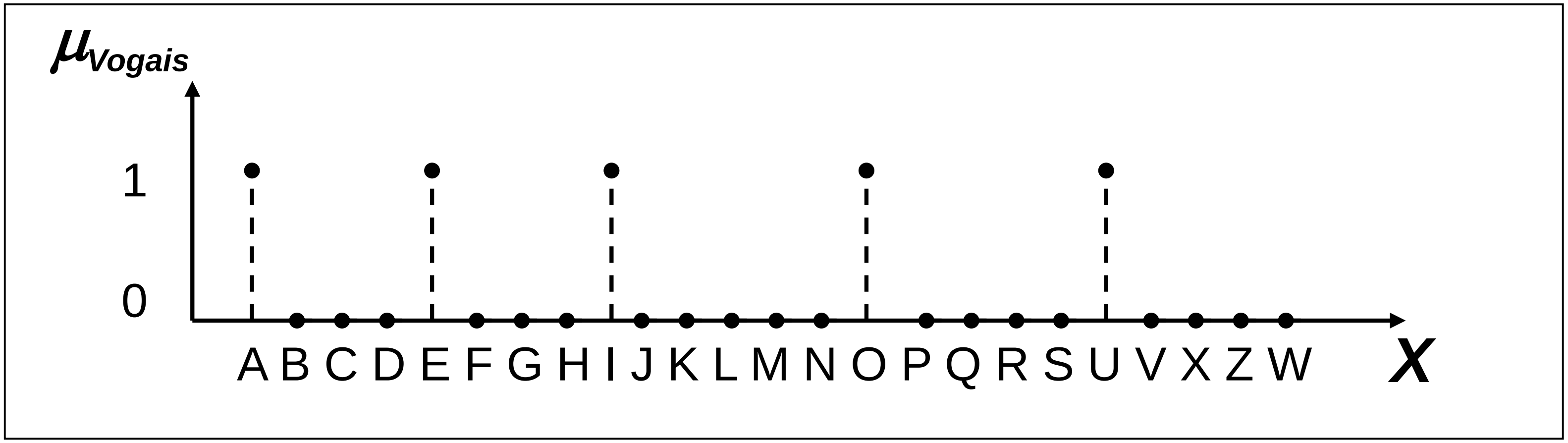
Seja *A* um conjunto difuso definido em *X*.

A função de pertença $\mu_A$ é definida como:

$$\boxed{\mu_A : X \longrightarrow [0,1]}$$

Seja $\boldsymbol{X}$={ A, B, C, …, Z, W}, o conjunto das letras do abecedário

O conjunto das vogais: ***Vogais*** = {A, E, I, O, U }, pode ser definido pela função de pertença:

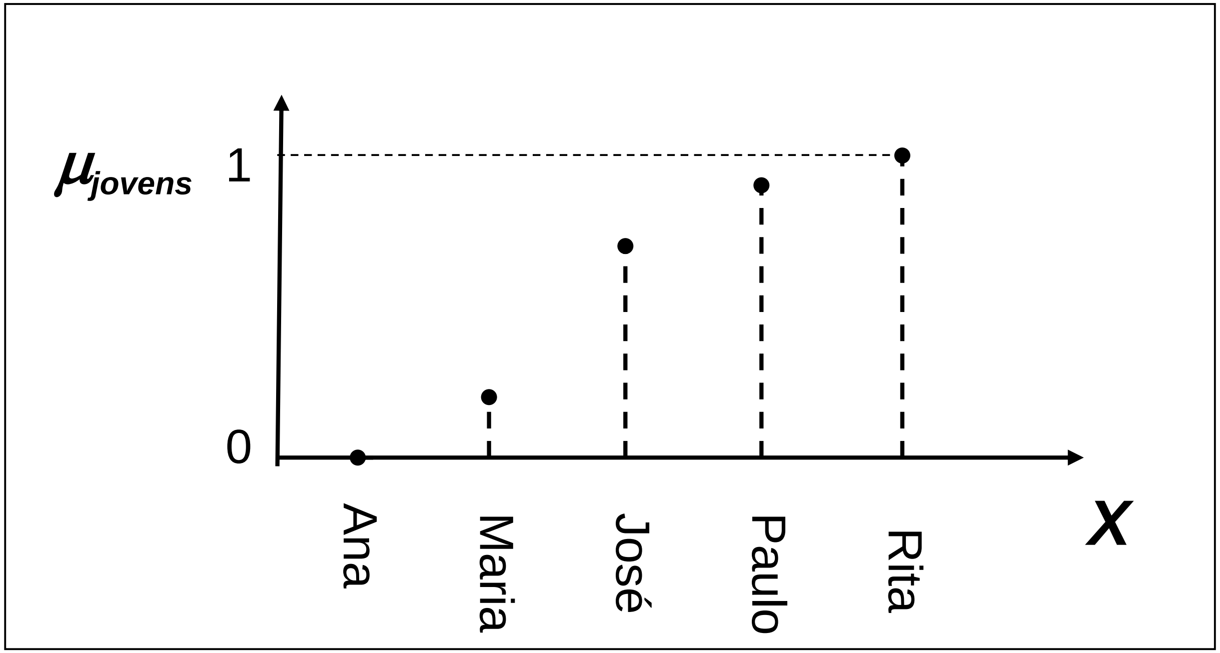



***Vogais*** = {A/1; B/0; C/0; D/0; E/1; …, H/0; I/1;…; N/0; O/1,…, U/1, V/0,…,W/0}

$$\boxed{\mu_A : X \longrightarrow \{0,1\}}$$

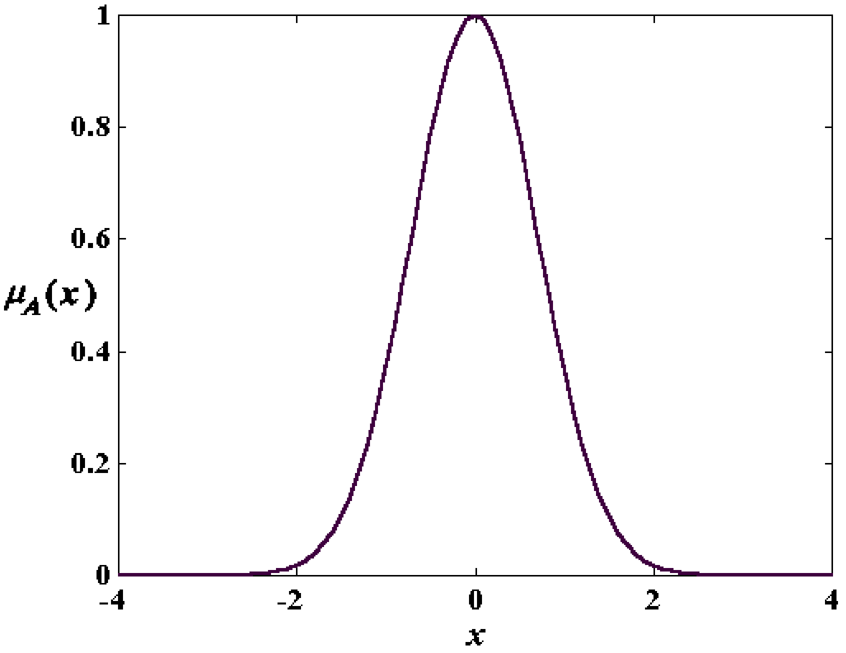
Seja um conjunto de alunos { Rita (1 ano); Paulo (15 anos); José (25 anos); Maria (50 anos), Ana (70 anos)}.

Uma forma possível de caracterizar o grau de "juventude" no Universo do discurso, seria:




***Jovens***={ Ana/**0.0**; Maria/**0.2**; José/**0.7**; Paulo/**0.9**, Rita/**1.0**}.

Exemplo: função pertença de um conjunto difuso de números reais próximos de 0, ver figura.




$$\mu_A(x) = \exp\left(-x^2\right)$$

Exemplo: A temperatura atmosférica pode ser interpretada como uma variável linguística, cujo conjunto dos seus termos T(temperatura) pode ser:

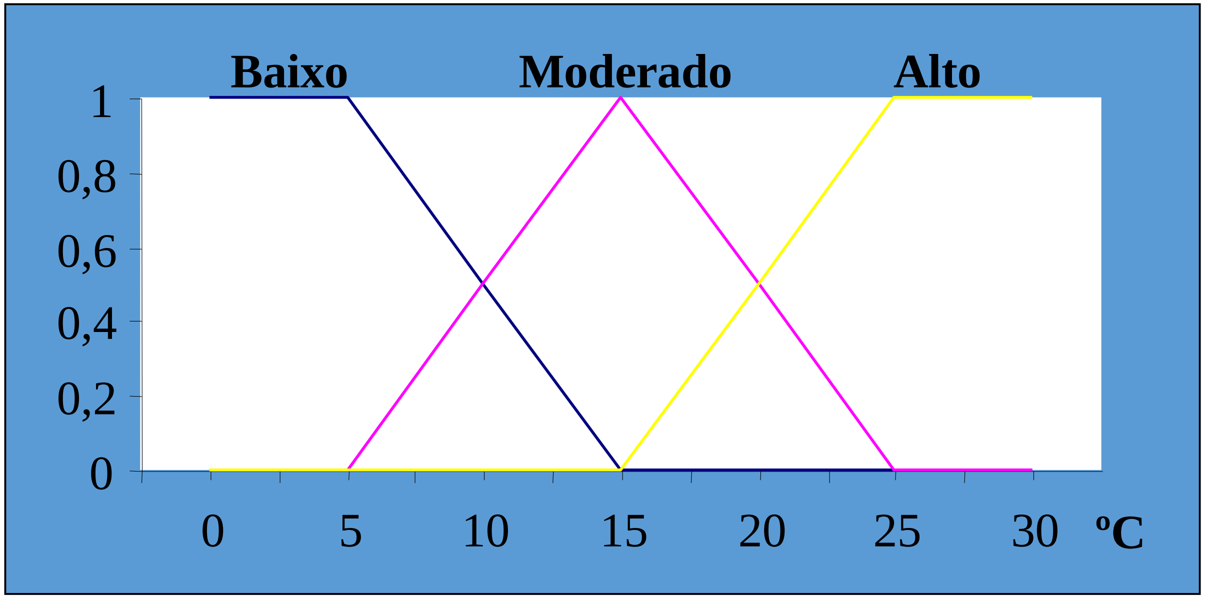
$$\mathrm{T}\left(\mathrm{Temperatura}\right) = \left\{\text{"baixo", "moderado","alto"}\right\}$$

em que, cada termo T é caracterizado por um conjunto difuso no universo do discurso U=[-10º, 40º].

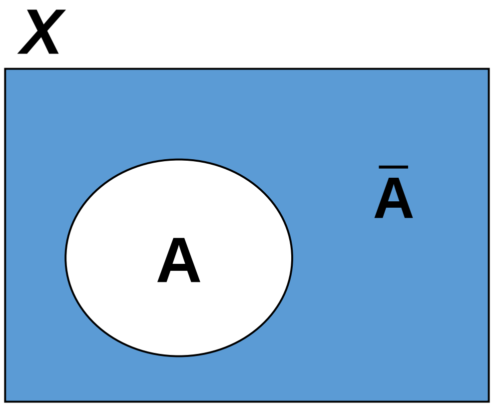
Podemos interpretar:

- "baixa" a temperatura abaixo dos 5º C;
- "moderada" em torno dos 15ºC;
- "alta" para temperaturas superiores a 25ºC.

*Podemos associar a esses termos os conjuntos difusos cujas funções de pertença estão representado na figura.*



## OPERAÇÕES SOBRE CONJUNTOS




Na lógica bivalente tem-se que:

$$\mathbf{A} \cup \bar{\mathbf{A}} = \mathbf{X} \qquad \mathbf{A} \cap \bar{\mathbf{A}} = \varnothing$$

Na lógica difusa geralmente tem-se que:

$$\mathbf{A} \cap \bar{\mathbf{A}} \neq \varnothing \qquad \mathbf{A} \cup \bar{\mathbf{A}} \neq \mathbf{X}$$

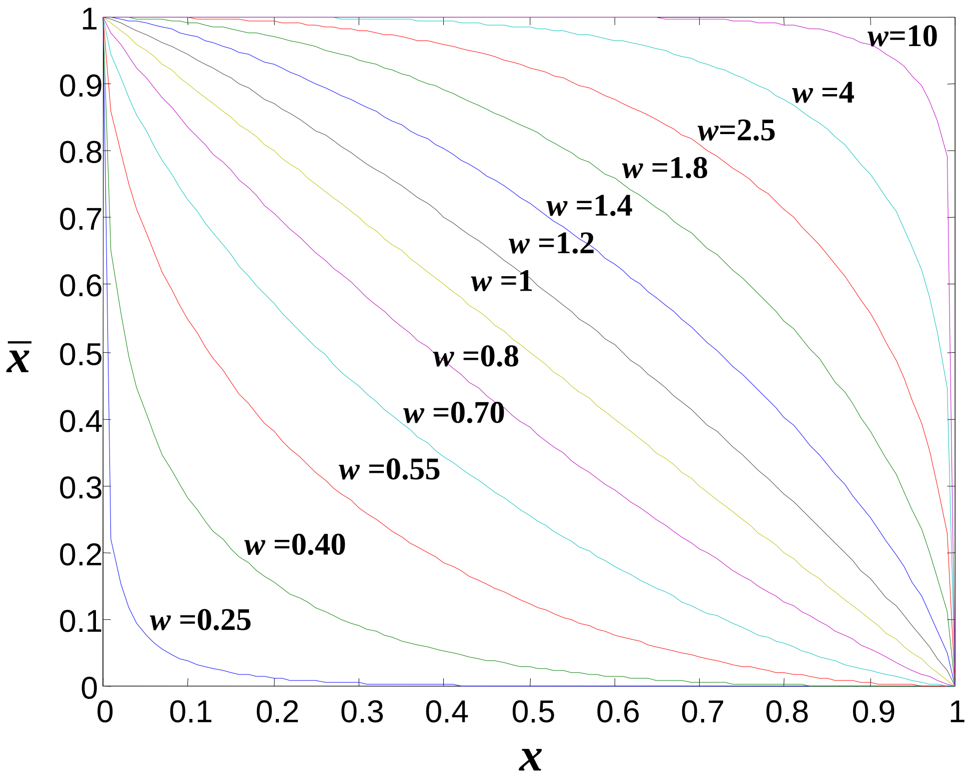
# Função Complemento

**Definição**: A função complemento do conjunto A, cA ou , pode ser definida pela função

c: [0,1] → [0,1]

que assinala um valor cada valor de pertença de um qualquer conjunto difuso *A*.

| | |
|---|---|
| $\bar{a} = 1 - a$ | Complemento de 1 |
| $\bar{a} = \dfrac{1-a}{1+\lambda a}, \quad -1 < \lambda < \infty$ | Complemento λ (Classe Sugeno) |
| $\bar{a} = \left(1 - a^{w}\right)^{\frac{1}{w}}, \quad 0 < w < \infty$ | Complemento *w* (Classe de Yager) |



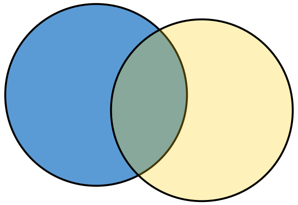

Funções Complemento *w* (Classe de Yager) para vários valores de *w*

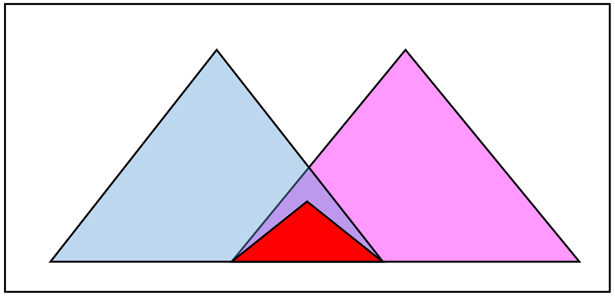
# Função de Intersecção de conjuntos difusos

**Definição**: A intersecção de dois conjuntos difusos A e B é especificada por uma operação binária no intervalo unitário

$$* : [0,1] \times [0,1] \to [0,1]$$

ou seja, para cada elemento *x* do conjunto Universo *X*, esta função toma como argumentos os par de pertença de *x* em *A* e *B*.

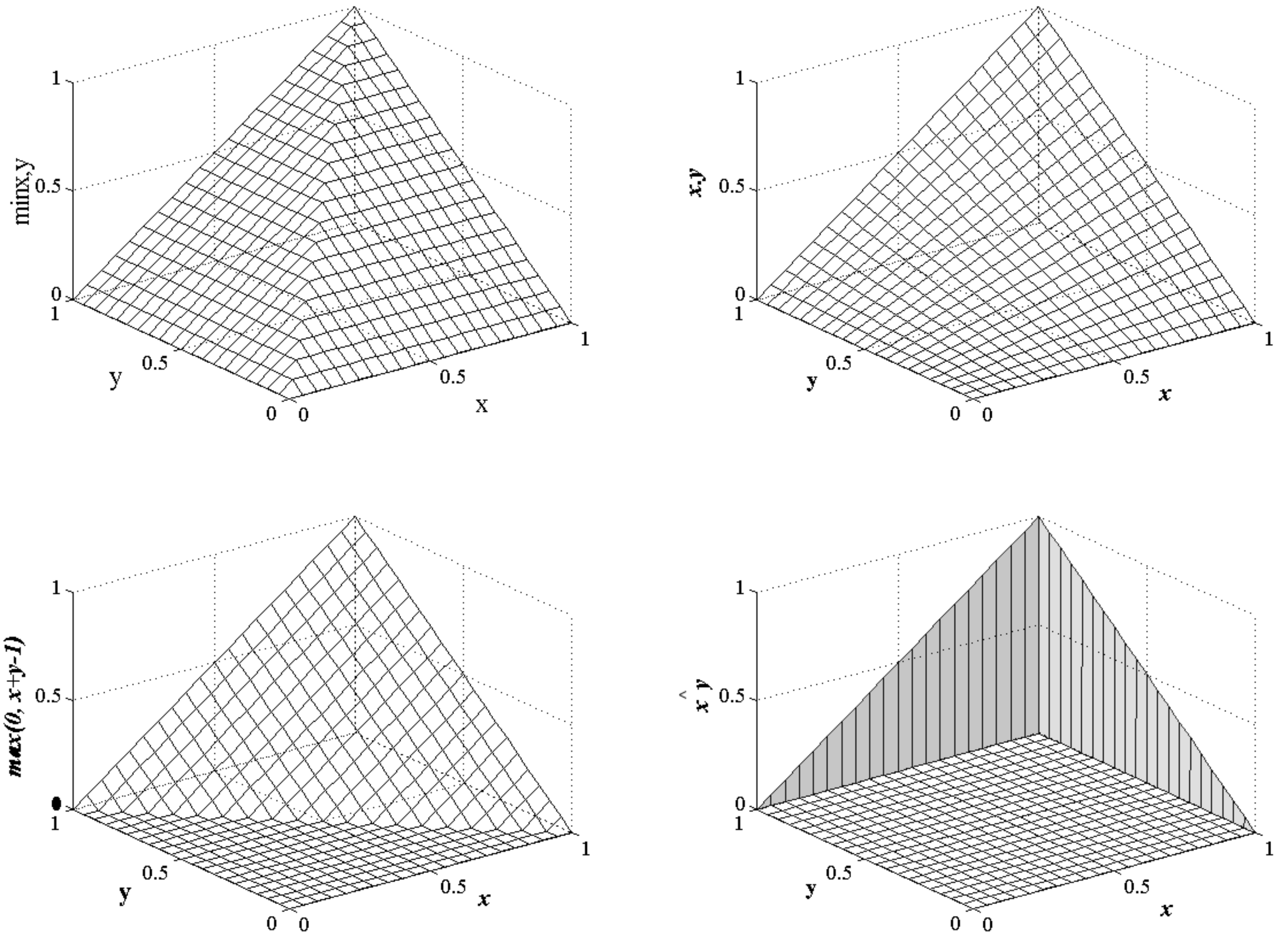




Resulta da aplicação desta função um grau de pertença de *x* no conjunto de intersecção de *A* com *B*, dado por:

$$\mu_{A \cap B}(x) = \mu_A(x) * \mu_B(x)$$

| | |
|---|---|
| $a \wedge b = min\{a,b\}$ | Produto Lógico ou Intersecção |
| $a \bullet b = a\ b$ | Produto Algébrico |
| $b\ = max\{\ 0\ ,\ a+b\text{-}1\}$ | Produto Limitado |
| $a \overset{\wedge}{\bullet} b = \begin{cases} a & se\ \ b = 1 \\ b & se\ \ a = 1 \\ 0 & se\ a,b < 1 \end{cases}$ | Produto Drástico |

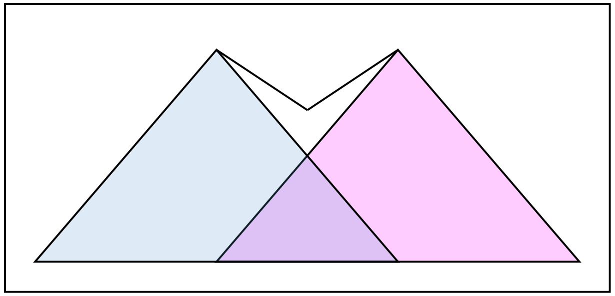

minx,y
x,y
max(0, x+y-1)
x ^ y
x
y
0
0.5
1

# Função de Reunião de conjuntos difusos

**Definição**: A união de dois conjuntos difusos *A* e *B* é especificado por uma operação binária no intervalo unitário

$$* : [0,1] \times [0,1] \to [0,1]$$

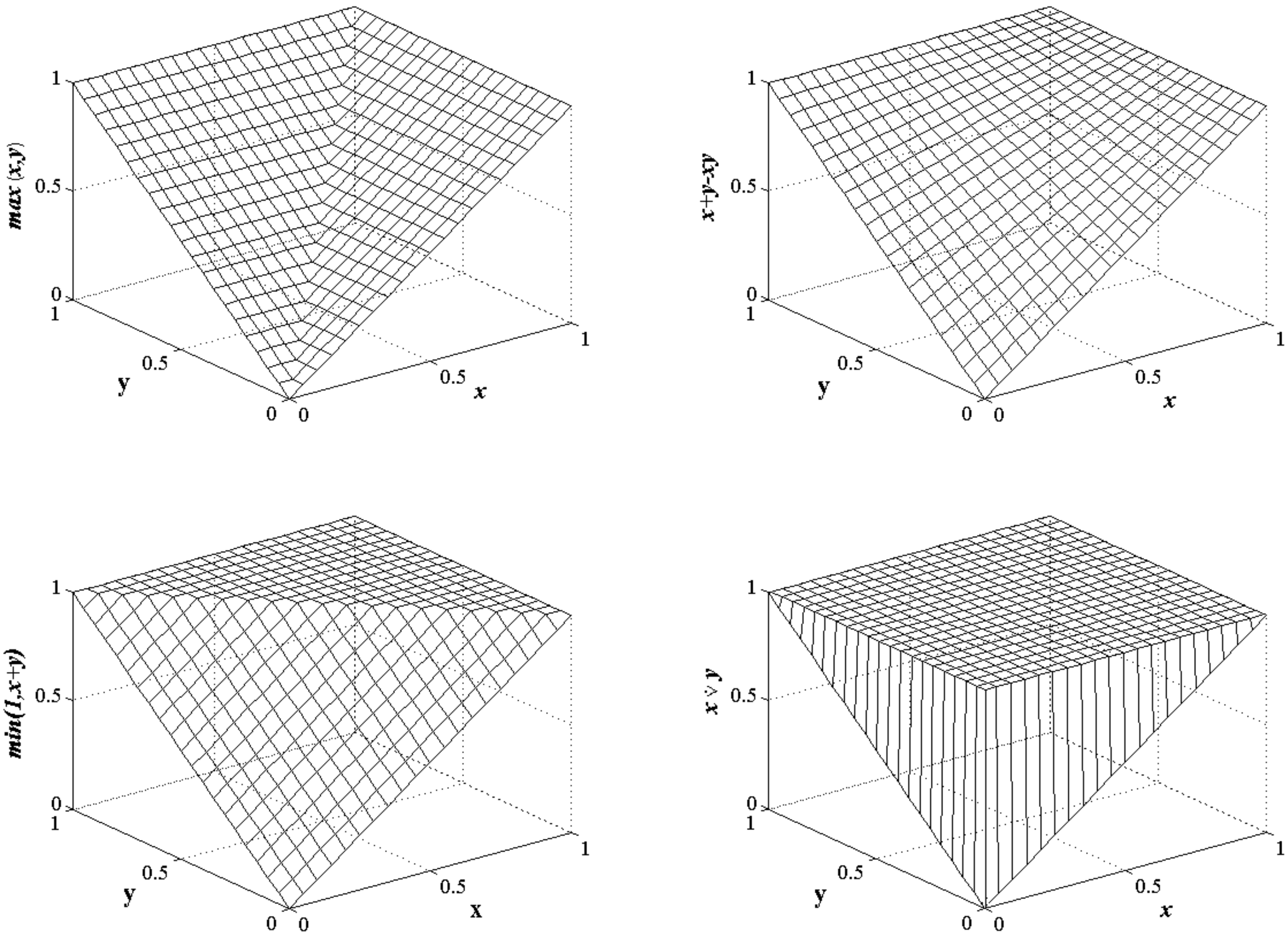
ou seja, para cada elemento *x* do conjunto Universo *X*, esta função toma como argumentos os par de pertença de *x* em *A* e *B*.



A função um grau de pertença de *x* no conjunto de união de *A* com *B* é dado por:

$$\mu_{A \cup B}(x) = \mu_A(x) \dot{+} \mu_B(x)$$

| | |
|---|---|
| $x \vee y = max\{x,y\}$ | Soma Lógica ou União |
| $x \widehat{+} y = x + y - x\,y$ | Soma Algébrica |
| $x \oplus y = min\{1\ ,\ x+y\}$ | Soma Limitada |
| $x \dot{\vee} y = \begin{cases} x & se\ \ y = 0 \\ y & se\ \ x = 0 \\ 1 & se\ x,y > 0 \end{cases}$ | Soma Drástica |


max (x,y)
x+y-xy
min(1,x+y)
x ∨ y
x
y
0
0.5
1

# Relação e composição

A relação dos conjuntos bivalentes $X_1, X_2, \cdots, X_n$ é um subconjunto do espaço produto cartesiano, referido como

$$R\left(X_1, X_2, \cdots, X_n\right) \subset X_1 \times X_2 \times \cdots \times X_n$$

A relação *R* é binária, pode ser definida por uma função característica, que assume apenas os valores 0 ou 1.

$$R\left(x_1, x_2, \cdots, x_n\right) = \begin{cases} 1\, sse \left(x_1, x_2, \cdots, x_n\right) \in R \\ 0 \;\; outros \end{cases}$$

*Exemplo: Seja R uma relação dos seguintes conjuntos:*
*X={Portugal (P), EUA, Inglaterra (UK), Canadá (C) } ;*
*Y={ libra, dólar, euro} ;*
*Z= {Europa, América}.*

*A relação R associa o país à moeda e ao continente. Assim, a relação R pode ser representada pela seguinte matriz tridimensional função de pertença:*

$$\begin{array}{cc} & \begin{array}{cccc} P & EUA & UK & C \end{array} \\ \begin{array}{c} \text{libra} \\ \text{dolar} \\ \text{euro} \end{array} & \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \\ & \text{Europa} \end{array} \qquad \begin{array}{cc} & \begin{array}{cccc} P & EUA & UK & C \end{array} \\ \begin{array}{c} \text{libra} \\ \text{dolar} \\ \text{euro} \end{array} & \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ & \text{America} \end{array}$$

Uma relação difusa é definida como um conjunto difuso no espaço produto dos conjuntos $X_1, X_2, \cdots, X_n$, em que cada elemento $\vec{x} \in X_1 \times X_2 \times \cdots \times X_n$ pode ter vários graus de pertença, geralmente representados por um valor real no intervalo [0,1]:

$$\mu_R : \times_{i \in N} X_i \rightarrow [0,1]$$

Exemplo: Seja R a relação difusa que relaciona a distância entre dois conjuntos de cidades:

X={Coimbra, Lisboa, Porto} ;

Y={Porto, Vila Real}:

$$\begin{array}{c c} & \begin{array}{cc} P & V \end{array} \\ \begin{array}{c} C \\ L \\ P \end{array} & \begin{bmatrix} 0.3 & 0.6 \\ 0.8 & 1 \\ 0 & 0.2 \end{bmatrix} \end{array}$$

$R(X,Z) = P(X,Y) \circ Q(Y,Z)$

## Composição *max-min*

$$\mu_{P \circ Q}(x,z) = \max_{y \in Y} \min\left[\mu_P(x,y), \mu_Q(y,z)\right]$$

## Composição *min-max*

$$\mu_{P \square Q}(x,z) = \min_{y \in Y} \max\left[\mu_P(x,y), \mu_Q(y,z)\right]$$

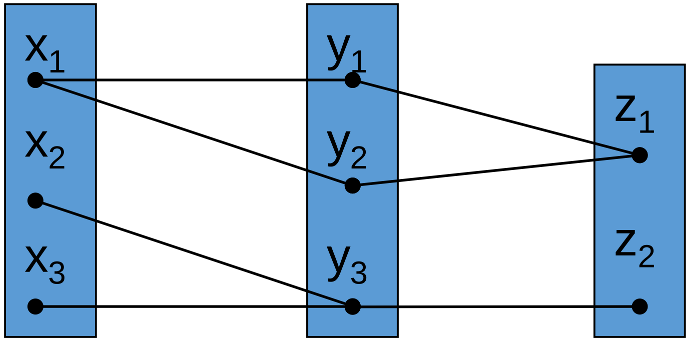
Para todo o *x X* e *z Z*.

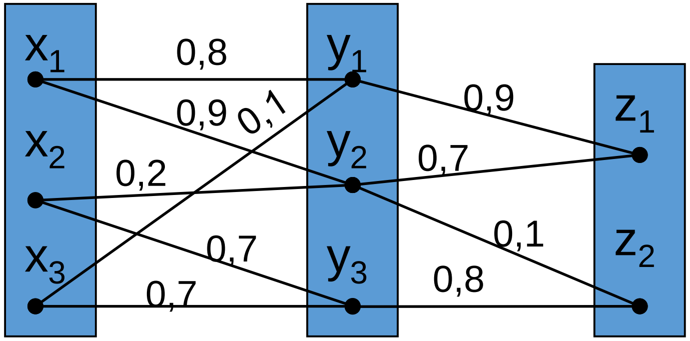
Sendo que:

$$\min_{y \in Y} \max\left[\mu_P(x,y), \mu_Q(y,z)\right] = \overline{\max_{y \in Y} \min\left[\overline{\mu_P(x,y)}, \overline{\mu_Q(y,z)}\right]}$$

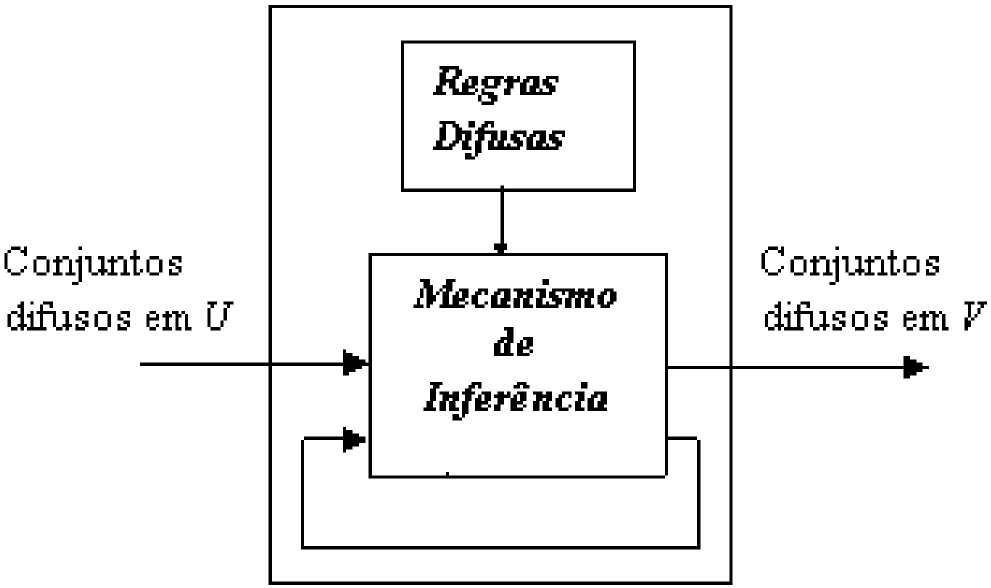
## Problema:

Encontre a relação R(X,Z)= P(X,Y)° Q(Y,Z) dos dois seguintes exemplos através da Composições max-min e min-max.



# INTERPRETAÇÃO DA REGRA DIFUSA "IF-THEN"




Um Sistema de Lógica Difusa mapeia os conjuntos difusos de entrada nos conjuntos difusos de saída.

$$R^{(l)} : IF \ x_1 \ is \ F_1^l \ and \ \ldots and \ x_n \ is \ F_n^l \ \ THEN \ \ y \ is \ G^l$$

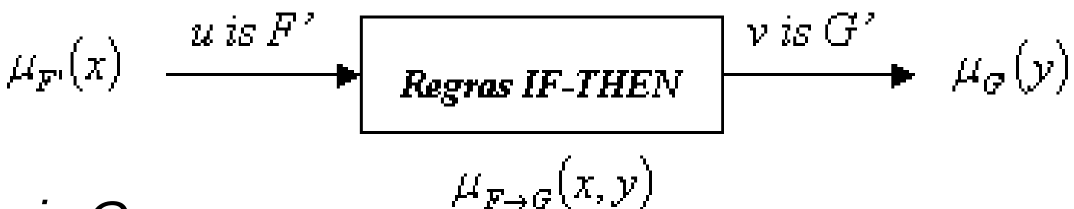
Cada regra *l*, *IF-THEN*, define um conjunto difuso

$F_1^l \times F_2^l \times \cdots \times F_n^l \;\to\; G^l$ no espaço produto $U \times V$.

Uma regra é interpretada como uma implicação difusa

$$F_1^l \times F_2^l \times \cdots \times F_n^l \;\to\; G^l$$

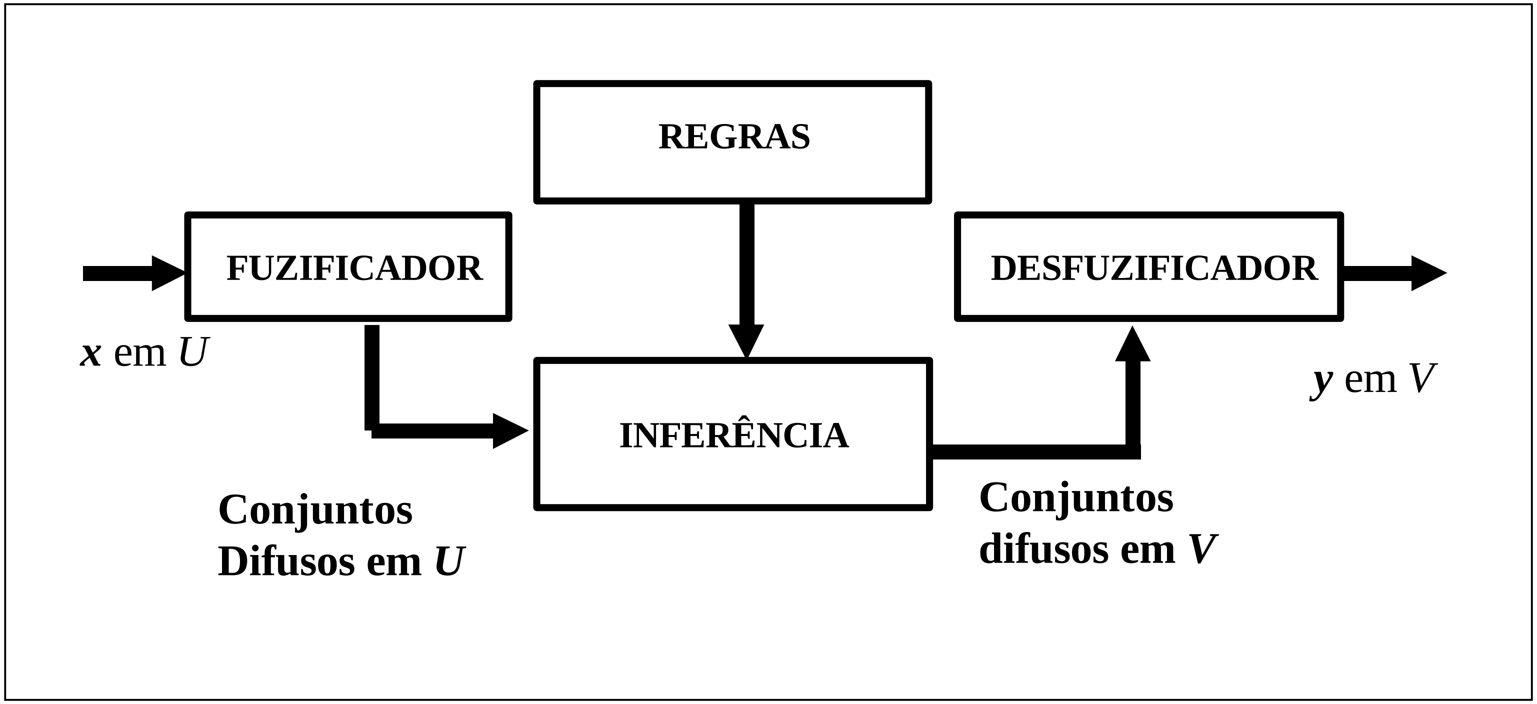
- *Modus Ponens* (MP):




premissa 1: *x is F'*
premissa 2: *IF x is F, THEN y is G*
consequência: *y is G'*

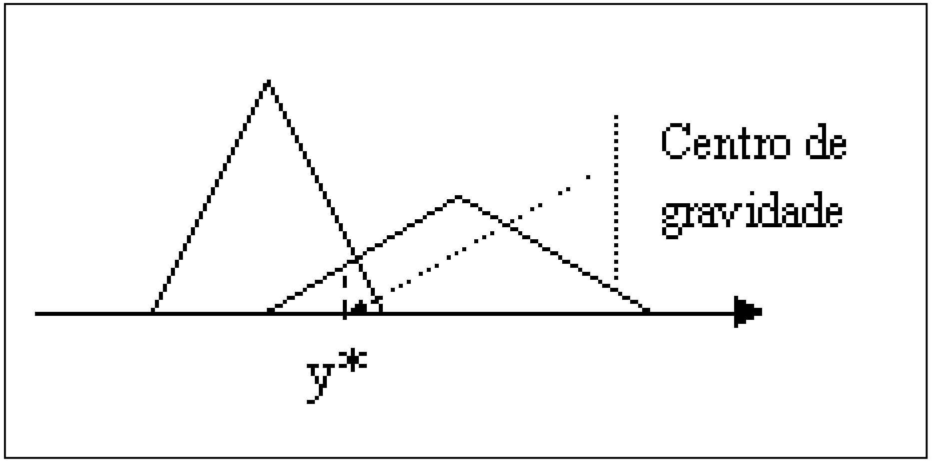
em que *F* e *G* são proposições.

# SISTEMA DE LÓGICA DIFUSA



O *desfuzificador de centro de gravidade*, especifica *y** como o centro da área coberta pela função de pertença *B'*, isto é:

$$y^* = \frac{\int_V y \cdot \mu_{B'}(y)\,dy}{\int_V \mu_{B'}(y)\,dy}$$

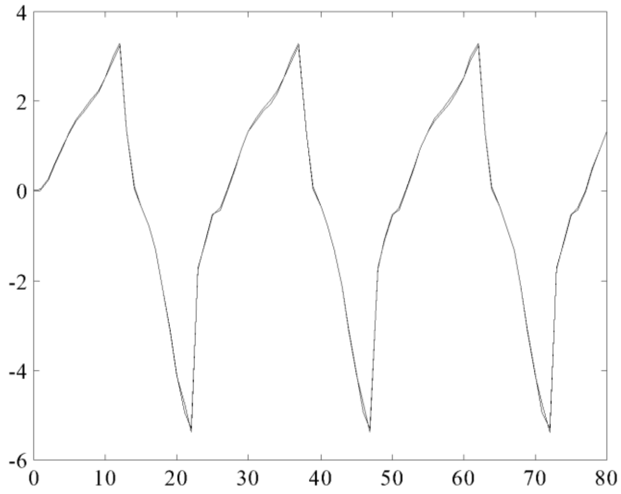


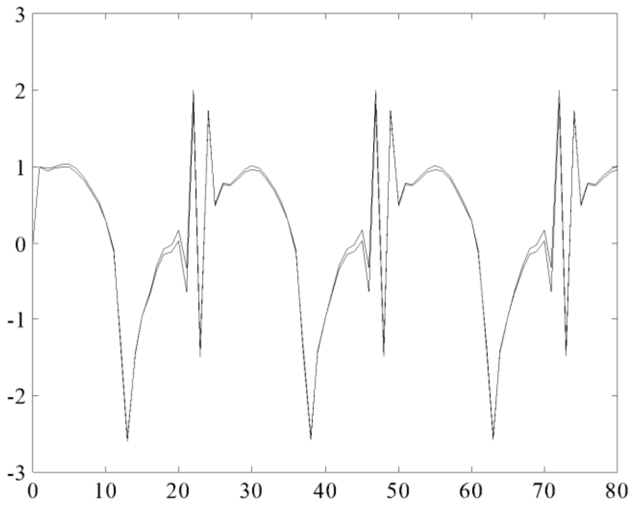
*Exemplo: Neste exemplo, pretende-se mostrar a habilidade do uso de um sistema difuso na modelação de um sistema de várias entradas e várias saídas (MIMO, Multi-Input-Multi-Output), descrito pelo seguinte sistemas de equações às diferenças:*

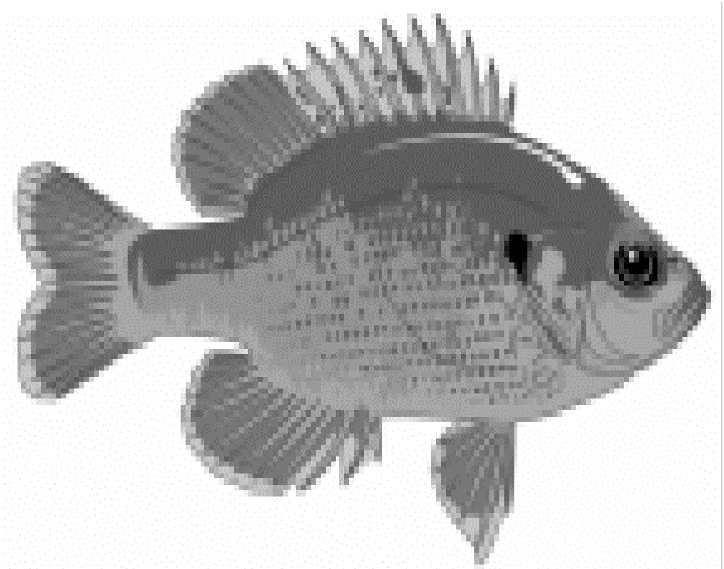
$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \end{bmatrix} = \begin{bmatrix} \dfrac{y_1(k)}{1+y_2^2(k)} \\ \dfrac{y_1(k)\,y_2(k)}{1+y_2^2(k)} \end{bmatrix} + \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}$$

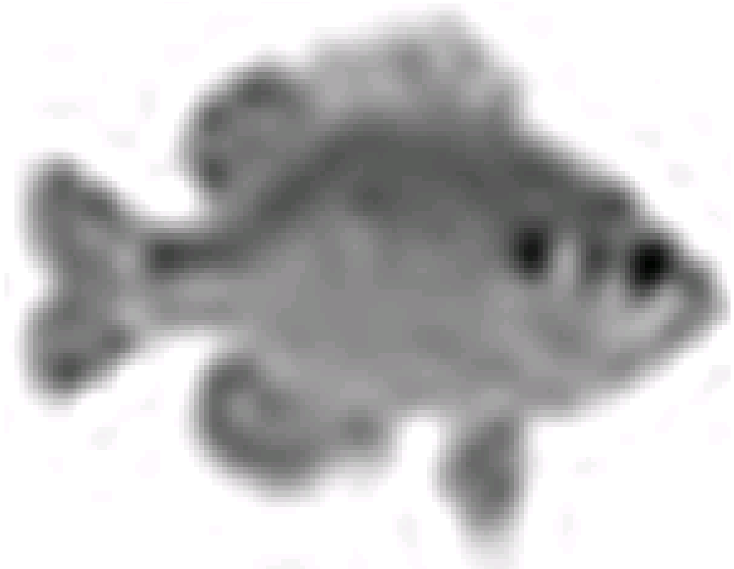
$$\begin{bmatrix} \hat{y}_1(k+1) \\ \hat{y}_2(k+1) \end{bmatrix} = \begin{bmatrix} \hat{f}_1\big(y_1(k), y_2(k)\big) \\ \hat{f}_2\big(y_1(k), y_2(k)\big) \end{bmatrix} + \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}$$

$$\left[u_1(k), u_2(k)\right] = \left[sin(2\pi k/25), cos(2\pi k/25)\right]$$



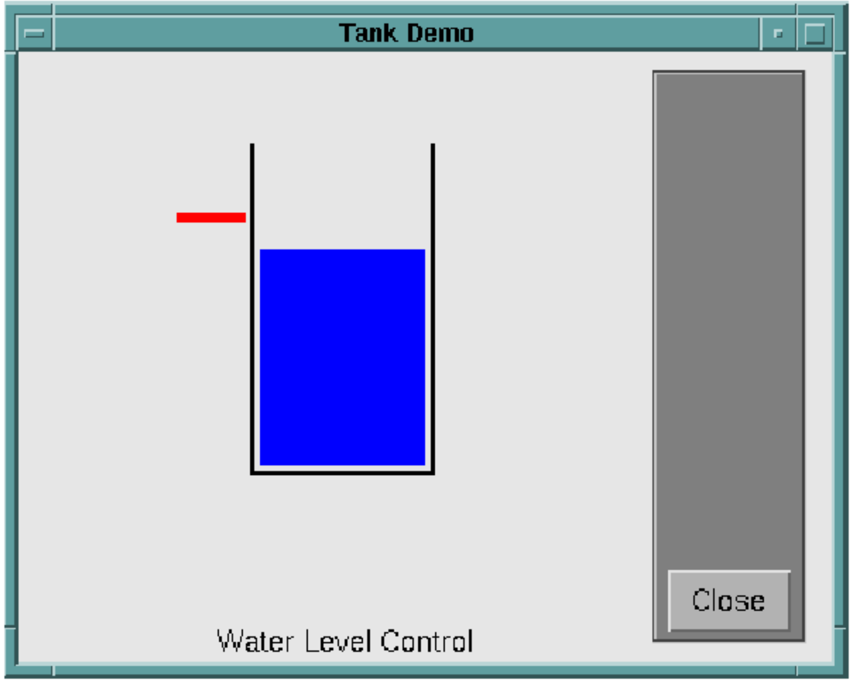




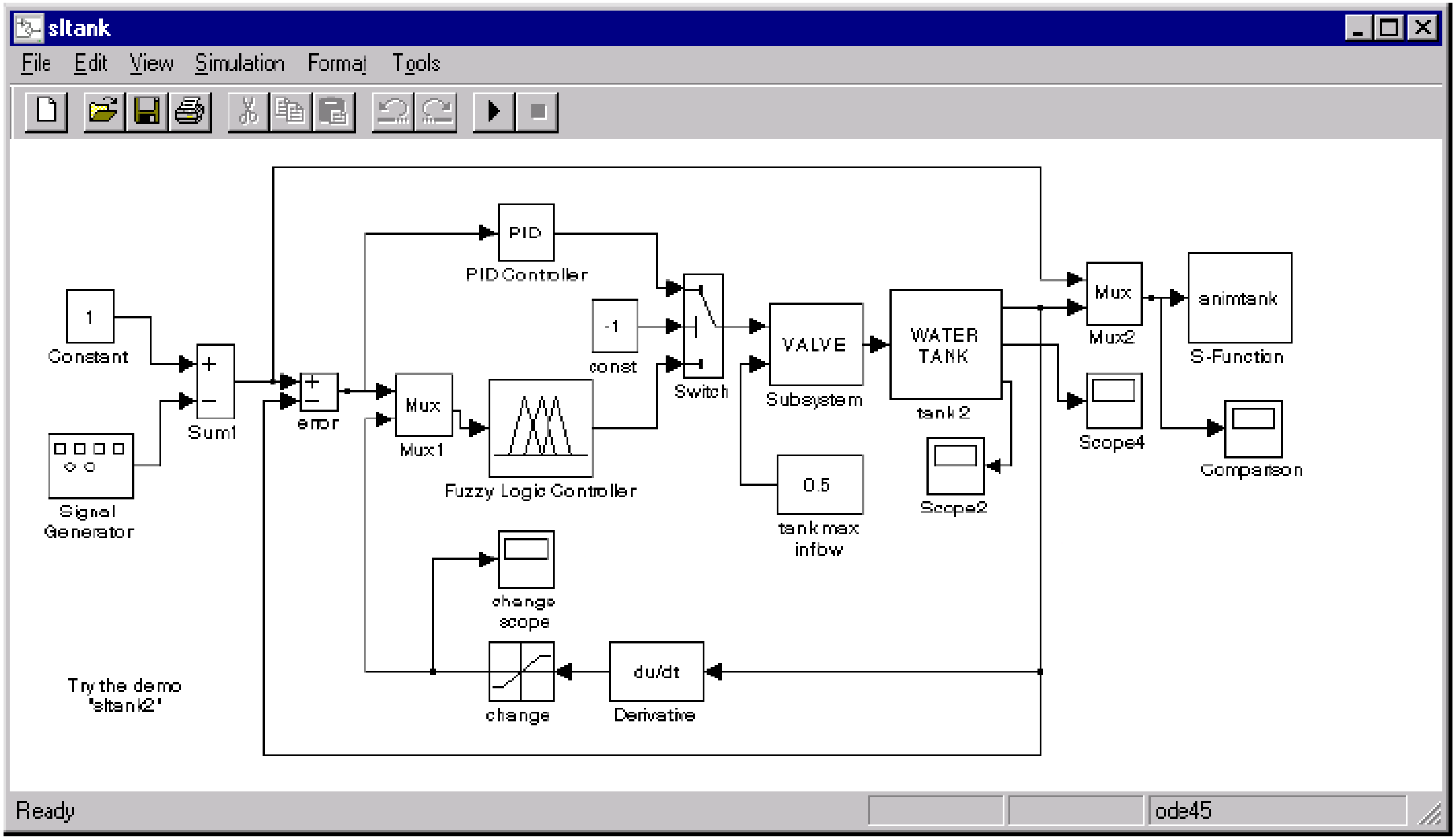
150×100 pixeis

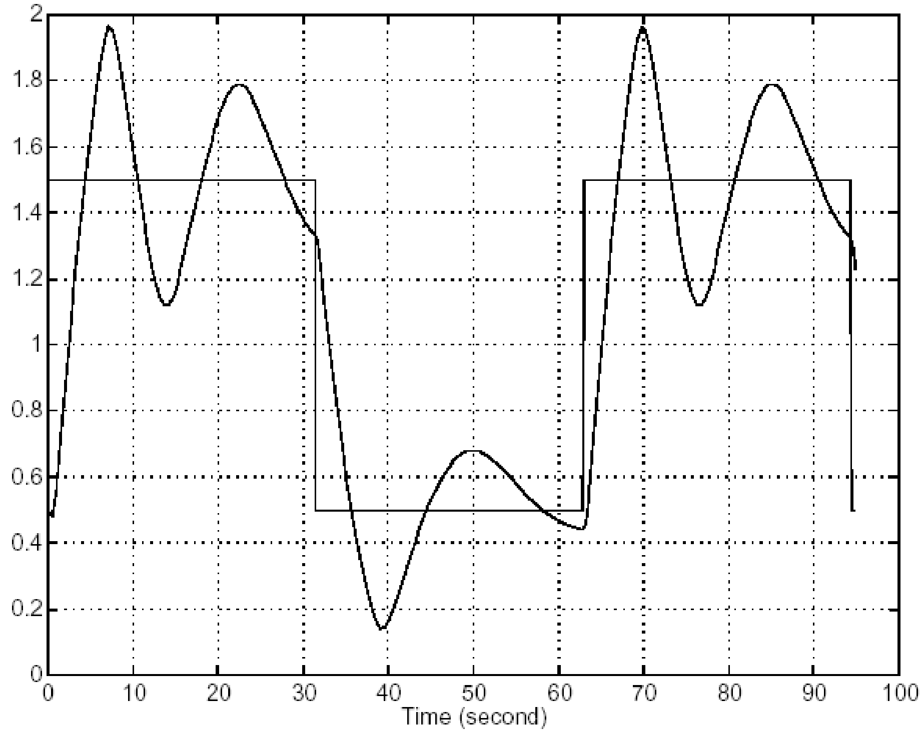
200 regras difusas

# APLICAÇÕES: Controlo do nível de um tanque




1. *If (level is okay) then (valve is no_change) (1)*
2. *If (level is low) then (valve is open_fast) (1)*
3. *If (level is high) then (valve is close_fast) (1)*

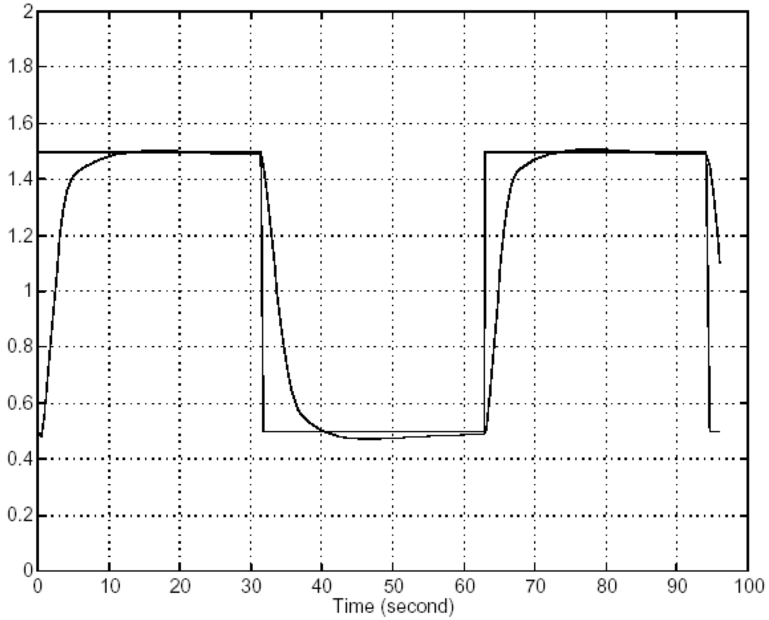

sltank
File Edit View Simulation Format Tools
PID
PID Controller
1
Constant
-1
const
Switch
VALVE
Subsystem
WATER
TANK
tank 2
Mux
Mux2
animtank
S-Function
+
−
Sum1
+
−
error
Mux
Mux1
Fuzzy Logic Controller
Scope4
Comparison
Signal
Generator
0.5
tank max
inflow
Scope2
change
scope
du/dt
Derivative
change
Try the demo
'sltank2'
Ready
ode45

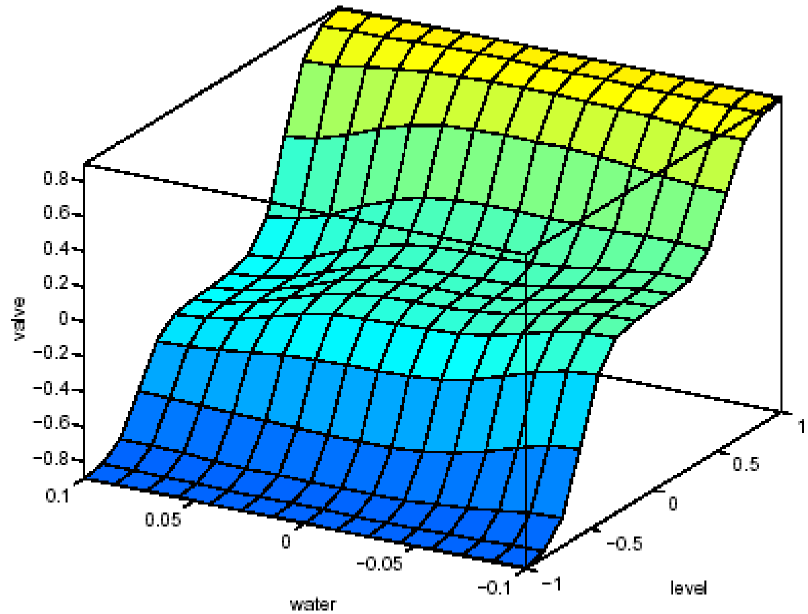


2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
10
20
30
40
50
60
70
80
90
100
Time (second)

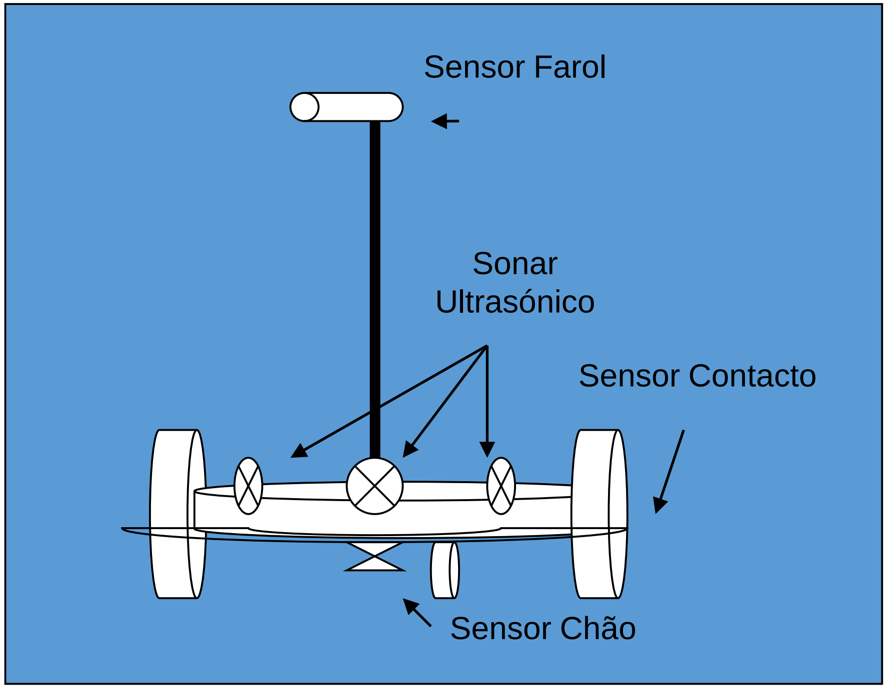
*4. If (level is good) and (rate is negative), then (valve is close_slow) (1)*

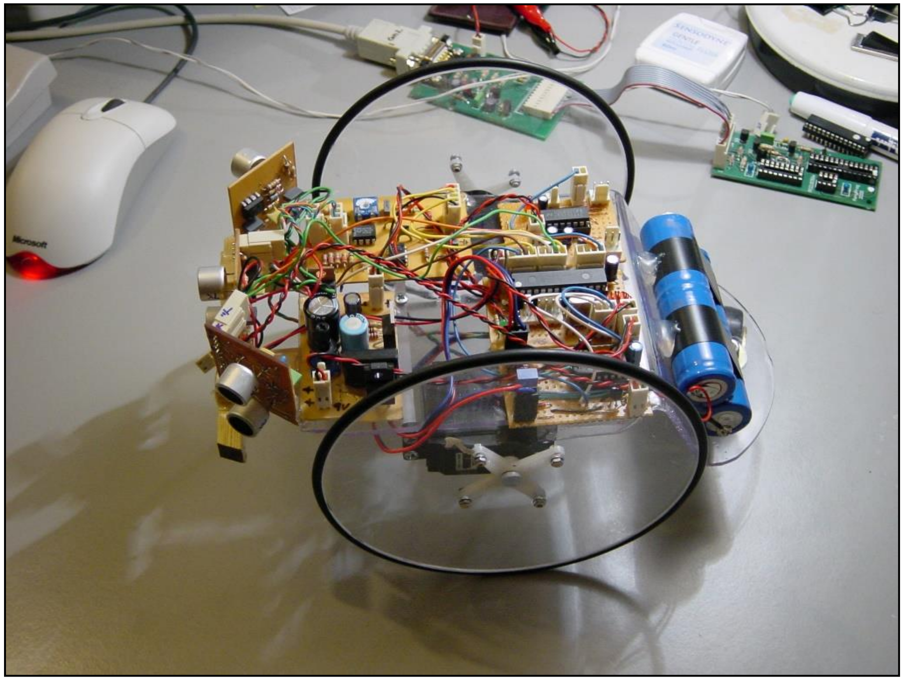
*5. If (level is good) and (rate is positive), then (valve is open_slow) (1)*

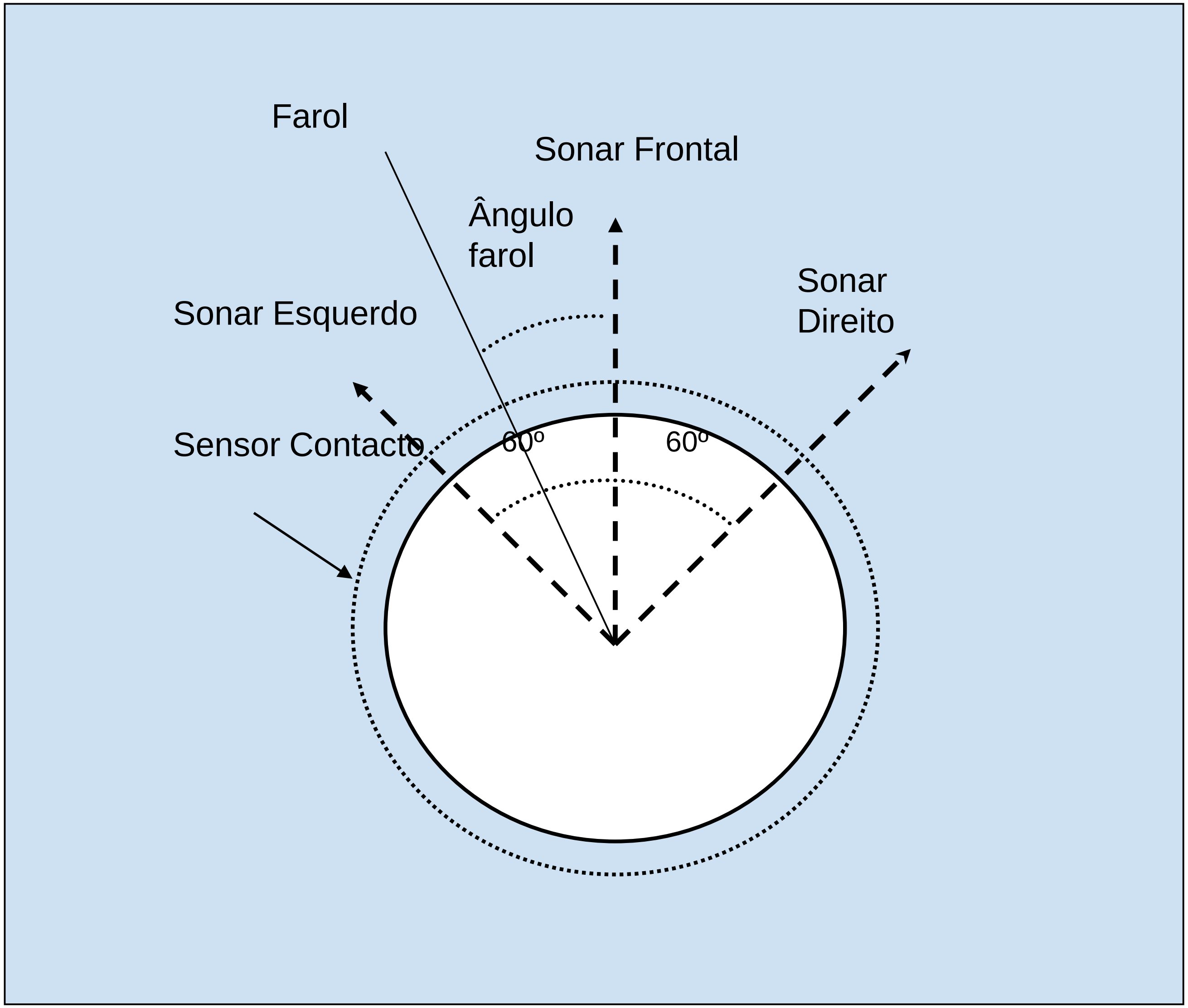


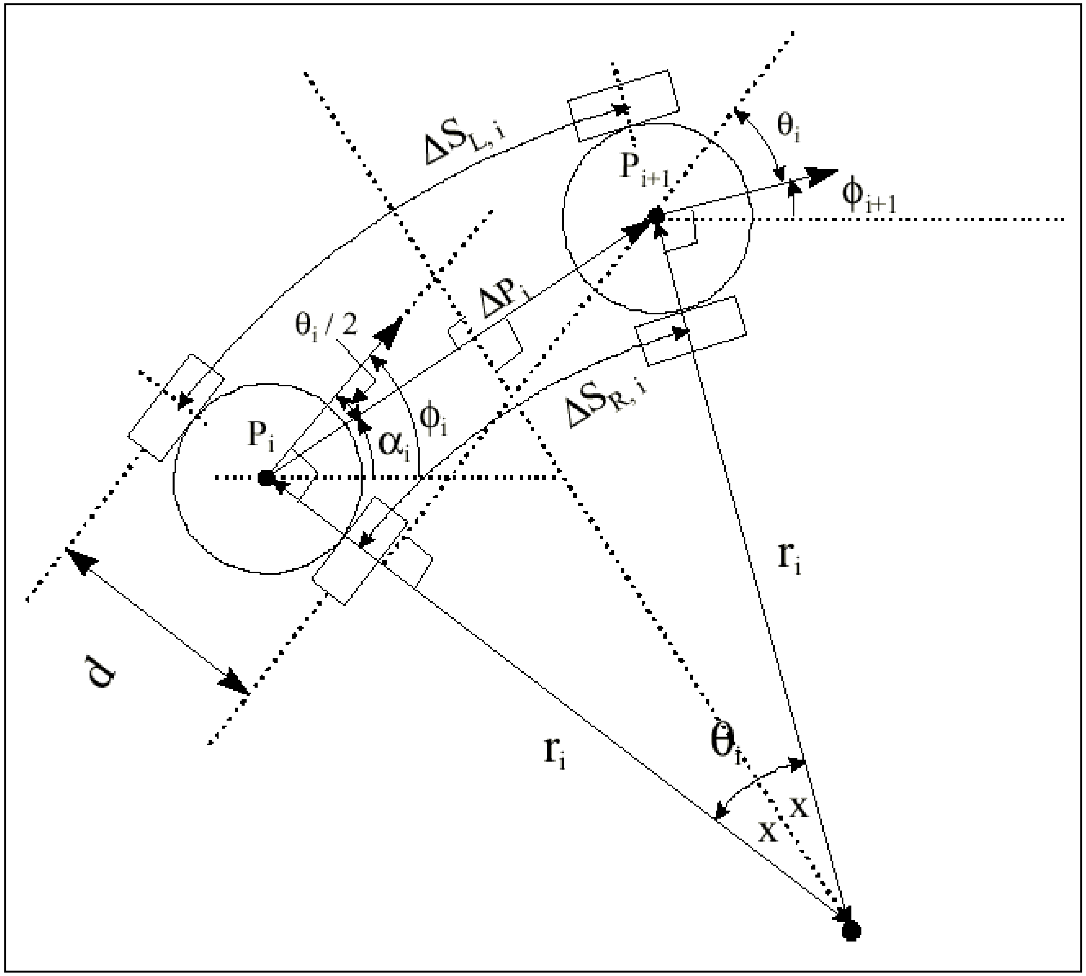



# Controlo na navegação de um robô autónomo móvel

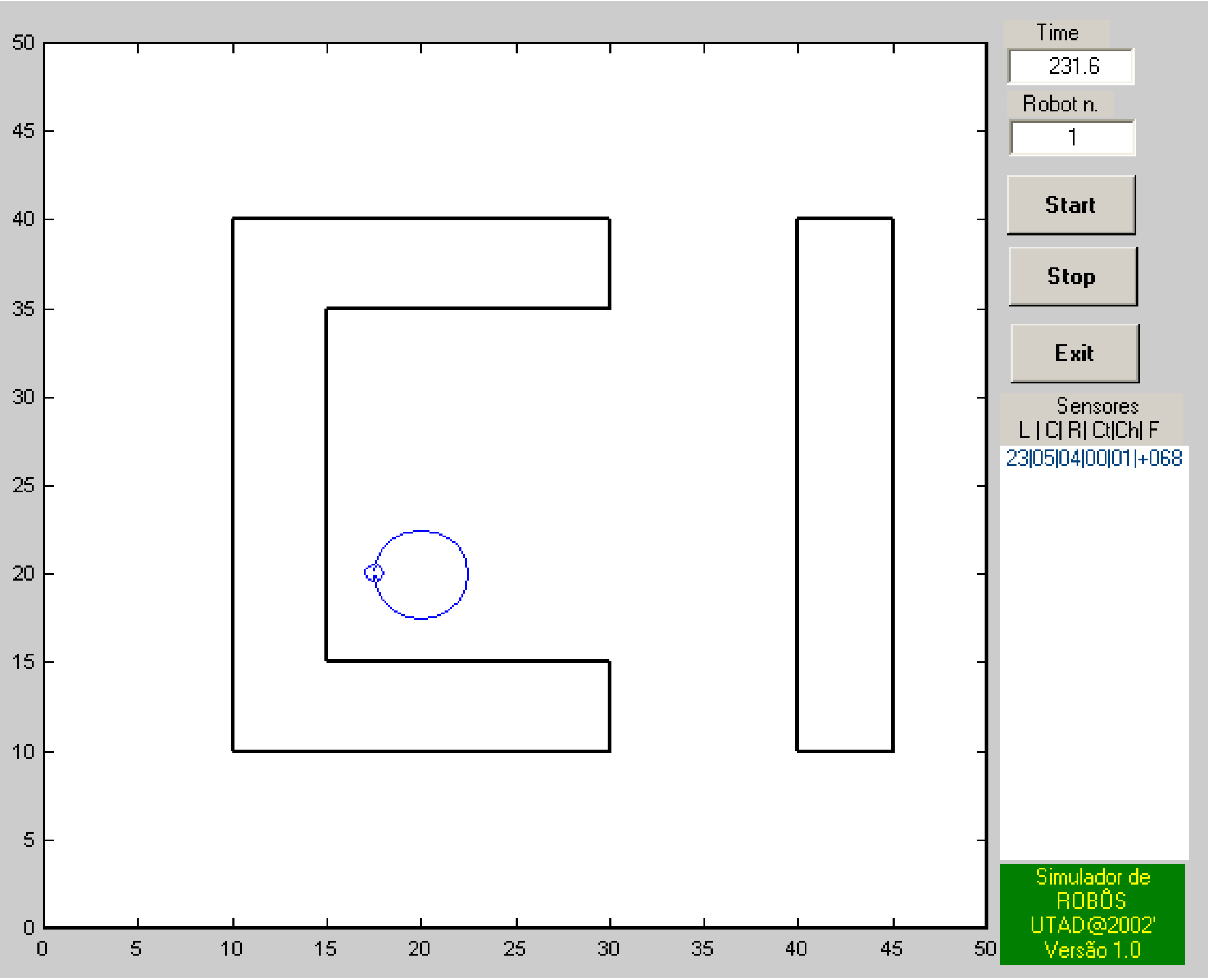


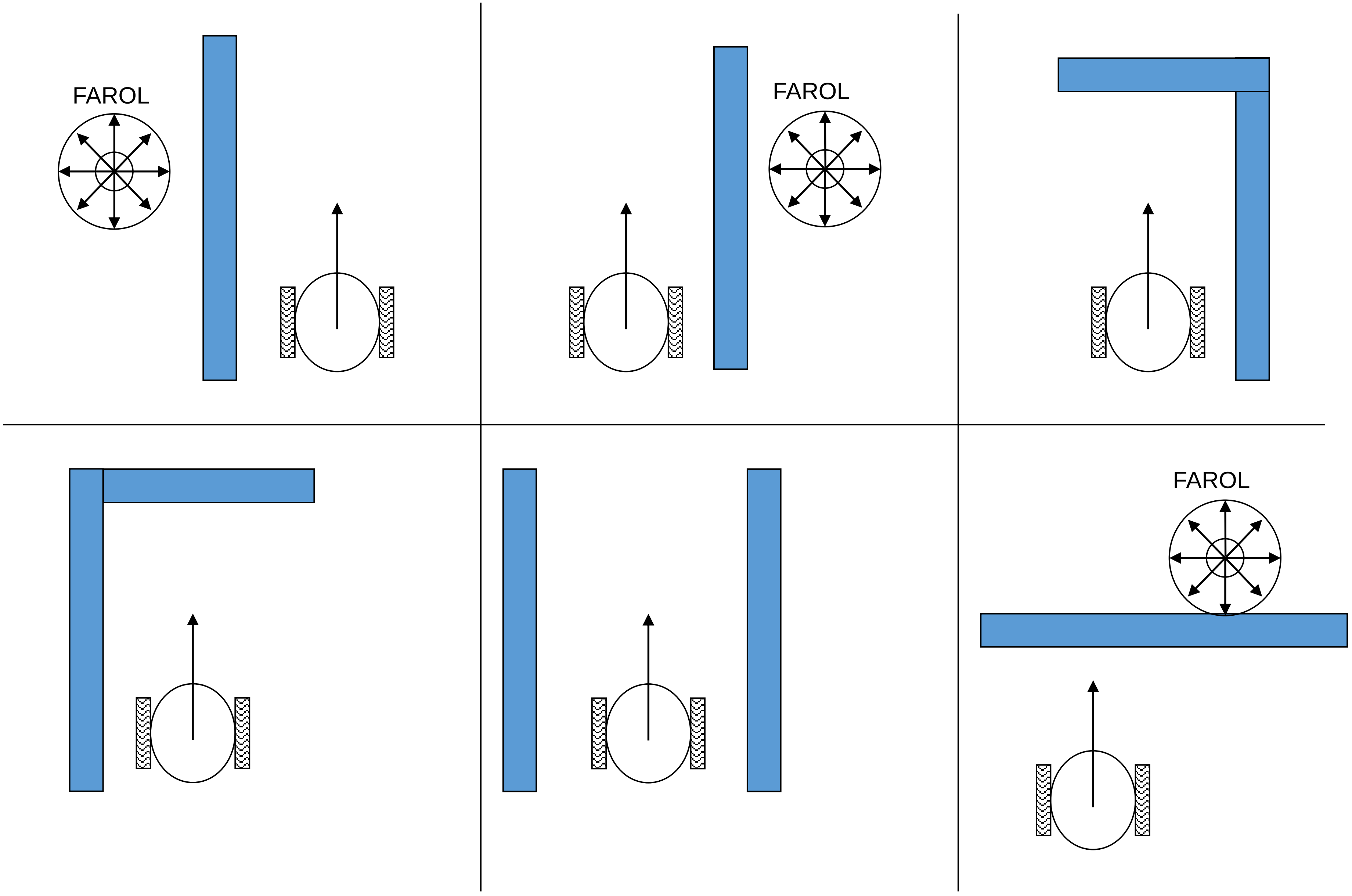




Farol
Sonar Frontal
Ângulo
farol
Sonar
Direito
Sonar Esquerdo
Sensor Contacto
60º
60º

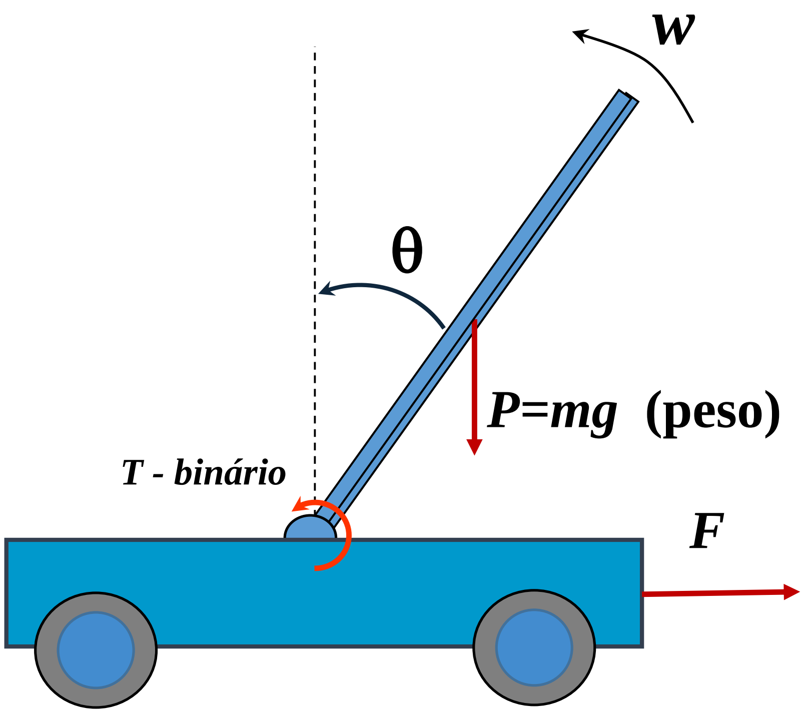


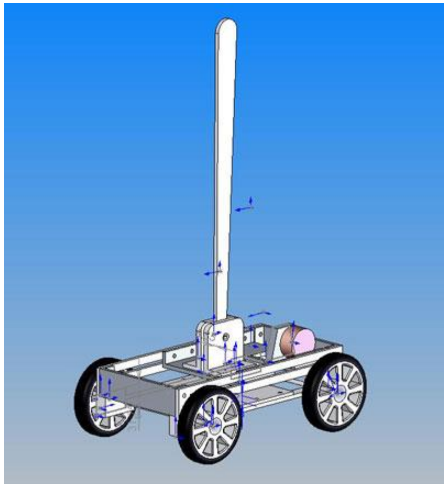
ΔS_{L, i}
θ_i
P_{i+1}
ϕ_{i+1}
ΔP_i
θ_i / 2
ϕ_i
α_i
P_i
ΔS_{R, i}
r_i
d
r_i
θ_i
x
x


Time
231.6
Robot n.
1
Start
Stop
Exit
Sensores
L | C| R| Ct|Ch| F
23|05|04|00|01|+068
Simulador de
ROBÔS
UTAD@2002'
Versão 1.0


FAROL
FAROL
FAROL

# CONTROLO DO PENDULO INVERTIDO

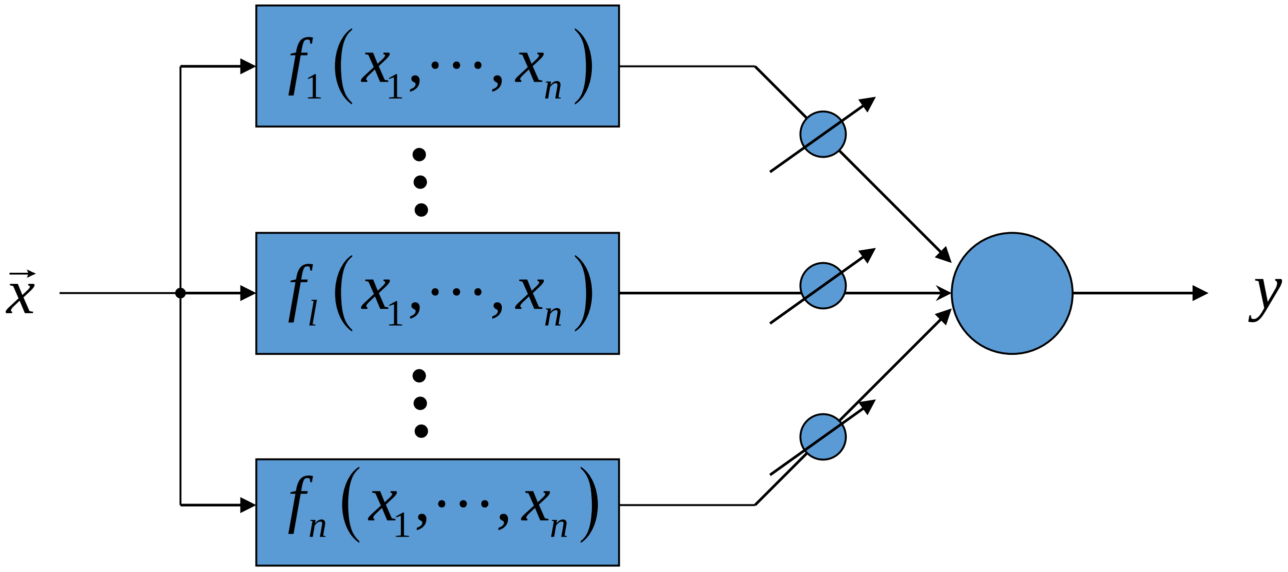


# MODELO DIFUSO TSK

$$R^{(l)}: IF\ x_1\ is\ F_1^l\ and\ \ldots and\ x_n\ is\ F_n^l\ \ THEN\ y = f_l\left(x_1,\cdots,x_n\right)$$

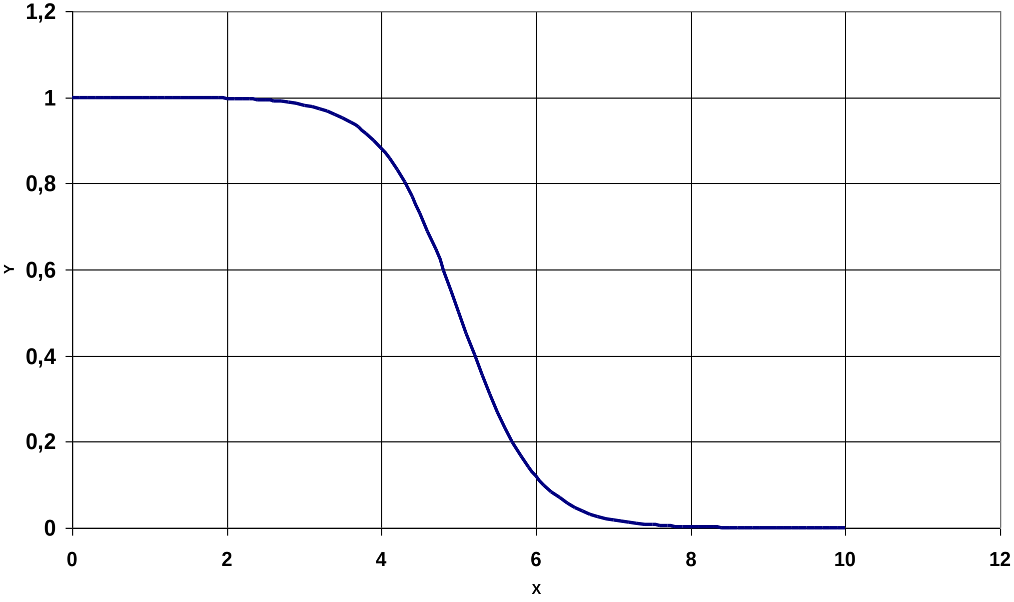
A saída é agora uma função das variáveis de entrada, em alternativa a um conjunto difuso



# Modelos Matemáticos

- Os modelos matemáticos são funções. Exige profundo conhecimentos matemáticos.
- Se o sistema é não-linear (tais como NN), a construção de um modelo exige um significativo esforço de calculo e esforço computacional.
- Os modelos lineares podem ser uma aproximação muito simplificada.
- Como encontrar as funções apropriadas?

Y=1-1./(1 + EXP(-2*(X-5)))

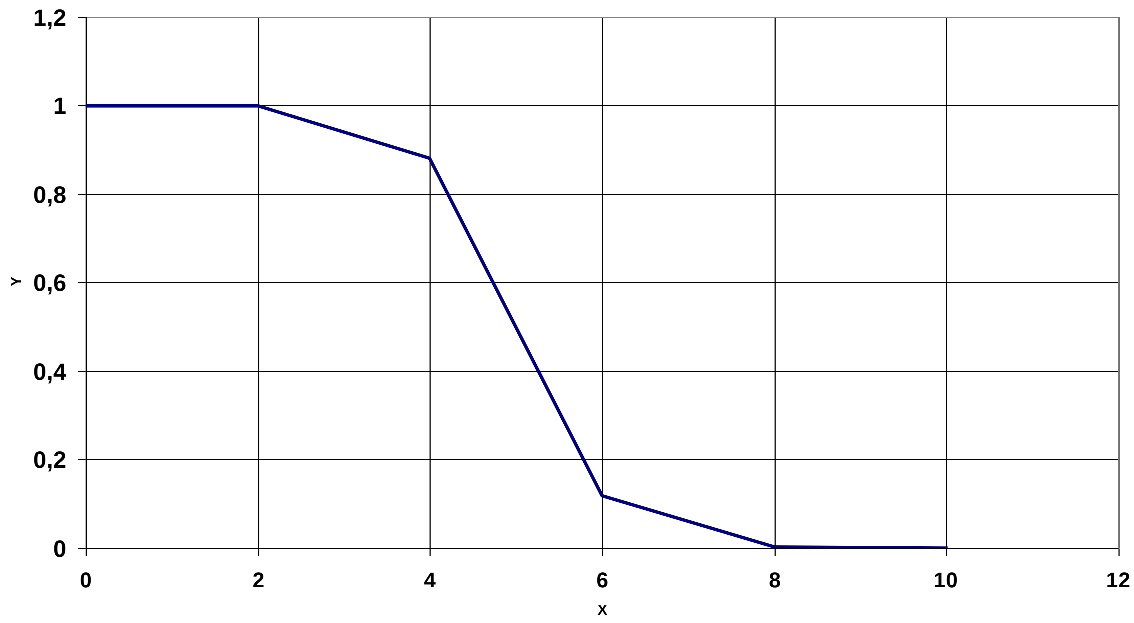


# Construção do Modelo (trad. regras )

- Regra para cada entrada => Largo nº de regras.
- Unicamente uma regra é disparada para cada entrada.
- Modelos não aperfeiçoados.

If $0 \leq x < 1$, then $y=1$
If $1 \leq x < 2$, then $y=0.99$
:
If $8 \leq x \leq 10$, then $y=0$

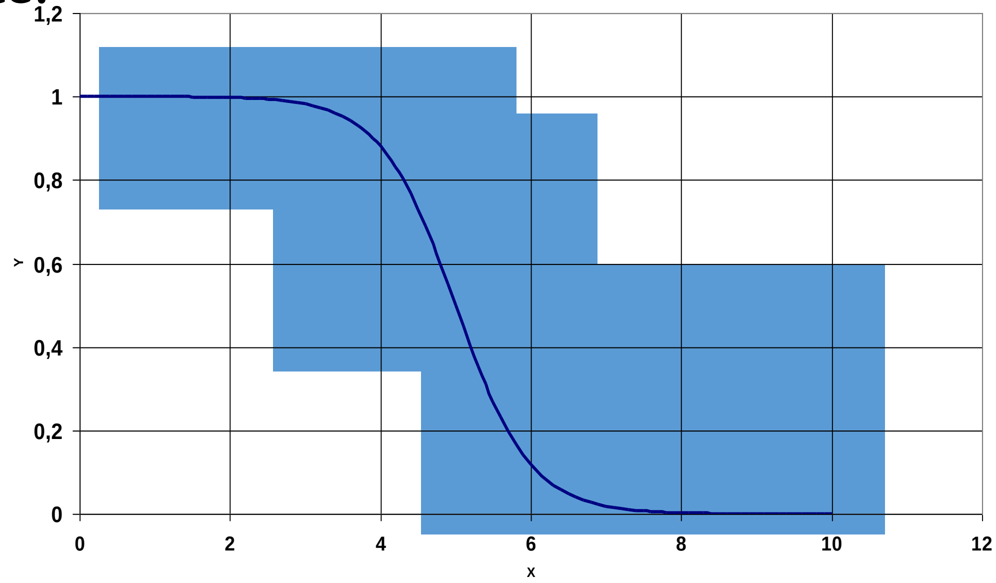
If $0 \leq x < 1$, then $y=f(x)$
If $1 \leq x < 2$, then $y=g(x)$
:
If $8 \leq x \leq 10$, then $y=h(x)$



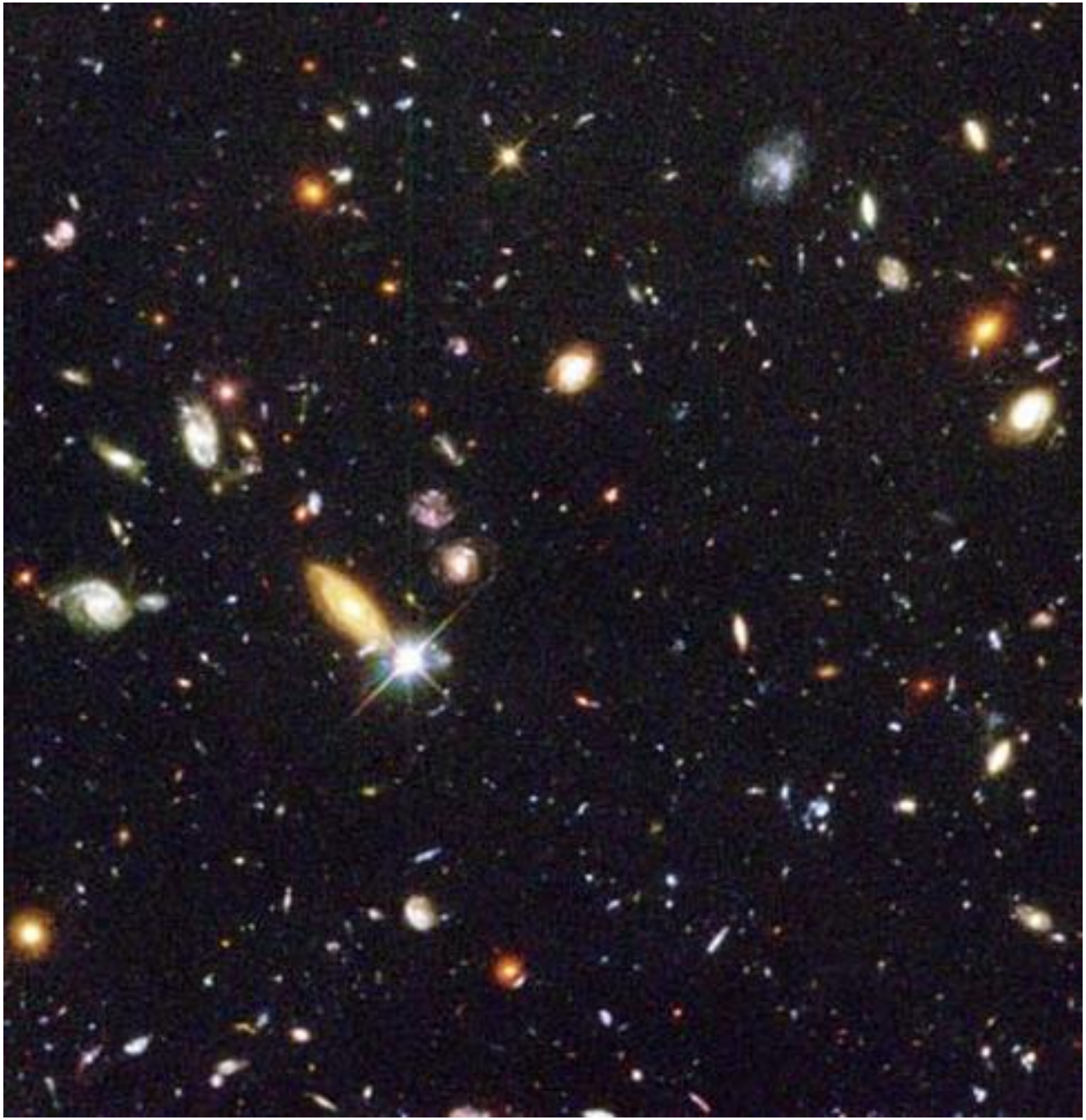
# Construção do Modelo (SC/fuzzy)

- Aproxima valores
- As regras descrevem casos típicos => Número de regras reduzido.
- Um grupo de regras “semelhantes” são “parcialmente” disparados simultaneamente.

If x≈0, then y≈1
If x≈5, then y≈0.5
If x≈10, then y≈0



# Agrupamentos Disjuntos
# e
# Agrupamentos Difusos



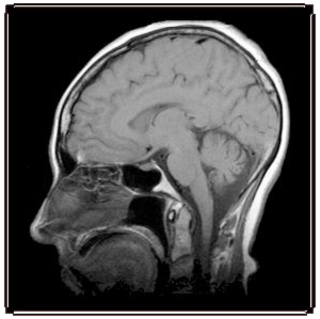
# Índice

- Introdução aos Agrupamentos
- Agrupamentos Hierárquicos
- Função Objectivo
  - Agrupamento rígido
  - Agrupamento difuso
  - Extensões dos agrupamentos difusos
- Sumário

# Motivação e objetivos

## Exemplos de Aplicações

- Segmentação de Imagens
  - Imagens Médicas
    - X-ray Computer Tomography (CT)
    - Magnetic Resonance Imaging (MRI)
    - Position Emission Tomography (PET)
  - Melhoramento de Imagens e som
  - Deteção de contornos
  - Deteção de rápidas mudanças nos Vídeos

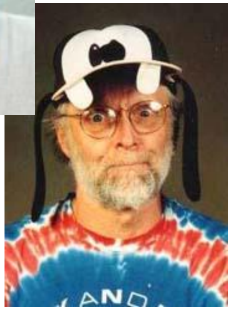


# Reconhecimento de Padrões (Pattern Recognition)

- Definição: Procura de estrutura nos dados
- Elementos para Reconhecimento Numérico de Padrões
  - Descrição de Processos
    - Denominação de Feature, Teste de dados, "Design Data"
  - "Feature Analysis"
    - Preprocessamento, Extracção, Selecção, …
  - Análise de agrupamentos
    - Rotulagem, Validação, …
  - Classificação
    - Classificação, Estimação, Predição, Controlo, …

# Agrupamento Difuso
## (Novos algoritmos de Agrupamento)

- Úteis na Modelação Difusa
  - Identificação de regras difusas necessárias para descrever o sistema, através da observação dos vetores de entrada e de saída
- Historia
  - FCM: Bezdek, 1981
  - PCM: Krishnapuram - Keller, 1993
  - FPCM: N. Pal - K. Pal - Bezdek, 1997



**Prof. Bezdek**

# O que é um agrupamento?

- Dados(Padrões)-> Agrupamentos(Classes)

- Propriedades desejadas

  homogeneidade dentro dos agrupamentos

  heterogeneidade entre agrupamentos

# Porquê agrupar?

- Melhora a compreensão dos dados
- Revela a estrutura interna dos dados
- Úteis para a análise e interpretação dos dados

# Como agrupar

- Medida de similaridade $\longrightarrow$ classes

- exemplos :

    distância, conectividade, intensidade...

https://en.wikipedia.org/wiki/Similarity_measure

# Categorias

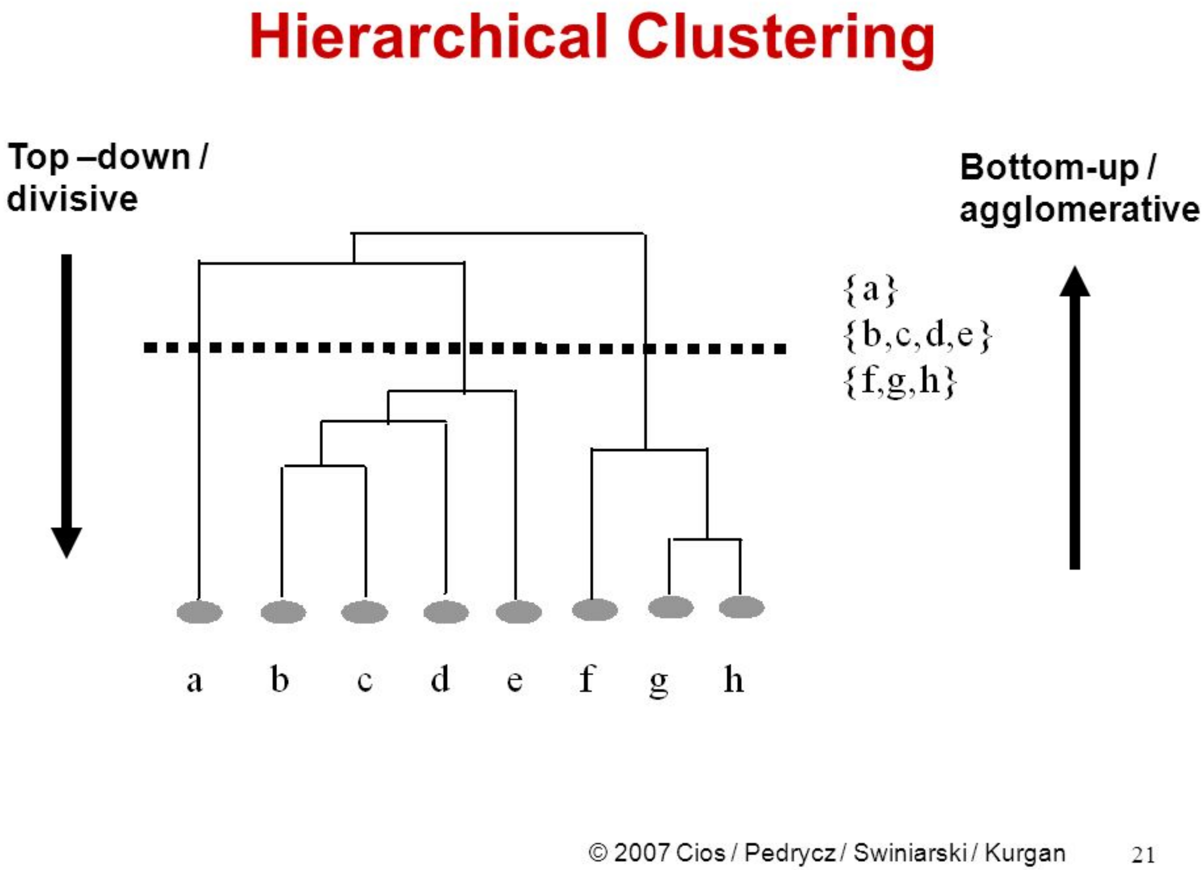
- Agrupamentos Hierárquicos
- Agrupamentos baseados em função objectiva

# Agrupamento Hierárquico

- Dois modos:

  -Bottom-up

  -Top-down

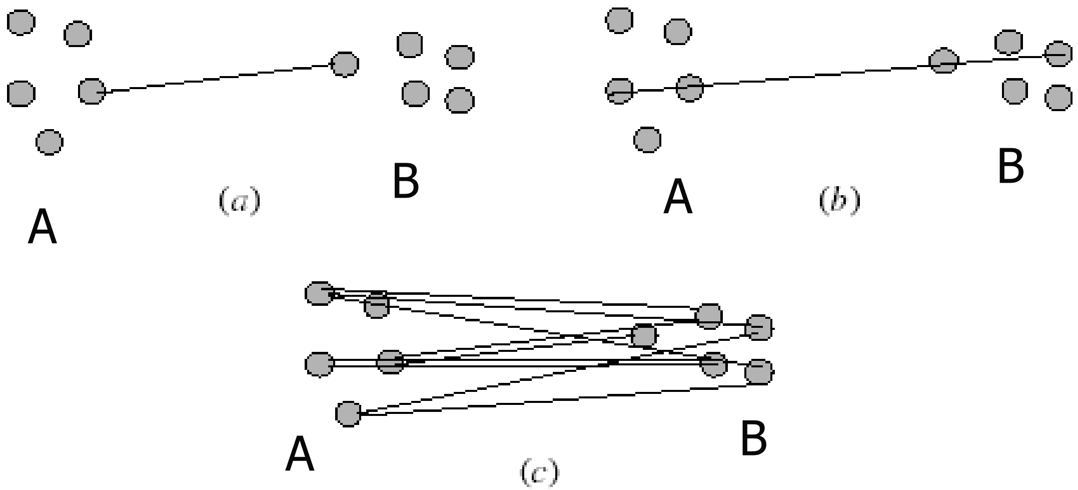


# Agrupamento Hierárquico (cont.)

- Três tipos de distância:

$$(a)\ \ d(A,B) = \min_{x \in A, y \in B} d(x,y) \qquad (b)\ \ d(A,B) = \max_{x \in A, y \in B} d(x,y)$$

$$(c)\ \ d(A,B) = \frac{1}{card(A)card(B)} \sum_{x \in A, y \in B} d(x,y)$$



# Agrupamento baseado em Função Objetiva (hard clustering)

- N padrões em $R^n$, *c* agrupamentos, um conjunto de protótipos $v_1, v_2, \ldots v_c$
- Função objetiva:

$$Q = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik} \left\| x_k - v_i \right\|^2$$

$\in$

matriz partição

$\| \quad \|^2$ uma certa distancia entre $x_k$ e $v_i$

Min Q com respeito a $v_1, v_2, \ldots, v_c$ e $U$ **U**

# Definição de U

- Se o padrão k pertence ao agrupamento i, então

$U_{ik}=1$

senão

$U_{ik}=0$

Os valores de U são binários.

# Restrições de U

- Cada agrupamento é não trivial

Exemplo N=8, c=3

$$0 < \sum_{k=1}^{N} u_{ik} < N \qquad i = 1,2,\ldots,c$$

$$U = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

- Cada padrão pertence a um único agrupamento

$$\sum_{i=1}^{c} u_{ik} = 1 \qquad k = 1,2,\ldots,N$$

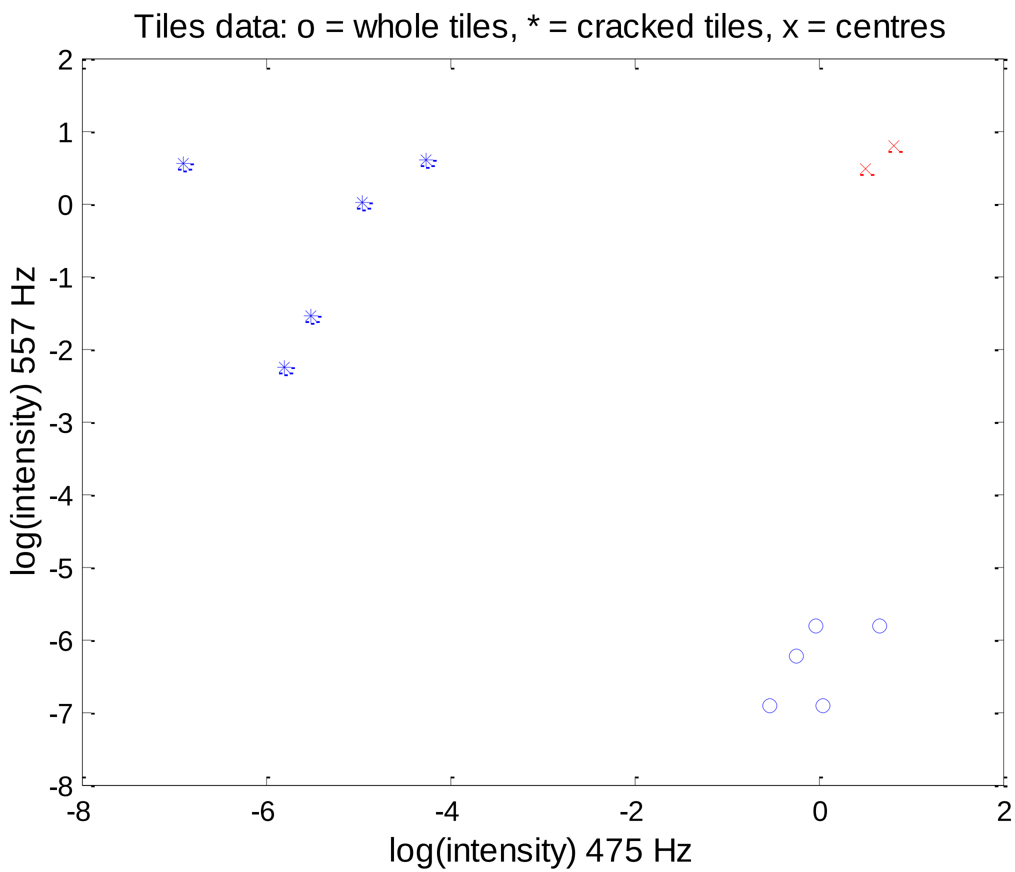
# Exemplo: Classificar azulejos defeituosas (hard C-Means – Partição rígida)

| 475Hz | 557Hz | Ok? |
|---|---|---|
| 0.958 | 0.003 | Yes |
| 1.043 | 0.001 | Yes |
| 1.907 | 0.003 | Yes |
| 0.780 | 0.002 | Yes |
| 0.579 | 0.001 | Yes |
| 0.003 | 0.105 | No |
| 0.001 | 1.748 | No |
| 0.014 | 1.839 | No |
| 0.007 | 1.021 | No |
| 0.004 | 0.214 | No |

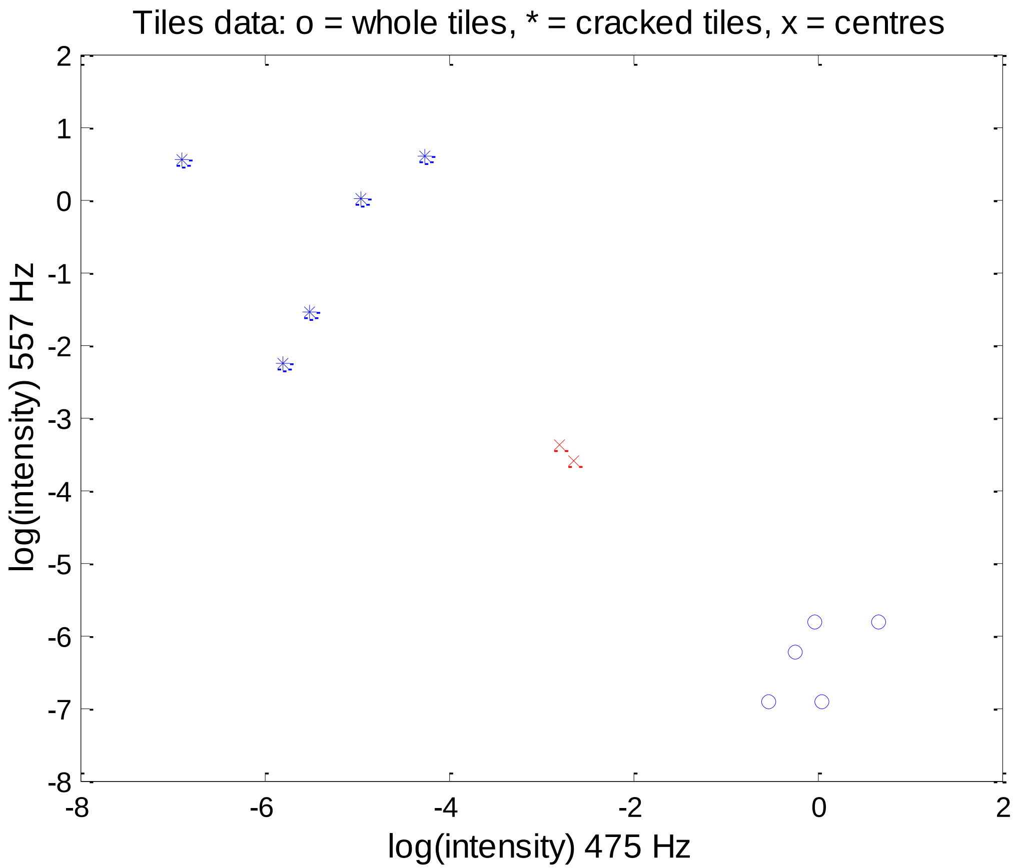
Table 1: frequency intensities for ten tiles.

Tiles are made from clay moulded into the right shape, brushed, glazed, and baked. Unfortunately, the baking may produce invisible cracks. Operators can detect the cracks by hitting the tiles with a hammer, and in an automated system the response is recorded with a microphone, filtered, Fourier transformed, and normalised. A small set of data is given in TABLE 1

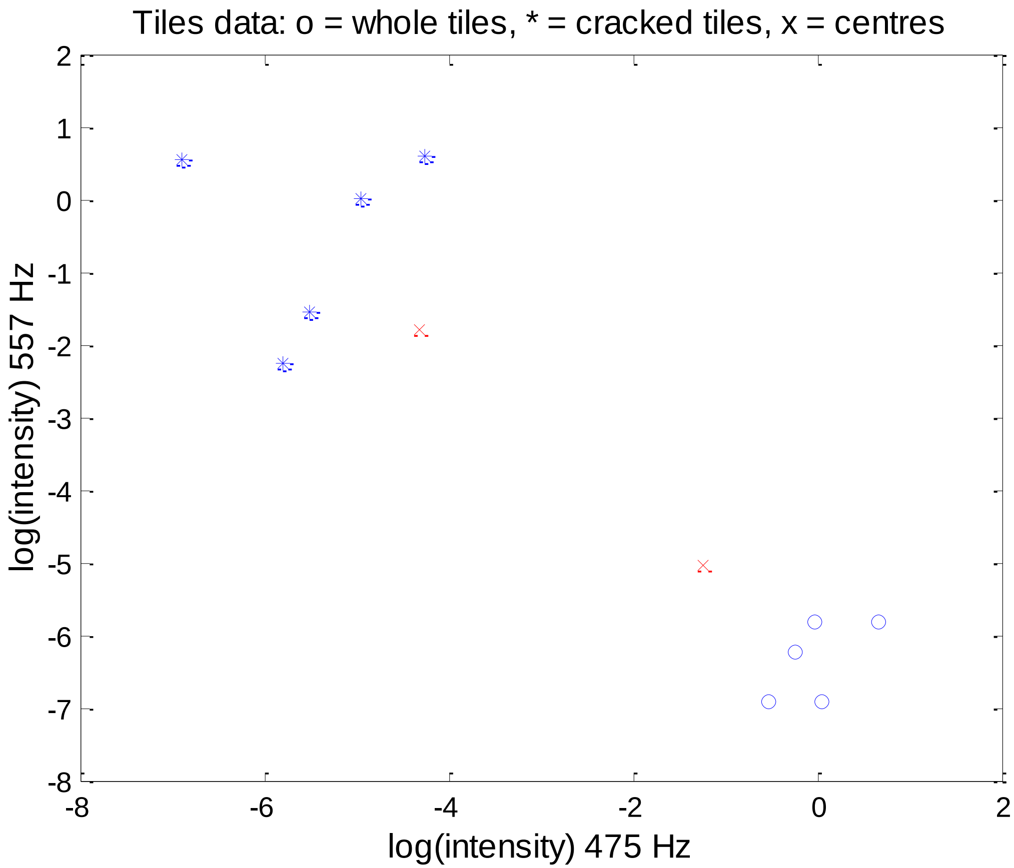
(adaptado de MIT, 1997).



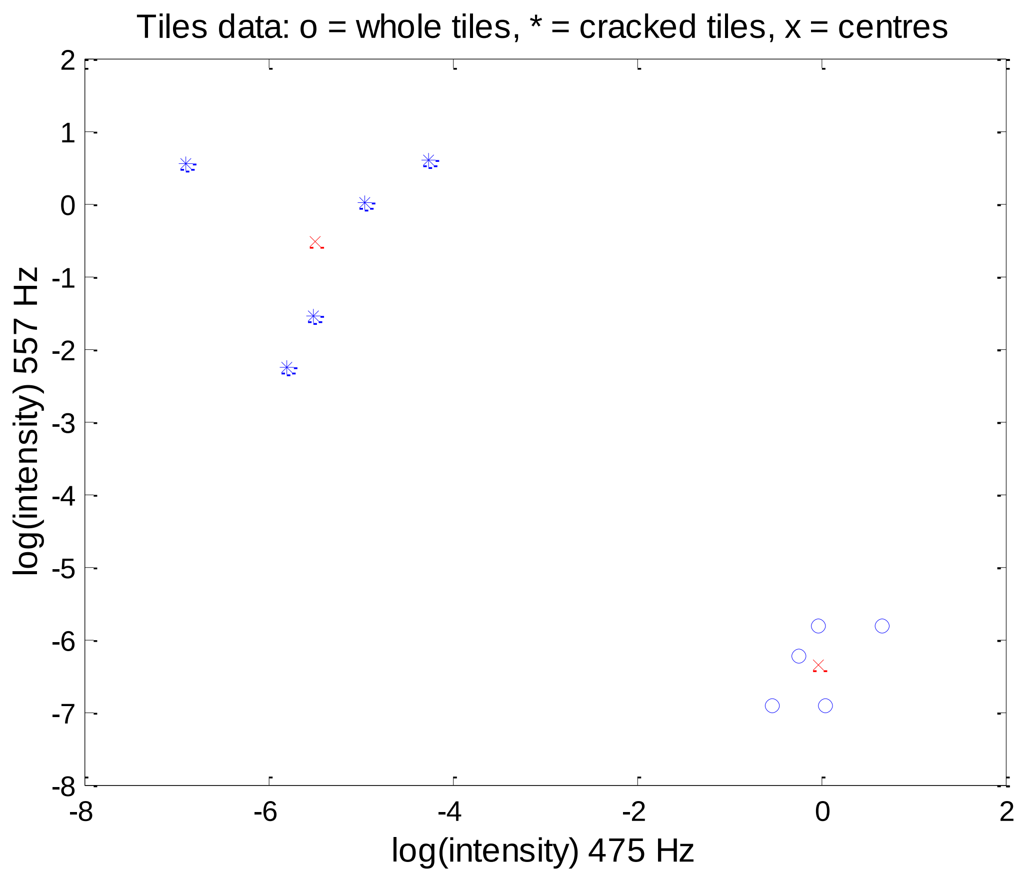

1. Place two cluster centres (x) at random.
2. Assign each data point (* and o) to the nearest cluster centre (x)



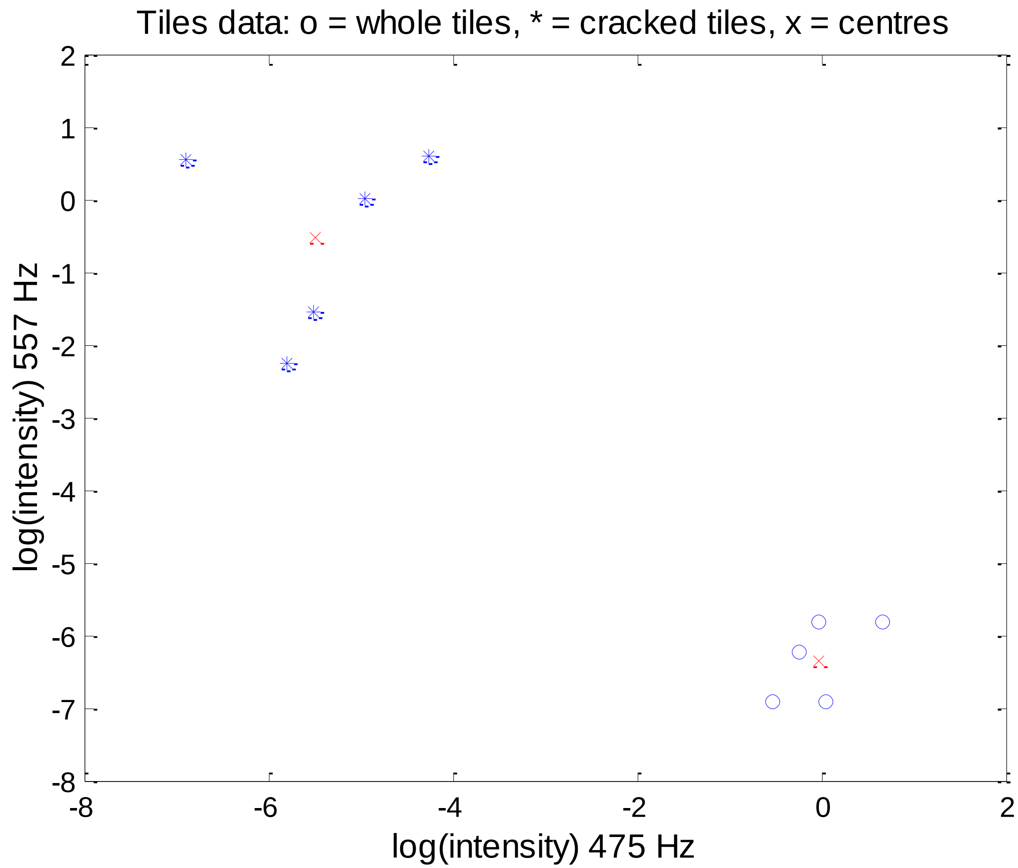

1. Compute the new centre of each class
2. Move the crosses (x)


Tiles data: o = whole tiles, * = cracked tiles, x = centres
log(intensity) 557 Hz
log(intensity) 475 Hz
2
1
0
-1
-2
-3
-4
-5
-6
-7
-8
-8
-6
-4
-2
0
2

Iteration 2


Tiles data: o = whole tiles, * = cracked tiles, x = centres
log(intensity) 557 Hz
log(intensity) 475 Hz
2
1
0
-1
-2
-3
-4
-5
-6
-7
-8
-8
-6
-4
-2
0
2

Iteration 3




Iteration 4 (then stop, because no visible change)

Each data point belongs to the cluster defined by the nearest centre

```
U =

    1     1     1     1     1     0     0     0     0     0
    0     0     0     0     0     1     1     1     1     1
```

The membership matrix U:

1. The last five data points (columns) belong to the first cluster (row)
2. The first five data points (column) belong to the second cluster (row)

# Porque agrupamentos difusos?

- Na maior parte das aplicações reais não existe uma separação rígida entre agrupamentos.

- Agrupamentos podem não ser separáveis devido ao ruído ou falta de capacidade discriminatória do espaço das características (“features”) que definem os padrões.

Agrupamento difusos suportam agrupamentos com fronteira sobrepostas (tal como nos conjuntos difusos).

# Agrupamentos Rígidos & Difusos

- Agrupamento rígido:

  - são agrupamentos disjuntos

  - cada dado pertence apenas a um agrupamento.

- Agrupamento Difusos:

  Cada ponto pode pertencer a mais que um agrupamento.

# Agrupamento baseado em Função Objetiva (Fuzzy Clustering)

- N padrões em $R^n$, c agrupamentos, um conjunto de protótipos $v_1, v_2, \ldots v_c$
- Função objectiva:

$$Q = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik}^{m} \left\| x_k - v_i \right\|^2$$

factor de fuzzification

$\| \quad \|^2$ distancia entre $x_k$ e $v_i$

U= matriz partição difusa

# Definições & Restrições de U

- U: uma matriz com valores confinados ao intervalo unitário

Restrições:

- Os agrupamentos são não triviais.

- A soma total dos graus de pertença de um dado ponto nos agrupamentos é 1.

# Fuzzy C-Means(FCM)

## Discrição do algoritmo

Algoritmo introduzido por Dunn em 1974 , melhorado por Bezdek em 1981

- Passo1 (Iniciar):

(a) selecionar os valores de: *c* (numero de agrupamentos), *m* (fator de fuzzificação), e ε (critério de paragem).

(b) escolher a função distância.

(c) Iniciar (aleatoriamente) a matriz partição.

- Passo2(ciclo principal):

  - calcular protótipos dos agrupamentos, $v_i$ para $i$ = 1, … , $c$

  - calcular a matriz de partição, $U=[u_{ik}]$ para $i$= 1,…,$c$ e $k$=1,…,$N$

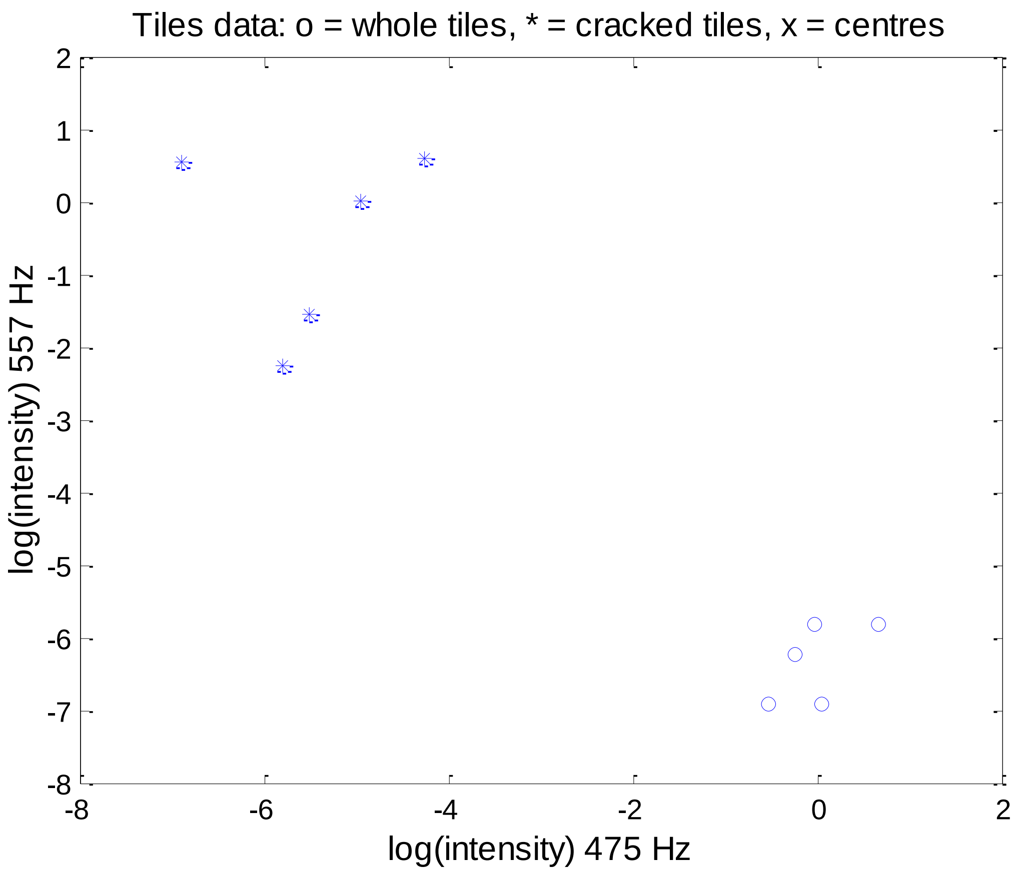
- Passo3: Critério de paragem:

$$\left\| U(iter+1) - U(iter) \right\| = \max_{i,k} \left| U_{ik}(iter+1) - U_{ik}(iter) \right| < \varepsilon$$

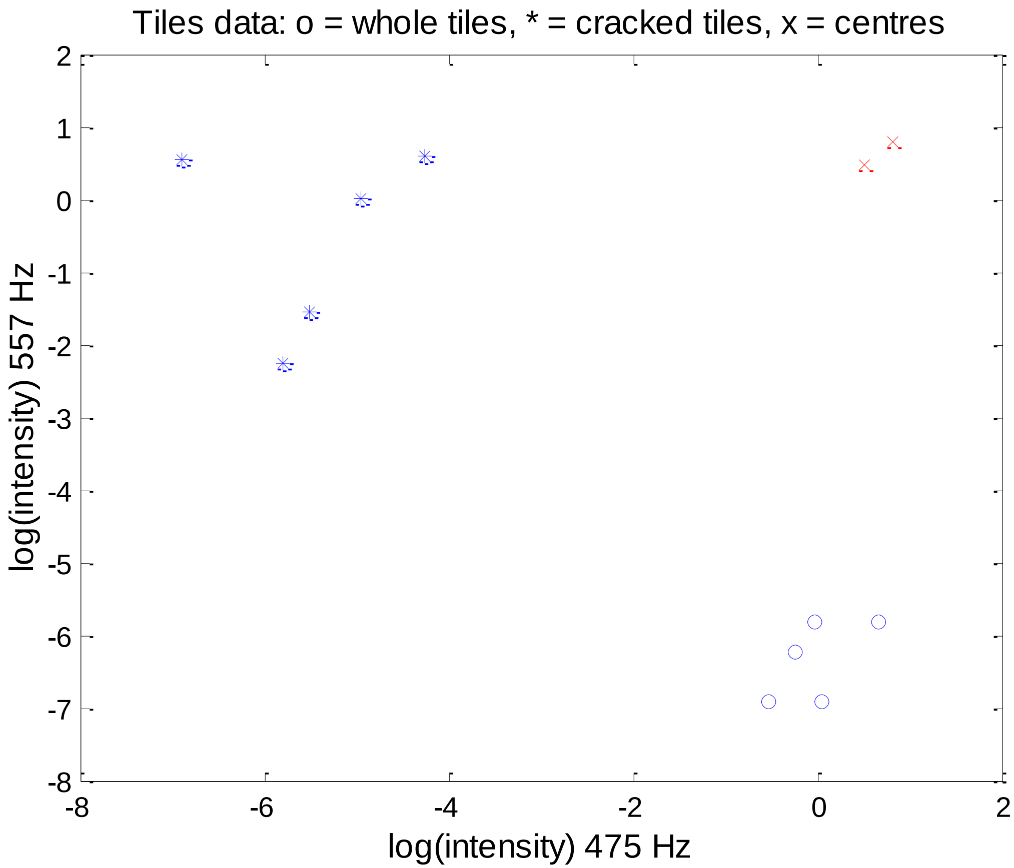
  - Se o critério de paragem for válido SAIR

  - Senão, voltar ao passo 2.

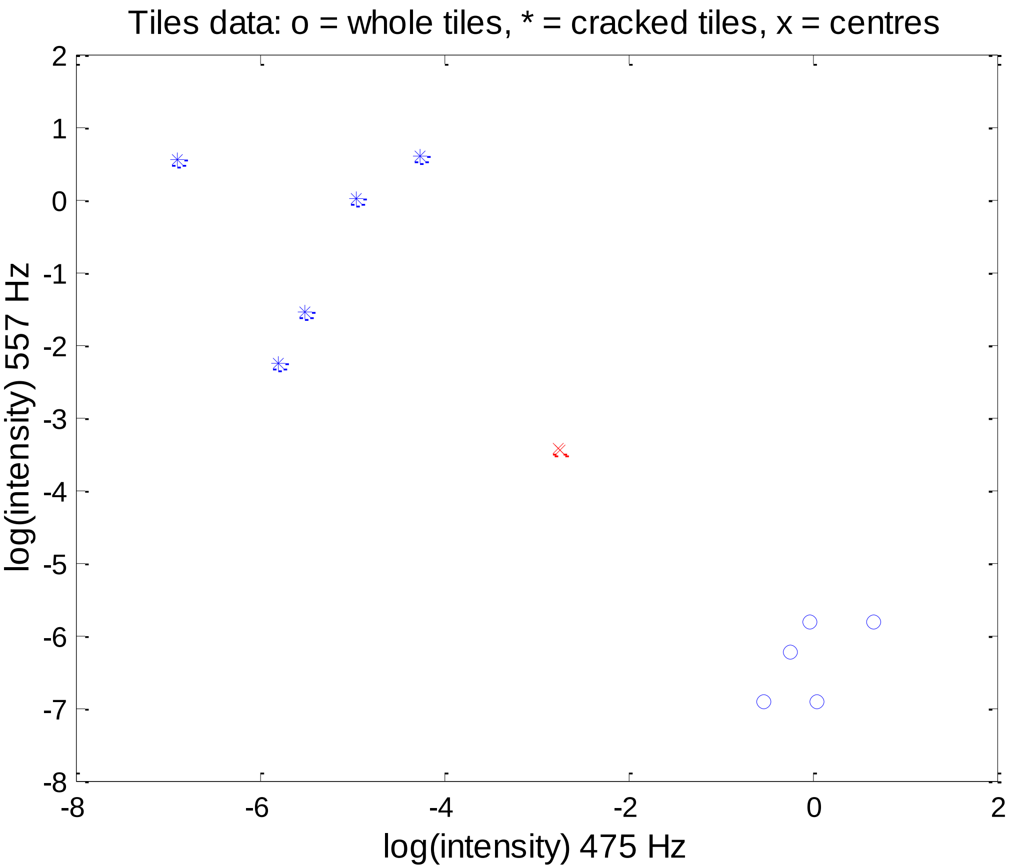
# Exemplo para FCM



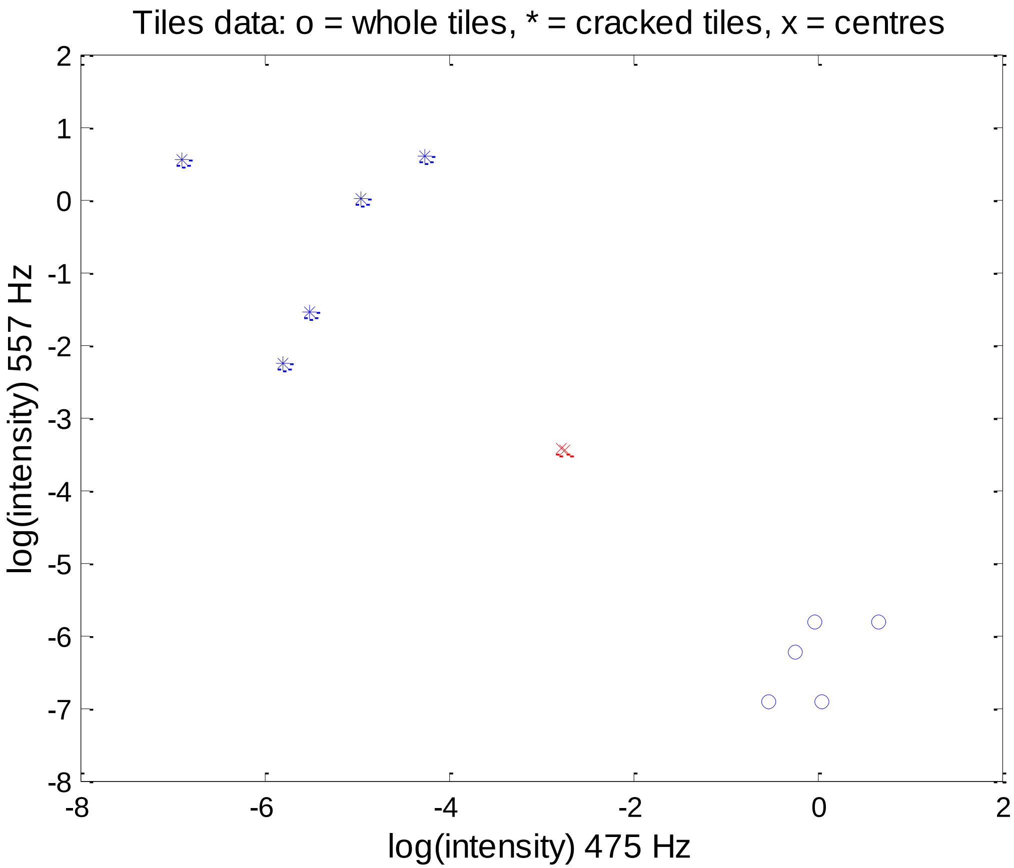

Each data point belongs to two clusters to different degrees




1. Place two cluster centres
2. Assign a fuzzy membership to each data point depending on distance

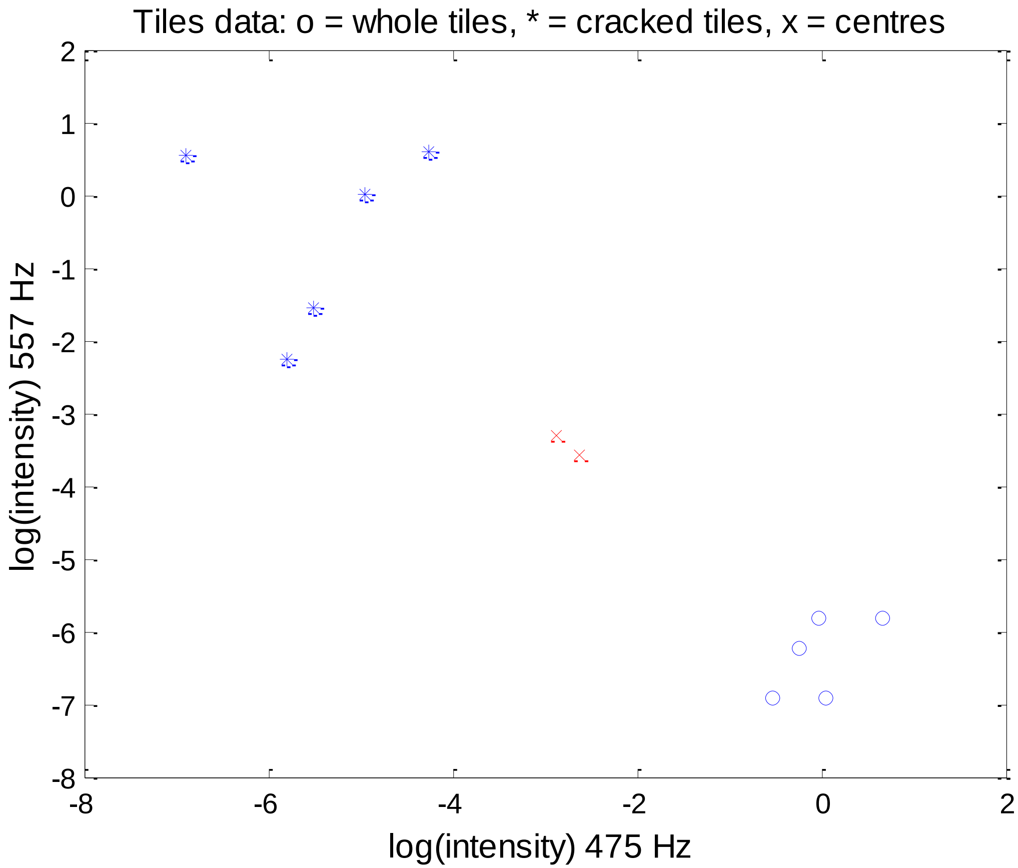



1. Compute the new centre of each class
2. Move the crosses (x)

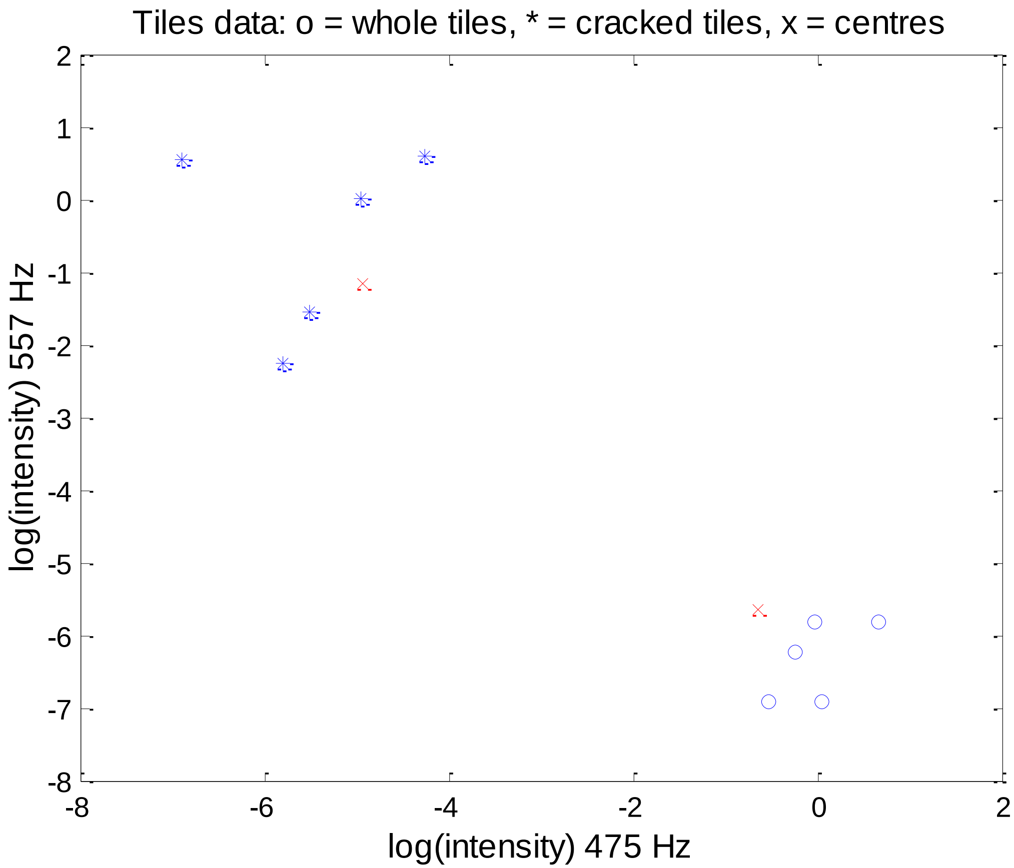


Tiles data: o = whole tiles, * = cracked tiles, x = centres
log(intensity) 557 Hz
log(intensity) 475 Hz
2
1
0
-1
-2
-3
-4
-5
-6
-7
-8
-8
-6
-4
-2
0
2

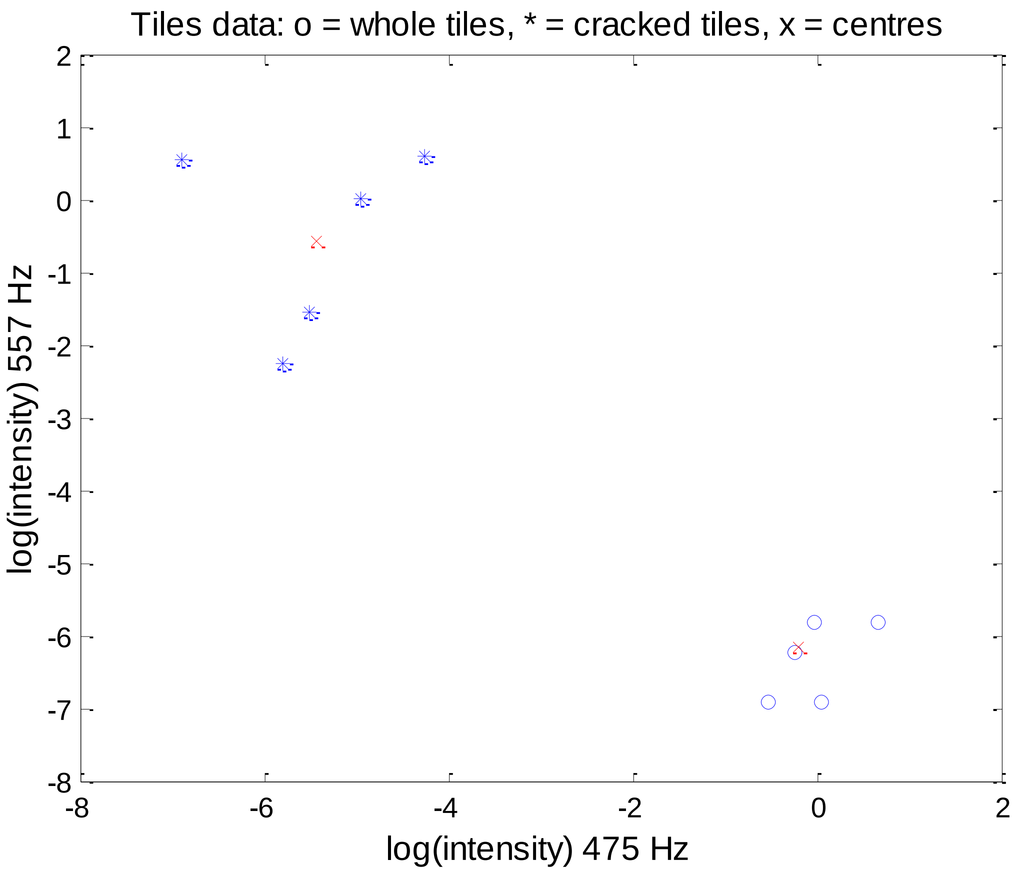

Iteration 2


Tiles data: o = whole tiles, * = cracked tiles, x = centres
2
1
0
-1
-2
-3
-4
-5
-6
-7
-8
log(intensity) 557 Hz
-8
-6
-4
-2
0
2
log(intensity) 475 Hz

Iteration 5




Iteration 10




Iteration 13 (then stop, because no visible change)

Each data point belongs to the two clusters to a degree

```
U =

0.0025  0.0091  0.0129  0.0001  0.0107  0.9393  0.9638  0.9574  0.9906  0.9807
0.9975  0.9909  0.9871  0.9999  0.9893  0.0607  0.0362  0.0426  0.0094  0.0193
```

The membership matrix U:

1. The last five data points (columns) belong mostly to the first cluster (row)
2. The first five data points (columns) belong mostly to the second cluster (row)

# Validação dos agrupamentos

Qual o número ótimo de agrupamentos?

- Índice de partição

[1/c, 1]

$$P(U) = \frac{1}{N}\sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^2$$

- Entropia da partição

[0, ln(c)]

$$H(U) = -\frac{1}{N}\sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}\ln(u_{ik})$$

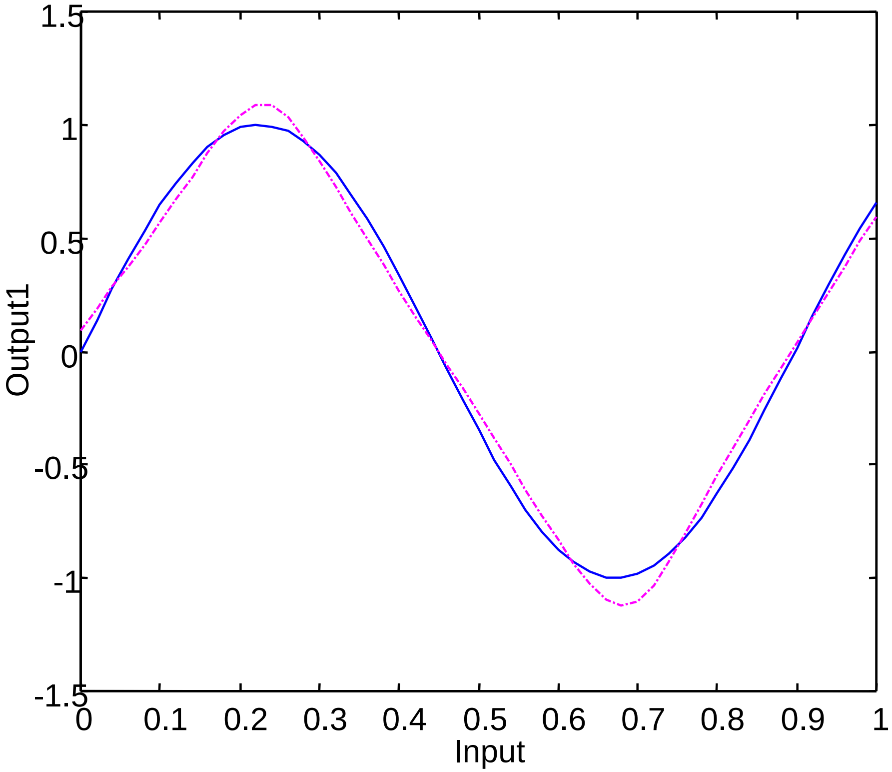
# Variantes de Agrupamentos Difusos

- Fuzzy C Variantes (Bezdek et al. 1981)

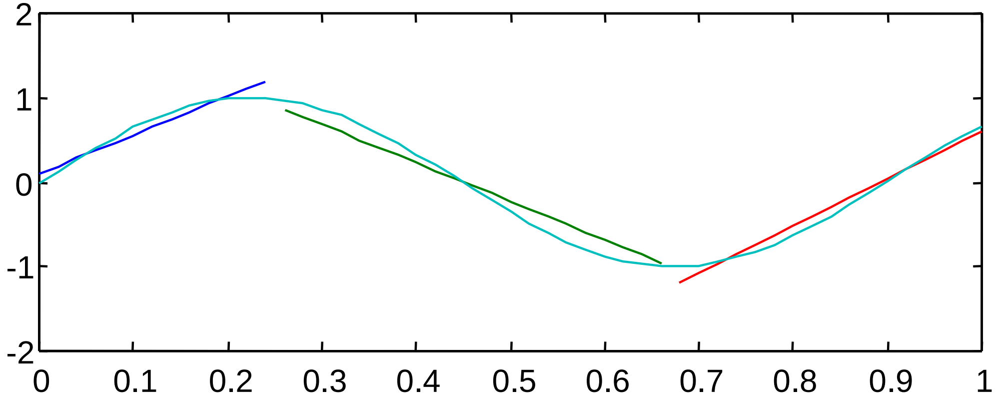
  pontos-> r- variedade dimensional
- Possibilistic Clustering(Krishnampuram 1993, keller 1996) sem a restrição da unicidade
- Noise Clustering (Ohashi 1984, Dave 1991)

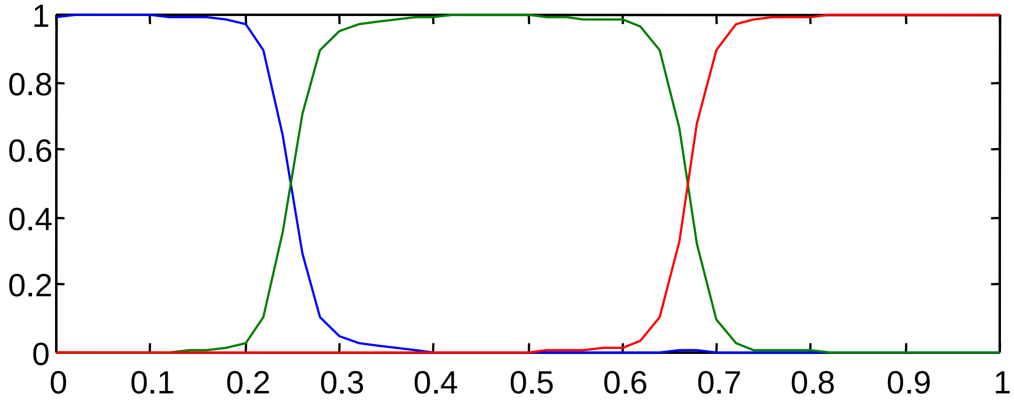
  “localiza” o ruído colocando-o num agrupamento à parte (passando a existirem c+1 agrupamentos)

# Exemplo: Aproximação de Funções

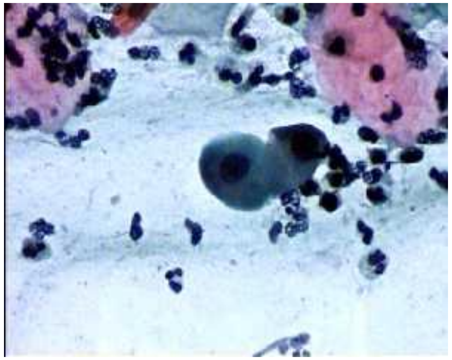


# Aproximação por conjuntos difusos (fuzzy sets)

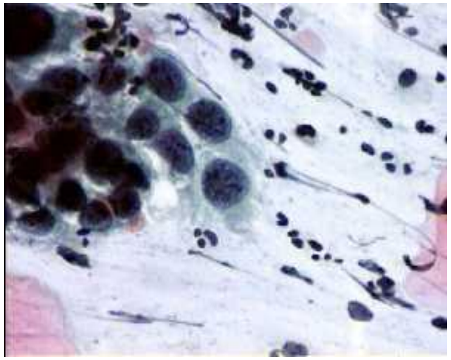


# Exemplo: Classifcar células cancerosas



Normal smear



Severely dysplastic smear

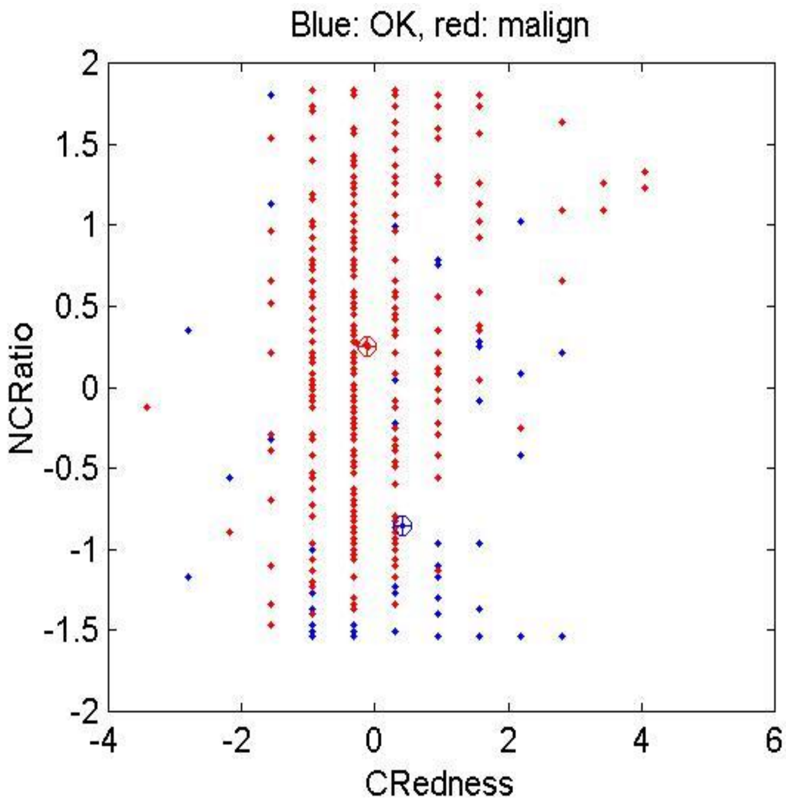
Using a small brush, cotton stick, or wooden stick, a specimen is taken from the uterin cervix and smeared onto a thin, rectangular glass plate, a slide. The purpose of the smear screening is to diagnose pre-malignant cell changes before they progress to cancer. The smear is stained using the Papanicolau method, hence the name *Pap smear*. Different characteristics have different colours, easy to distinguish in a microscope. A cyto-technician performs the screening in a microscope. It is time consuming and prone to error, as each slide may contain up to 300.000 cells.

Dysplastic cells have undergone precancerous changes. They generally have longer and darker nuclei, and they have a tendency to cling together in large clusters. Mildly dysplastic cels have enlarged and bright nuclei. Moderately dysplastic cells have larger and darker nuclei. Severely dysplastic cells have large, dark, and often oddly shaped nuclei. The cytoplasm is dark, and it is relatively small.

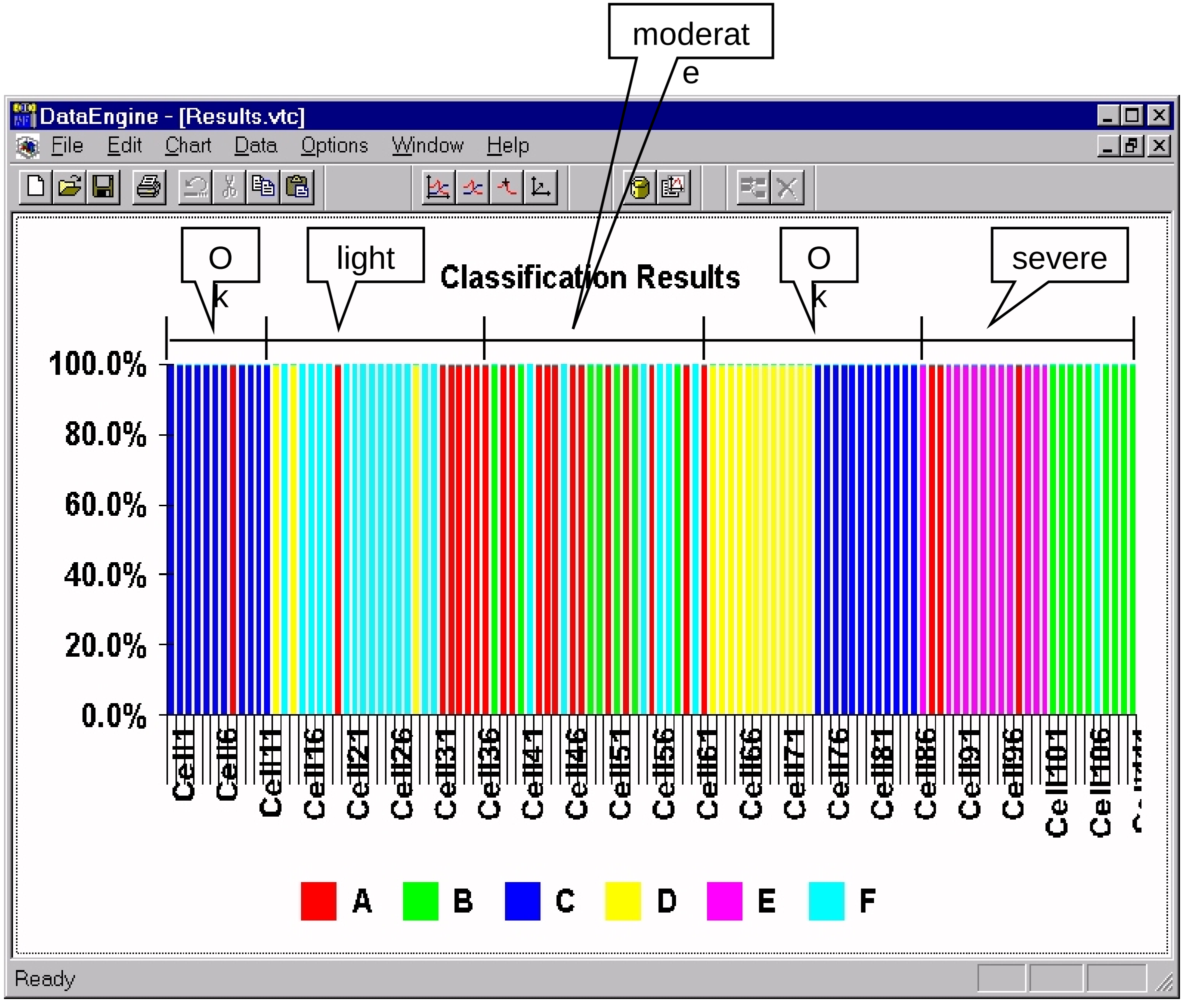
# Aspectos possíveis

- Nucleus and cytoplasm area
- Nucleus and cyto brightness
- Nucleus shortest and longest diameter
- Cyto shortest and longest diameter
- Nucleus and cyto perimeter
- Nucleus and cyto no of maxima
- (...)

# Classes que não são separáveis!

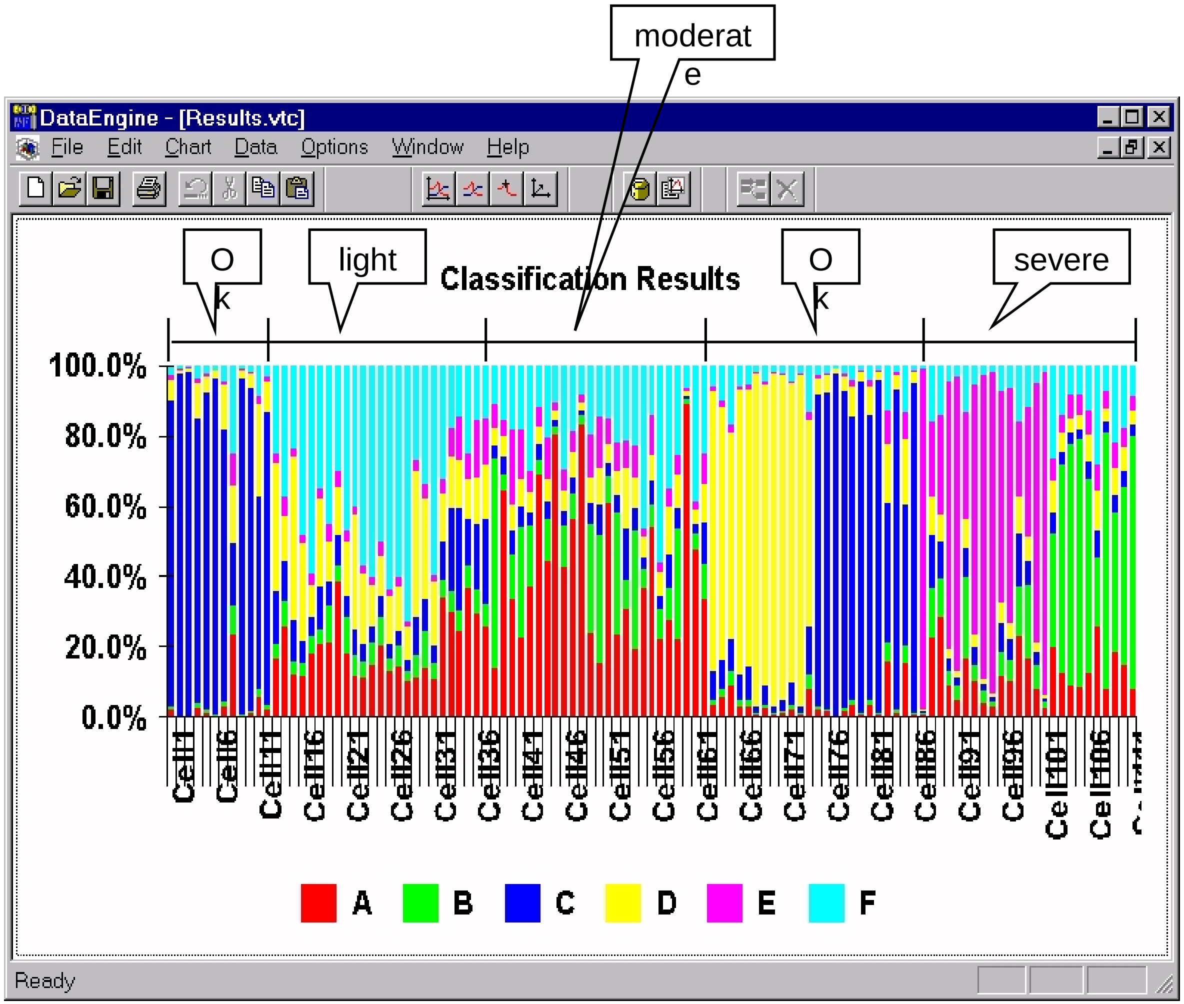


# Classificador Rígido (HCM)




A cell is either one or the other class defined by a colour.

# Cassificador Difuso (FCM)




A cell can belong to several classes to a Degree, i.e., one column may have several colours.

# Sumario

- Agrupamento Hierárquico & por Função-Objectiva

| **Agrupamento Hierárquico** | **Agrupamento baseado em função-objetiva** |
|---|---|
| O número de agrupamentos não é especificado à priori | É necessário especificar no inicio do processo o nº de agrupamentos |
| Não existe conhecimento acerca da forma ou tamanho do agrupamento | É conhecido a forma global ou tamanho do agrupamento. |
| Agrupamento estático | Agrupamento dinâmico |

# Referências


| | |
|---|---|
| [Bez81] | J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum, NY, 1981. |
| [BezKKP99] | James C. Bezdek, James Keller, Raghu Krishnapuram and Nikhil R. Pal, *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*, Kluwer Academic Publishers, TA 1650.F89, 1999. |
| [KriK93] | R. Krishnapuram and J. M. Keller, "A possibilistic approach to clustering," *IEEE Transactions on Fuzzy Systems*, Vol. 1, No. 2, pp. 98-110, May 1993. |
| [PalPB97] | N. R. Pal, K. Pal and J. C. Bezdek, "A mixed c-means clustering model," *Proceedings of the Sixth IEEE International Conference on Fuzzy Systems*, Vol. 1, pp. 11-21, Jul. 1997. |
| [YanRP94] | Jun Yan, Michael Ryan and James Power, *Using fuzzy logic Towards intelligent systems*, Prentice Hall, 1994. |