Conhecimento e Raciocínio

Aula 4

Lógica Difusa

Viriato A.P. Marinho Marques
DEIS - ISEC
2020 / 2021

9.1 Métodos de Representação

9.1 Métodos de Representação

1. **Probabilidade**
    1. Redes Bayesianas
    2. Árvores de Decisão Bayesianas (PROSPECTOR)
    3. Cadeias de Markov

2. **Teoria de Dempster-Shaffer**

3. **Factores de Certeza** (*Certainty Factors - CF*)
    1. MYCIN
    2. CLIPS
    3. EXSYS: Soma, Multiplicação, Max, Min, Média, Mycin...

Módulo II

4. **Lógica Difusa**
    1. Inferência de **Mamdani** / *JFS freeware*
    2. Regra Composicional de Inferência / *Sistema CADIAG-2*
    3. *Fuzzy Pattern Matching / Fuzzy CLIPS*

9.2 Conjuntos Difusos

9.2 Conjuntos Difusos

O conceito de **Conjunto Difuso** foi introduzido por L.A. Zadeh nos anos 60. Com base neste conceito surgiram posteriormente:

- **Números** e **Intervalos Difusos**
- **Computação com Palavras** (*Computing with Words* - ***CW***)
- **Teoria da Possibilidade**
- **Lógica Difusa**.

A Lógica Difusa provou ser uma base de trabalho muito consentânea com a realidade. É utilizada com grande sucesso como base de processos de inferência em

- **Sistemas de Controlo**
- **Sistemas Periciais**

9.2 Conjuntos Difusos

Conjunto Difuso: Um conjunto a que cada elemento pode pertencer "não completamente" mas apenas "parcialmente"

- O grau de pertença é expresso por um número ou por uma função $\mu(x)$
- Varia entre 0 e 1.

Formalmente:

- Se X é uma colecção de objectos denotados genericamente por x, um **conjunto difuso $\tilde{A}$** definido em X é um conjunto de pares ordenados da forma

$$\left(x, \mu_{\tilde{A}}(x)\right) \qquad \mu_{\tilde{A}}(x) \in [0,1]$$

- A função μ chama-se Função de Pertença
- Representa o grau de pertença do elemento x ao conjunto difuso

Exemplo

Conjunto de inteiros próximos de 1: $\tilde{A} = \{(0,0.5),(1,1),(2,0.5),(3,0.25)\}$

9.2 Conjuntos Difusos

A Função de Pertença pode ser contínua:

- Neste caso o conjunto difuso representa-se na forma $\tilde{A} = \left(x, \mu_{\tilde{A}}(x)\right)$

Exemplo

Conjunto de reais próximos de 25

$$\tilde{A} = \left\{x, \mu(x) \mid \mu(x) = 1/\left[1 + (x-25)^2\right]\right\}$$

(Portanto, se $x=25$, $\mu=1$)

Operações com Conjuntos Difusos

Segundo Zadeh:

- Intersecção $\qquad \tilde{C} = \tilde{A} \cap \tilde{B} \qquad \mu_{\tilde{C}}(x) = \min\left\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\right\}$
- Reunião $\qquad \tilde{D} = \tilde{A} \cup \tilde{B} \qquad \mu_{\tilde{C}}(x) = \max\left\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\right\}$
- Complementação $\qquad \tilde{E} = \not\subset \tilde{A} \qquad \mu_{\tilde{B}}(x) = 1 - \mu_{\tilde{A}}(x)$

NOTA: Portanto, dado um Universo U, um seu elemento que não figure num conjunto nele definido, deverá aparecer no seu complementar com grau de pertença 1.

9.2 Conjuntos Difusos

Portanto:

- **A∩B**
  Elementos comuns a A e B sendo μ de cada um $\mu_{A \cap B} = min(\mu_A(x), \mu_B(x))$
- **A∪B**
  Elementos comuns e não comuns a A e B sendo $\mu_{A \cup B} = max(\mu_A(x), \mu_B(x))$
- **/A**
  Todos os elementos do **Universo** de A, sendo μ de cada um $\mu_{/A} = 1 - \mu_A$

Exemplos

- Seja X o conjunto de modelos de casas possíveis, considerando o número de quartos de 1 a 10.
- Seja $\mu(x)$ a função de pertença de $\tilde{A}$, definida pela condição difusa "*a casa do tipo x é confortável para uma família de 4 pessoas*":

Ã = {(1,0.2), (2,0.5), (3,0.8), (4,1), (5,0.7), (6,0.3)}

9.2 Conjuntos Difusos

Dados Ã e Ñ

Ã = Casas confortáveis para 4 pessoas =

= {(1,0.2), (2,0.5), (3,0.8), (4,1), (5,0.7), (6,0.3)}

Ñ = Casas grandes =
= { (3,0.2), (4,0.4), (5,0.6), (6,0.8), (7,1), (8,1), (9,1), (10,1)}

os seguintes conjuntos difusos têm os respectivos significados:

Casas grandes E confortáveis para 4 pessoas =
= Ã∩Ñ =
= {(3, 0.2), (4,0.4), (5,0.6), (6,0.3)}

Casas grandes OU confortáveis para 4 pessoas=

=Ã∪Ñ =
={(1,0.2), (2,0.5), (3,0.8), (4,1), (5,0.7), (6,0.8), (7,1), (8,1), (9,1), (10,1) }

Casas pequenas=

= /Ñ =

= {(1,1), (2,1), (3,0.8), (4,0.6), (5,0.4), (6,0.2)}

9.3 Números Difusos

9.3 Números Difusos

Um Número Difuso $\tilde{N}$ representa um valor aproximado de N, em que N se chama Valor Médio

- A função de pertença de um número difuso tem apenas um máximo, e sempre de valor 1
- As funções de pertença normalmente utilizadas para números difusos são triangulares
- As trapezoidais usam-se para Intervalo Difusos.

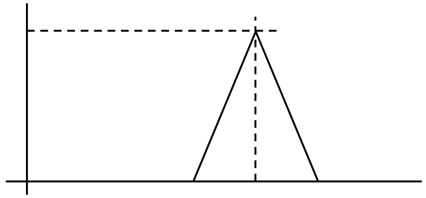
Exemplo

Fuzzy Number (cerca de) 49:

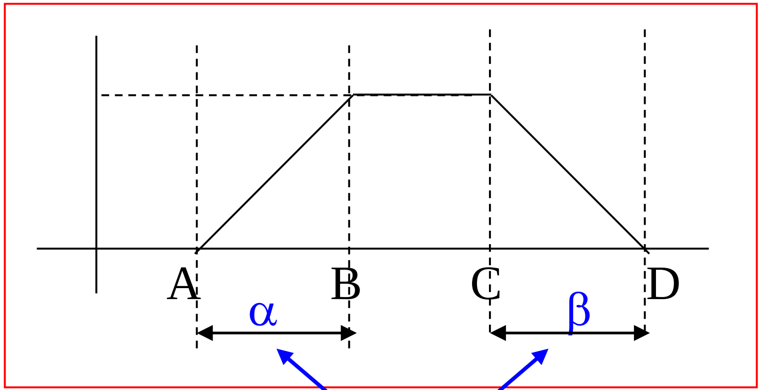
$\sim 49 = \{(x, \mu_{\sim 49}(x)) \mid x \in |R, \mu_{\sim 49}(x)\}$

Aproximação

$\Rightarrow$



9.3 Números Difusos



Na notação LR (Left-Right) são dados B, C e as aberturas α e β

A notação ***Alfa-Cut*** é aplicável **apenas** a funções de pertença trapezoidais (e triangulares) e simplifica muito os cálculos:

- Um intervalo difuso é representado pelas 4 abcissas A, B, C e D dos pontos em que $\mu=0$, $\mu=1$, $\mu=1$ e $\mu=0$ (trapézios).
- Para um número difuso de forma triangular, $B=C$

Operações Aritméticas

Por aplicação do chamado Teorema da Extensão, as operações com **Números Crespos** (*Crisp Numbers*) podem ser estendidas aos **Números Difusos**:

9.3 Números Difusos

O cálculo pela definição é muito moroso:

Exemplo

Adição: $\mu_{\tilde{A} \oplus \tilde{N}} = \sup_{z = x + y} \min [\mu_{\tilde{A}}(x), \mu_{\tilde{N}}(y)]$

Sejam os números difusos Ã e Ñ

Para cada x_i de $\tilde{A}$

  Para cada y_j de Ñ

    Calcular $z_i = x_i + y_j$

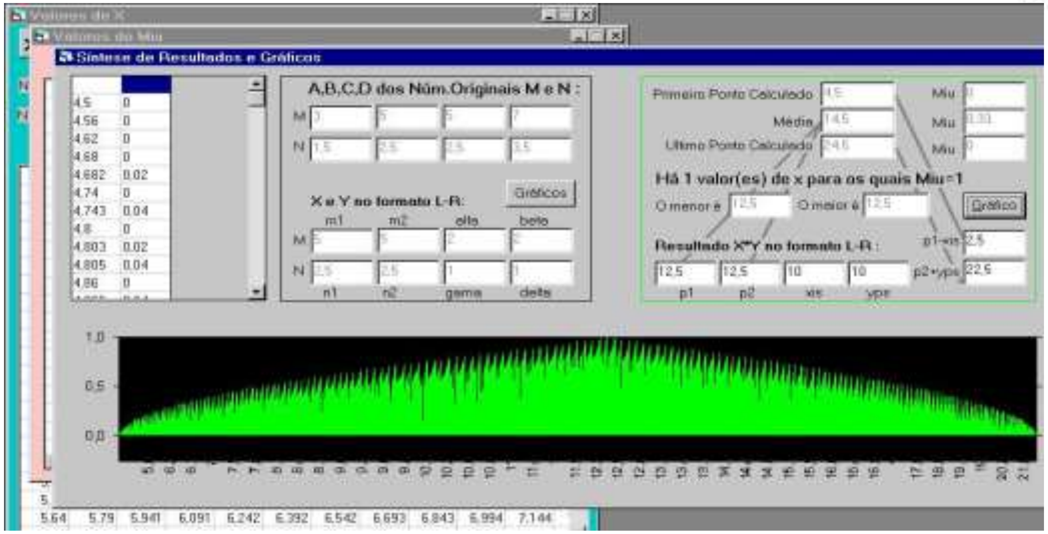
    Calcular $min(\mu(x)_i , \mu(y)_j)$

Fim Para

Para cada z_i

  Determinar $\mu(z)_j = max(\mu(z)_i)$

Resultado: O conjunto dos $\{z_i, \mu(z)_i\}$



Prós:

- Funciona para qualquer forma de função de pertença

Contras:

- Muito moroso
- Obriga a discretizar as funções de pertença contínuas

Aspecto gráfico da adição de dois números difusos executada pela definição

9.3 Números Difusos

As operações na notação Alfa-Cut ou LR, com formas trapezoidais, são

- Compreensíveis a partir da aritmética de intervalos ***crespos***
- Facilmente implementadas

Para intervalos crespos [a,b] e [c,d] os extremos dos intervalos resultantes são:

$[a,b] + [c,d] = [a+c, b+d]$

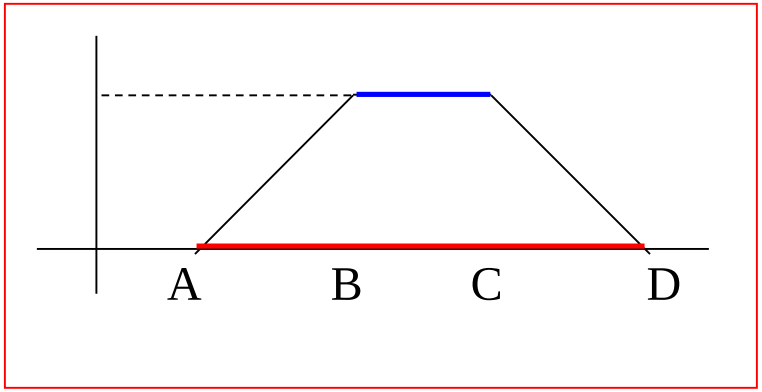
$[a,b] - [c,d] = [a-d, b-c]$

$[a,b] * [c,d] = [\min(ac, ad, bc, bd), \max(ac, ad, bc, bd)]$

$[a,b] / [c,d] = [a,b] * [1/d, 1/c] = [\min(a/d, a/c, b/d, b/c), \max(a/d, a/c, b/d, b/c)]$

$$\alpha.[a,b] \begin{array}{ll} = [\alpha a, \alpha b] & \text{se } \alpha>0 \\ = [\alpha b, \alpha a] & \text{se } \alpha<0 \end{array}$$

Para efeitos de cálculo, um intervalo difuso pode ser encarado como definido pelos extremos de dois intervalos crespos:



9.3 Números Difusos

Portanto, aplicando "duas vezes" as fórmulas anteriores, obtém-se:

1. Adição:

(m0, m1, m2, m3) + (n0, n1, n2, n3) = (m0+n0, m1+n1, m2+n2, m3+n3)

"o limite menor" menos *"o limite maior"* dá *"o limite menor"* (o mais à esquerda)

2. Subtracção:

(m0, m1, m2, m3) - (n0, n1, n2, n3) = (m0-n3, m1-n2, m2-n1, m3-n0)

Este processo repete-se para *"o intervalo interno [m1,m2]"*

Simétrico para *"o limite maior"*

3. Multiplicação:

(m0, m1, m2, m3) x (n0, n1, n2, n3) =
= (**min**(m0n0, m0n3, m3n0, m3n3), **min**(m1n1, m1n2, m2n1, m2n2),
max(m1n1, m1n2, m2n1, m2n2), **max**(m0n0, m0n3, m3n0, m3n3))

Na multiplicação nada se pode dizer acerca do resultado dos produtos. Portanto, *"os limites menores"* têm de recorrer ao ***min*** ...

...e os "maiores" ao ***max***

4. Divisão :

(m0, m1, m2, m3) / (n0, n1, n2, n3) =
= (min(m0/n0, m0/n3, m3/n0, m3/n3), min(m1/n1, m1/n2, m2/n1, m2/n2), min(m1/n1, m1/n2, m2/n1, m2/n2), max(m0/n0, m0/n3, m3/n0, m3/n3))

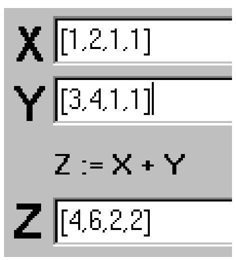
A divisão obtém-se a partir da multiplicação

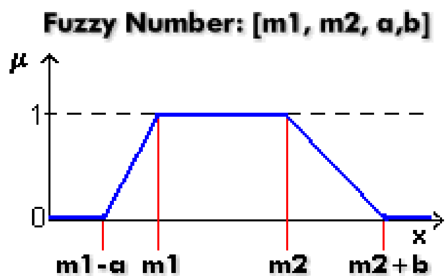
9.3 Números Difusos

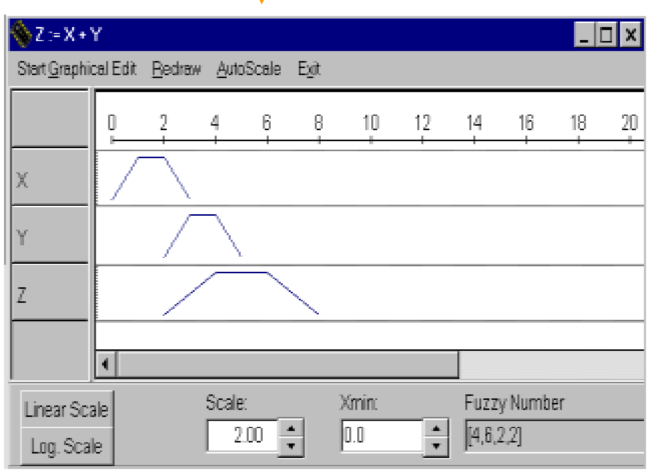
Exemplo:

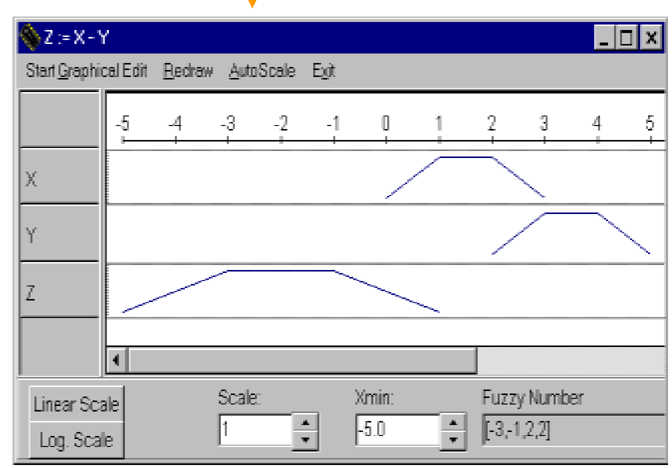
Adição e Subtracção de 2 números difusos

α-cut	(0,1,2,3) + (2,3,4,5) = (2,4,6,8)	(0,1,2,3) - (2,3,4,5) = (-5,-3,-1,-2)
LR	(1,2,1,1) + (3,4,1,1) = (4,6,2,2)	(1,2,1,1) - (3,4,1,1) = (4,6,2,2)









9.4 Computação com Palavras

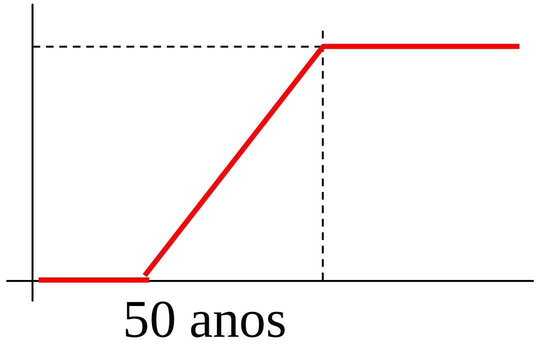
9.4 Computação com Palavras (*Computing with Words - CW*)

Conjuntos difusos podem representar termos linguísticos: basta atribuir um **valor médio** e uma **difusão** à tradução do "conceito" pretendido

Exemplo

- Velho: $\tilde{N}_{\mathbf{(velho)}} = \{(u, \mu_{velho}(u)) \mid u \in [0,100]\}$
  $\mu_{velho}(u) = 0 \;\text{ se } \mu \in [0,50]$
  $\mu_{velho}(u) = 1 / (1+((u-50)/5)^{-2}) \text{ se } \mu \in]50,100]$

Na aproximação por trapézios teríamos qualquer coisa deste tipo:



Neste exemplo a **Variável Linguística** *"idade"* foi adjectivada pelo termo Alta gerando o termo Velho.

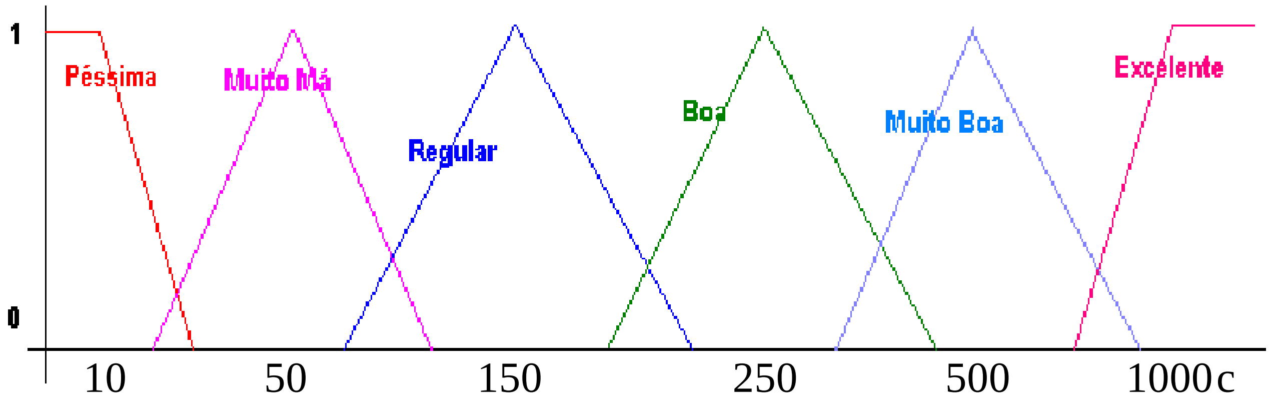
9.4 Computação com Palavras

Deste modo um **vocabulário** pode ser associado a qualquer variáv linguística:

- Cada termo linguístico é uma **graduação da variável**
- Cada termo linguístico é representado por um **conjunto difuso**

Exemplo:

Seja a variável **profissão** classificada em função do **rendimento** que gera. Poderia ser:



(não desenhado à escala)

A classificação é subjectiva: pode recorrer-se a estatísticas e existem métodos para desenhar os conjuntos difusos a partir dos resultados obtidos *(ver Zimmerman)*

9.4 Computação com Palavras

Numa escala unitária, de modo semelhante, podem também representar-se:

- **Quantificadores**: Todos, quase todos...alguns...nenhum
- **Qualificadores de Frequência**: Sempre,...Por Vezes...Nunca

Além de qualificadores e quantificadores, a linguagem comum recorre a **Modificadores** (*hedge modifiers*):

- **Muito**: É normalmente traduzido pela **função condensação**, que eleva ao quadrado a função μ do termo original:

  Muito(velho) = $\{(u, \mu^{2}_{velho}(u)) \mid u \in [0,100]\}$

  $\mu_{velho}(u) = 0$ se $\mu \in [0,50]$

  $\mu_{velho}(u) = 1 / (1+((u-50)/5)^{-2})^{2}$ se $\mu \in]50,100]$

- **Aproximadamente ou Cerca de**: É normalmente traduzido pela **função dilatação**, que calcula a raiz da função μ do termo original:

  Aproximadamente(velho) = $\{(u, \mu^{1/2}_{velho}(u)) \mid u \in [0,100]\}$

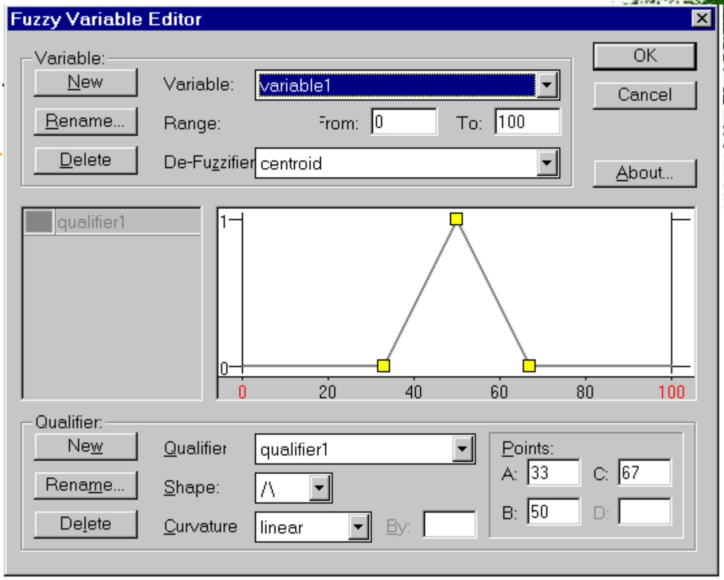
  $\mu_{velho}(u) = 0$ se $\mu \in [0,50]$

  $\mu_{velho}(u) = 1 / (1+((u-50)/5)^{-2})^{1/2}$ se $\mu \in]50,100]$

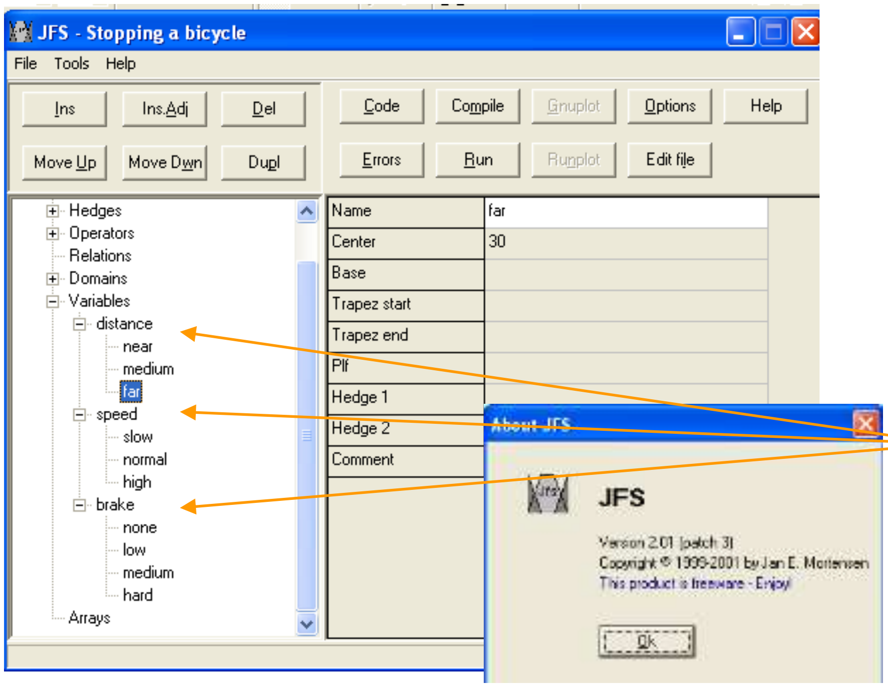
9.4 Computação com Palavras

Exemplo: FLINT (Uma ferramenta do LPA-Prolog)

Definição das funções de pertença de Termos Linguísticos usando uma opção do FLINT



Exemplo: JFS (*freeware*)



Controlar a travagem de um veículo
Definição de termos linguísticos como graduações das variáveis *distância*, *velocidade* e *intensidade de travagem*

9.5 Inferência Difusa

9.5 Inferência Difusa

Com conjuntos, números difusos e termos linguísticos, torna-se fácil representar regras **If...Then** em correspondência directa com a realidade. Duas aplicações típicas são Sistemas de Controlo e Sistemas Periciais:

Sistemas Periciais:

- **Diagnóstico:** SE a febre é muito(alta) e a tosse é seca ENTÃO é gripe
- **Crédito:** SE o rendimento é (baixo OU muito(baixo)) OU
  SE o rendimento é (médio OU alto) E o débito é muito alto
  ENTÃO o risco é grande

Sistemas de Controlo:

- ABS: SE roda bloqueada ENTÃO eliminar pressão no disco
  SE roda por vezes(bloqueda) ENTÃO diminuir pressão
- Caldeira: SE temperatura alta E pressão alta
  ENTÃO baixar muito a alimentação de gás

9.5 Inferência Difusa

Têm sido definidos várias inferências, em si. Vamos ver duas:

- A **Inferência de Mandani**, muito usada em sistemas de controlo
- A **Regra Composicional de Inferência** (RCI), de Zadeh, usada p.e. em sistemas de diagnóstico

Qualquer dos métodos permite obter conclusões com um *grau de verdade* variável, em função dos graus de verdade da premissa.

Exemplo:

SE o **rendimento** é **baixo** E o débito é **muito(alto)** ENTÃO o risco é **grande**

Premissa:

Duas cláusulas cujos valores de μ (termo *"rendimento baixo"* e *"débito muito(alto)"*) expressam *"quanto o rendimento é baixo"* e *"o débito é alto"*

Conclusão:

Alguma forma de expressar *"quanto o risco é grande"*

9.5 Inferência Difusa

Note-se que:

- A semântica *"quanto o rendimento é baixo"* é a mesma de *"quanto é verdadeiro que o rendimento seja baixo"*
- A semântica *"quanto o débito é alto"* é a mesma de *"quanto é verdadeiro que o débito seja alto"*
- A semântica *"quanto o risco é grande"* é a mesma de *"quanto é verdadeiro que o risco seja grande"*

Estamos em presença de um tipo de lógica no qual, entre verdadeiro e falso, se admite a existência de graduações:

- As inferências são feitas com base em termos linguísticos
- A incerteza inerente a este tipo de representações e o **Raciocínio Aproximado** (*aproximate reasoning*) nada tem a ver com probabilidade. Esta incerteza expressa **possibilidade**

De termos linguísticos, graus de verdade e possibilidade, derivam a **Lógica Difusa** e a **Teoria da Possibilidade**.

9.5.1 Inferência de Mamdani

9.5.1 Inferência de Mamdani

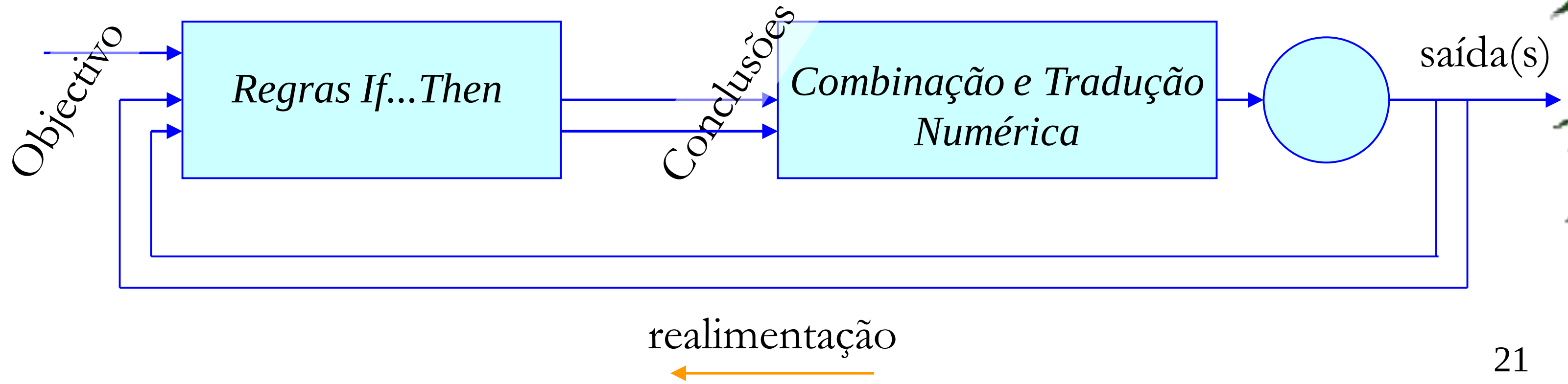
Num sistema de controlo os sinais de erro, numéricos, representam uma diferença entre o objectivo pretendido e a saída actual:

1) Para realizar as inferências com base em termos linguísticos, os sinais de erro são transformados em termos linguísticos

2) Realiza-se uma inferência para cada regra IF...THEN

3) O resultado de cada uma tem de ser combinado num único conjunto difuso

4) Esse conjunto difuso tem de ser transformado num valor numérico para aplicação aos actuadores



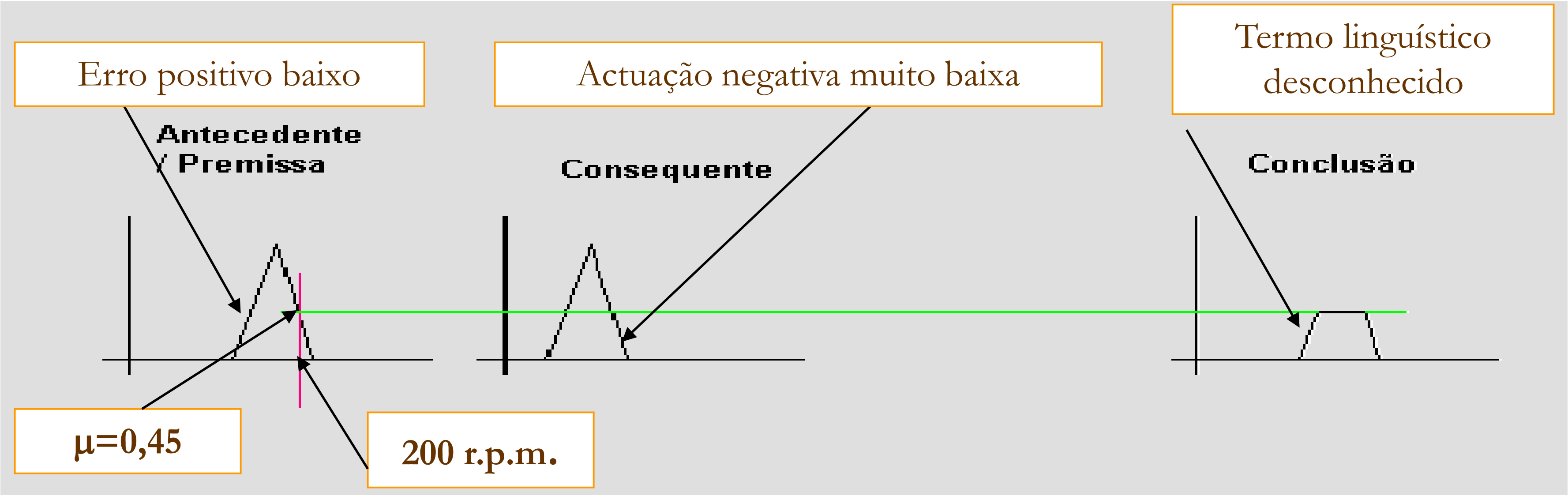
9.5.1 Inferência de Mamdani

Na inferência segundo Mandani o μ da **conclusão** obtém-se **cortando o termo linguístico do consequente** pelo valor de μ do **antecedente**

- Se houver mais que 1 antecedente (ligados pelo operador E) determina-se o mínimo desses μ

Exemplo: Num sistema de controlo de velocidade de um motor, suponhamos que 200r.p.m. corresponde a um *erro positivo baixo* e que a regra é

"Se erro positivo baixo então diminuir muito pouco a velocidade"



9.5.1 Inferência de Mamdani

Exemplo: Inferência de Mandani num Fuzzy ES

Suponhamos um ES destinado a calcular o limite de crédito a conceder a um cliente de um banco. Sejam 4 regras:

1. SE o candidato é novo E o rendimento é baixo ENTÃO o crédito é baixo
2. SE o candidato tem outros créditos, ENTÃO o crédito é baixo
3. Se o candidato tem bens próprios, ENTÃO o crédito é alto
4. Se o candidato sofre de doença crónica, ENTÃO o crédito é médio

Fuzzificação

Regra 1:	Fuzzificar idade do candidato (dada em anos)	μ(idade)=0,8
	Fuzzificar rendimento (dado em \$/mês)	μ(\$)=0,2
Regras 2, 3 e 4:	Fuzzificadas por método análogo	

Inferência

Regra 1:	A condição **E** é avaliada como min(μ(idade), μ(\$))=min(0.8,0.2)=0,2 Obtém-se assim μ(Regra1)=0,2 porque o termo crédito baixo é "cortado" por μ(Regra1), o ***min*** de entre os dois
Regras 2, 3 e 4:	O valor de outros créditos, bens próprios e gravidade da doença crónica são fuzzificados pelo mesmo processo. Suponhamos que para a regra 2. se obteve μ(Regra2)=0,3

9.5.1 Inferência de Mamdani

Agregação

As regras 1 e 2 têm a mesma conclusão (crédito baixo). O operador subentendido entre elas é o OU:

1. SE o candidato é novo E o rendimento é baixo ENTÃO o crédito é baixo

OU

2. SE o candidato tem outros créditos, ENTÃO o crédito é baixo

Como se trata de uma disjunção, entre as conclusões geradas por uma e outra determina-se $\mu(1,2) = \max(\mu(\text{Regra1}), \mu(\text{Regra2})) = \max(0.2, 0.3) = 0.3$

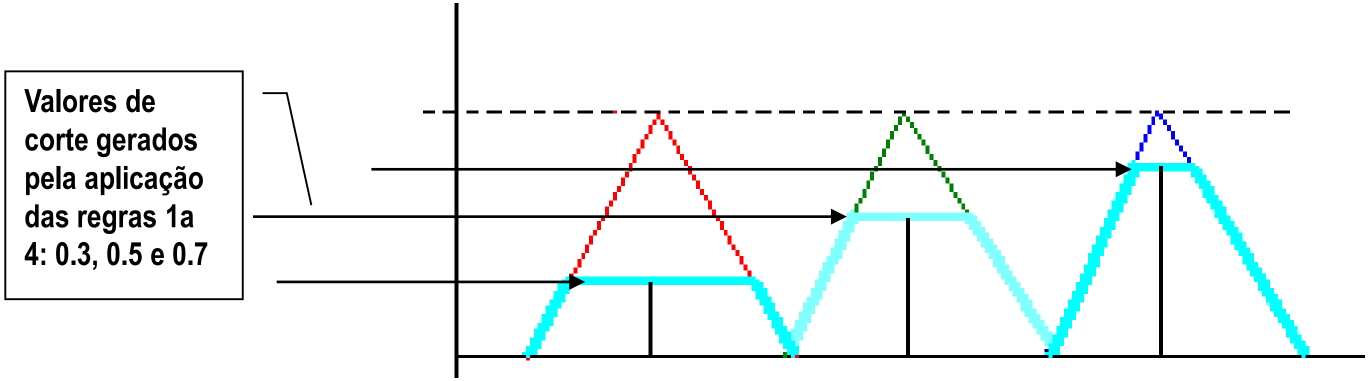
O conjunto de operações ***min*** e ***max*** utilizadas para interpretar as regras 1. e 2. designa-se por Método MAX-MIN: o ***min*** é utilizado nos antecedentes e o ***max*** na agregação de conclusões de duas (ou mais) regras que têm igual conclusão

Desfuzzificação

As regras 1 e 2 geraram um único termo conclusão (crédito baixo cortado a 0.3). Suponhamos que:

- A regra 3 gerou como conclusão crédito alto cortado a 0.7
- A regra 4 gerou como conclusão crédito médio cortado a 0.5

9.5.1 Inferência de Mamdani



Para determinar o crédito a conceder, estas conclusões têm de ser combinadas num só número crespo final. É isto a **Desfuzzificação**.

O método mais usado é o **COA** (*Center of Area*) ou Centróide. Consiste em determinar a **abcissa do baricentro**. As dimensões do resultado são $ e portanto representa o valor do crédito a conceder.

Para μ discreta: $$\Re = \frac{\sum_{i=0}^{n} d_i . \mu_A(d_i)}{\sum_{i=0}^{n} \mu_A(d_i)}$$

Para μ contínua: $$\Re = \frac{\int_{u1}^{u2} u.\mu(u)du}{\int_{u1}^{u2} \mu(u)du}$$

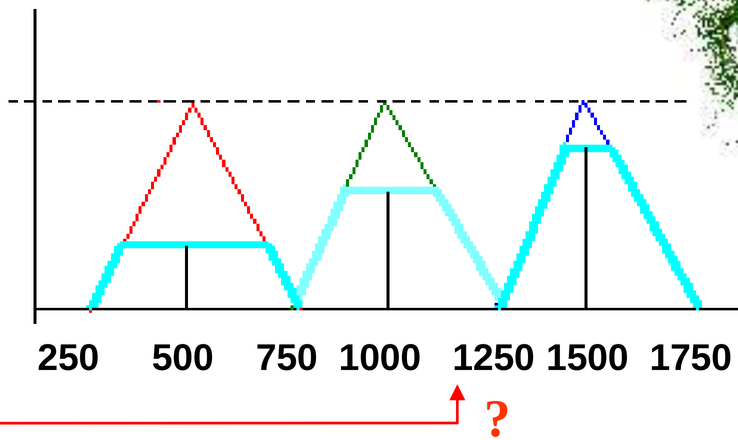
μ=abcissas

9.5.1 Inferência de Mamdani

O COA representa o Centro de Gravidade de uma figura geométrica:

- Neste caso pretendemos o Centro de Gravidade da figura delimitada pelo azul claro, que representa a conclusão das regras 1 a 4.

- Este centro de gravidade é uma abcissa *x*: Ela representa o valor do crédito a conceder



- Note-se que, como Alto possui maior massa (tem valores de μ mais altos) é de esperar que o COA obtido se situe acima do ponto médio (1000)

Como os trapézios são simétricos e o corte superior é plano , podemos calcular uma aproximação de COA designada por Weighted Average. Para cada trapézio tem-se

média(Baixo)=500 média(Médio)=1000 média(Alto)=1500

... e depois combinam-se os valores assim obtidos:

$$\text{Crédito} = \frac{500x0.3 + 1000x0.5 + 1500x0.7}{0.3 + 0.5 + 0.7} = 1333.33\,€$$

9.5.2 Regra Composicional de Inferência

9.5.2 Regra Composicional de Inferência (RCI - *Zadeh*)

A Regra Composicional de Inferência baseia-se em **Relações Difusas**

- Sejam as seguintes relações:

Relação Crespa "X é casado com Y"

Só utiliza 0's e 1's (verdadeiro / falso)

É casado com	Carla	Jacinta	Maria
João	1	0	0
Manuel	0	1	0
Luís	0	0	1

Relação Difusa "X é muito maior que Y"

Utiliza Graus de Pertença à Relação

X >> Y	1	2	3
3	0.2	0.05	0
20	1	0.9	0.85
89	1	1	1

9.5.2 Regra Composicional de Inferência

Uma relação difusa pode ser dada por uma tabela ou por uma função:

Exemplo

A seguinte função pode traduzir a relação *"x é bastante maior que y"*

0 se $x \leq y$

$\mu_{\sim R}(x,y) = (x-y)/10y$ se $y < x \leq 11y$

1 se $x > 11y$

A RCI permite efectuar inferências a partir de duas relações difusas, recorrendo à operação de **Composição de Relações**. Cada inferência equivale a um ***Modus Ponens Generalizado***:

Premissa: S_1 is Q_1

Implicação: $S_1 \sim R\ S_2$

Conclusão: S_2 is Q_2 em que ~R é uma relação difusa

9.5.2 Regra Composicional de Inferência

Exemplo
MPG (exemplo do Zimmerman):

SE o tomate está vermelho ENTÃO está maduro
O tomate está amarelado

O tomate está pouco maduro

Exemplo
Aplicação ao Diagnóstico:

SE o sintoma é {X} ENTÃO o diagnóstico é Y
O sintoma é semelhante a {X} num grau λ_1

O diagnóstico é Y num grau λ_2

Composição de Relações Difusas (Zadeh):

$$R_1 \circ R_2 = \left\{ \left[(x,z), \max_{y} \{\min \{\mu_{R1}(x,y), \mu_{R2}(y,z)\}\} \right] \mid x \in X, y \in Y, z \in Z \right\}$$

9.5.2 Regra Composicional de Inferência

Exemplo Seja o universo X={1,2,3,4}

Seja ~R1 = "*x é pequeno*" = {(1,1), (2,.6), (3,.2), (4,0)}

Seja ~R2 = "*x e y são aproximadamente iguais*"

A Composição de Relações obtém-se (quase) como uma multiplicação de matrizes (o número de colunas da primeira matriz tem de ser igual ao número de linhas da segunda) :

{(1,1), (2,.6), (3,.2), (4,0)} ∘

x \ y	**1**	**2**	**3**	**4**
1	1	.5	0	0
2	.5	1	.5	0
3	0	.5	1	.5
4	0	0	.5	1

=

=

1	**2**	**3**	**4**
Max(min(1,1),min(.6,.5),min(0.2,0),min(0,0)	max(min(1,.5),min(.6,1),min(.2,.5),min(0,0)	max(min(1,0),min(.6,.5),min(.2,1),min(0,.5)	max(min(1,0),min(.6,0),min(.2,.5),min(0,1)

=

	1	**2**	**3**	**4**
y	1	.6	.5	.2

A relação resultado traduz ***"y é aproximadamente pequeno"***. Note-se que para "3" e "4" o valor μ subiu devido à "dilatação" originada pelo termo "aproximadamente" que em ~R1 não existia.

9.6 Possibilidade e Lógica Difusa

9.6.1 Possibilidade e Necessidade

Possibilidade: uma forma de incerteza que exprime uma avaliação (eventualmente *) subjectiva acerca de certa questão

Exemplos

1. Uma pessoa é "nova" ou "velha" mas não se sabe exactamente a sua idade
2. Um forno está "quente" ou "frio" mas não se sabe a exactamente a sua temperatura
3. Num semáforo:
   - A probabilidade de verde, amarelo e vermelho é p.e. *5/10*, *1/10* e *4/10*
   - A possibilidade de verde, amarelo ou vermelho é *1* (para qualquer das cores) porque se sabe que é "completamente possível" que a cor visível seja uma destas

Dado um conjunto difuso, a possibilidade de um valor x é numericamente igual a $\mu(x)$

(*) No CADIAG-2 o autor defende que no domínio do diagnóstico possibilidade e probabilidade têm o mesmo valor. Normalmente possibilidade > probabilidade.

9.6.1 Possibilidade e Necessidade

Proposições da forma "x é $\widetilde{F}$", em que x é um **objecto**, implicitamente contêm uma referência a um **atributo** A do **objecto** X (i.e. são representáveis por **ternos OAV** em que **valor** é um *conjunto difuso* que representa um *termo linguístico*)

Exemplo

"a água está quente" é equivalente a "a temperatura da água é alta", em que "temperatura" é um **atributo** do **objecto** "água" e **alta** o valor desse atributo.

A Equação de Atribuição Relacional (**RAE**) de Zadeh formaliza este processo de atribuição, baseando-se numa relação difusa ~R:

$$\widetilde{R}(A(X)) = \widetilde{F}$$

A RAE induz uma distribuição de possibilidade cuja função de distribuição, $\pi_X(u)$, tem valores numericamente iguais aos da função de pertença que define o termo linguístico. Ou seja:

$$\pi_X \triangleq \mu_{\widetilde{F}}(u)$$

em que o símbolo " $\triangleq$ " significa "é definida como".

9.6.2 Variável Linguística Verdade

A **necessidade**, *N*, está intimamente relacionada com a possibilidade:

$$\pi(\tilde{A}) = 1 - N(\not\subset \tilde{A})$$

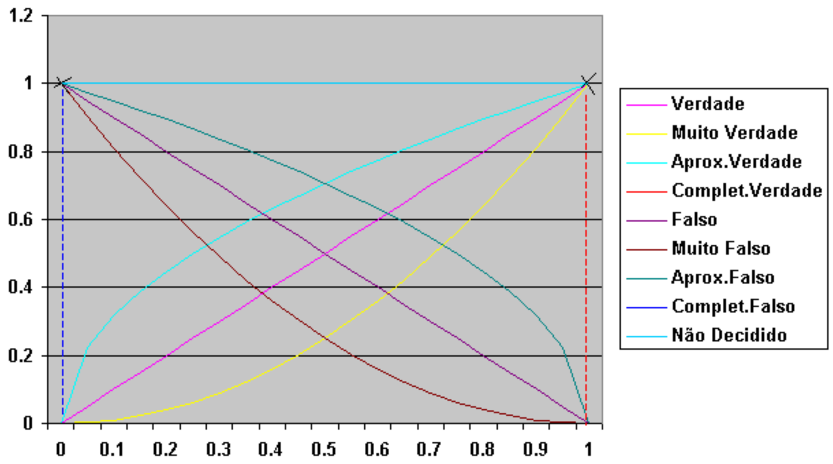
- Possibilidade e Necessidade são **Medidas Difusas**
- A necessidade é "mais exigente" e restritiva no sentido em que **mede quão necessariamente um facto é possível**, em função de outro conhecido.
- $N=1$ indica que um facto é "necessariamente verdadeiro"

9.6.2 Variável Linguística Verdade

De todas as variáveis linguísticas, uma assume particular importância: a variável **Verdade**

Admitindo graduações desta variável (*completamente falso, falso ... pouco verdadeiro, verdadeiro, completamente verdadeiro*) diferencia-se claramente a lógica clássica da lógica difusa: agora há graus de verdade expressos por termos linguísticos que são representáveis por conjuntos difusos.

9.7 Sistema Pericial CADIAG-2



Termos da variável linguística "verdade" segundo Baldwin

$$\mu_{\text{Muito Verdadeiro}} = \left(\mu_{\text{Verdadeiro}}(v)\right)^2$$

$$\mu_{\text{Aproximadamente Verdadeiro}} = \left(\mu_{\text{Verdadeiro}}(v)\right)^{1/2}$$

9.7 CADIAG-2

A teoria dos conjuntos difusos, possibilidade e composição de relações permitem tratar os quadros médicos, inerentemente inexactos, de forma adequada:

- Suportam aproximações linguísticas dos termos médicos
- Suportam métodos de raciocínio aproximado através da lógica difusa

O CADIAG-2 (*Computer Assisted Diagnosis*) é um Sistema Pericial de Diagnóstico Médico baseado em na Teoria da Possibilidade e na Composição de Relações Difusas.

9.7.2 CADIAG-2: Conhecimento

Fuzzy Set Theory in Medical Diagnosis, *Adlassnig, K.P.*
IEEE Transactions on Systems, Man and Cybernetics Vol.SMC-16, Nº2, March/April 1986

9.7.2 CADIAG-2: Representação do Conhecimento

Conhecimento médico e dados dos pacientes são representados por 8 relações difusas:

Conhecimento médico

- Relações sintoma+testes - doença $R(S_i , D_j)$, μ_o
- Relações sintoma+testes - doença $R(S_i , D_j)$, μ_c
- Relações comb.de sintomas+testes - doença $R(Sc_i , D_j)$, μ_o
- Relações comb.de sintomas+testes - doença $R(Sc_i , D_j)$, μ_c
- Relações sintoma - sintoma $R(S_i , S_j)$, μ_o
- Relações sintoma - sintoma $R(S_i , S_j)$, μ_c
- Relações doença - doença $R(D_i , D_j)$, μ_o
- Relações doença - doença $R(D_i , D_j)$, μ_c

Relações Difusas

Dados dos pacientes

- Relação Paciente - sintomas R_{PS}
- Relação Paciente - comb. de sintomas R_{PSC}

o = grau de ocorrência = probabilidade de sintoma suposto doença (*a priori*)
c = grau de confirmabilidade = probabilidade de doença suposto sintoma (*a posteriori*)

Probabilidades iniciais obtidas por análise de históricos e experiência médica.

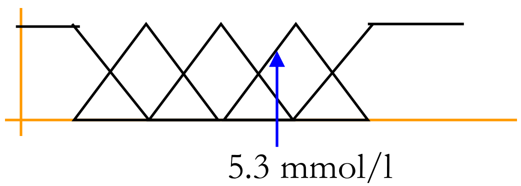
9.7.2 CADIAG-2: Conhecimento

Exemplo: Fuzzificação de Sintomas

Suponhamos que o paciente fez uma análise ao nível de potássio. O resultado é guardado no *Medical Information System* como valor crespo, por exemplo 5.3 mmol/l. O interpretador difuso, *Fuzzy Interpreter of Patient Data*, traduz este valor para os sintomas

potássio muito reduzido	$\mu_{Si} = 0.0$
potássio reduzido	$\mu_{Si} = 0.0$
potássio normal	$\mu_{Si} = 0.4$
potássio elevado	$\mu_{Si} = 0.6$
potássio muito elevado	$\mu_{Si} = 0.0$

Fuzzificação



Trata-se de uma fuzzificação: um valor crespo é representado por um valor μ de um termo de uma variável linguística.

- A relação assim obtida chama-se R_{PS} (de patient/symptom) e pertence ao produto cartesiano do conjunto de pacientes Π pelo conjunto de sintomas, Σ. Portanto, $R_{PS} \subset \Pi \times \Sigma$, em que $\mu_{Rps}(P_q,S_i)=\mu_{Si}$ para cada paciente P_q. O valor de $\mu \in [0,1]$ é determinado conforme exemplo acima.

- Caso um determinado teste não tenha sido efectuado, μ toma o valor *v* significando "desconhecido". Portanto, $\mu \in [0,1] \cup v$.

9.7.2 CADIAG-2: Conhecimento

Exemplo: Relação Paciente - Sintomas

Pacientes =

= Π = {João, Manuel} = {J,M}

Sintomas =

= Σ = {hipotermia, temp.normal, febre alta, febre muito alta, potássio reduzido, potássio elevado, potássio muito elevado}

$R_{PS} = \Pi \times \Sigma$

= {(J, hipotermia),0.0; (J, temp.normal),0.1; (J,febre alta),0.8; (J,febre muito alta),0.3; (J,potássio reduz.), 0.0; (J,potássio elevado), 0.5; (J,potássio mt.elevado), 0.0; (M, hipotermia),0.0; (J, temp.normal),1.0; (J,febre alta),0.0; (J,febre muito alta),0.0; (M,potássio reduz.), 0.0; (M,potássio elevado), 0.5; (M,potássio mt.elevado), 0.0 }

(os termos com grau de pertença 0 não são, na prática, representados)

O CADIAG utiliza ainda relações R_{PSC} de Pacientes / Combinações de Sintomas que provêm dos sintomas actuais, análises e história clínica:

dores nas costas $\wedge$ limitação do movimento de coluna $\wedge$ ... $\wedge$ sexo masculino $\wedge$ idade entre 20 e 40 anos

9.7.3 CADIAG-2: Inferência

9.7.3 CADIAG-2: Inferência

O objectivo (*goal*) último do sistema é concluir qual o grau de pertença de cada diagnóstico em consideração, a cada paciente.

Esta "associação" entre pacientes e diagnósticos também é traduzida por uma relação difusa R_{PD}, (de *patient/diagnostic*) definida em $R_{PD} \subset \Pi \times \Delta$, sendo Δ o conjunto dos diagnósticos possíveis, D_j. Na relação R_{PD}:

- μ_{Rpd} representa o grau de pertença do diagnóstico D_j ao paciente P_q
- Se um diagnóstico tem μ=1.0, está confirmado
- Se tem μ=0.0 está excluído
- Entre 0 e 1, μ traduz o grau de possibilidade de cada diagnóstico
- Se μ=*v*, isso indica que este diagnóstico não foi considerado

Exemplo: Relação Paciente - Diagnóstico

R_{PD} = {(J, gripe), 0.8; (J, amigdalite), 0.3; (J, sinusite), 0.2; (M, gripe), 0.0; (J, amigdalite),0.5; (J, sinusite), 0.1}

Os valores de μ exprimem quanto o CADIAG concluiu que os sintomas do João e do Manuel sugerem que se trata de gripe, amigdalite ou sinusite.

9.7.3 CADIAG-2: Inferência

As regras são do tipo

IF antecedent THEN consequent WITH (o,c)

- *o* = frequência de ocorrência: traduz a frequência com que o sintoma X ocorre quando a doença Y está presente **causas → sintomas**
- *c* = grau de confirmabilidade: traduz o grau em que a presença do sintoma X significa a presença da doença Y **sintomas → causas**

o e ***c*** são pois aproximações (representadas por frequências relativas) de probabilidades *a priori* e *a posteriori*

ocorrência $= p(Y|X) = p(Y \cap X)/p(X)$ *a priori*
confirmabilidade $= p(X|Y) = p(Y \cap X)/p(Y)$ *a posteriori*

Estas aproximações são adquiridas:

1. Por avaliação numérica ou linguística dada por peritos médicos
2. Por avaliação estatística de uma base de dados contendo os dados dos pacientes com diagnósticos já confirmados

9.7.3 CADIAG-2: Inferência

O CADIAG utiliza pois Relações Difusas e sua composição para inferir baseado em factores de base probabilística. Isto é possível ?

De facto, probabilidade e possibilidade são diferentes. Veja-se por exemplo que:

- $p(A \cap B)=p(A|B)/p(B)$ para probabilidade
- $\mu(A \cap B)=\min(\mu_A,\mu_B)$ para possibilidade

É possível porque o autor do CADIAG-2 admite que

"As relações do CADIAG-2 têm a importante propriedade de poderem ser interpretadas probabilisticamente. Os valores da **frequência de ocorrência** μ_o e do **grau de confirmabilidade** μ_c são definidos da seguinte forma:

$$\mu_o = \frac{F(S_i \cap D_j)}{F(D_j)} \qquad \mu_c = \frac{F(S_i \cap D_j)}{F(S_i)}$$

em que F representa frequência absoluta. Com estas definições, avaliações estatísticas das relações médicas já conhecidas ou ainda desconhecidas pode ser levada a cabo usando os dados dos pacientes com diagnósticos confirmados"

9.7.3 CADIAG-2: Inferência

A partir das relações Paciente/Sintomas e Paciente/Combinações de Sintomas e das 8 relações difusas atrás apresentadas e representativas do conhecimento médico, o CADIAG-2 realiza 12 operações de composição:

- Estas composições são realizadas pela RCI, cuja aplicação "equivale" a vários *modus ponnens generalizados (MPG)*:

Premissa:	*x* é $\tilde{A}'$
Implicação:	Se *x* é $\tilde{A}$, então *y* é $\tilde{B}$
Conclusão:	*y* é $\tilde{B}$ (no grau *z*)

Composição de Relações Difusas segundo Zadeh (repetição):

$$R_1 \circ R_2 = \left\{ \left[(x,z), \max_y \{\min \{\mu_{R1}(x,y), \mu_{R2}(y,z)\}\} \right] \mid x \in X, y \in Y, z \in Z \right\}$$

As 12 composições de relações descrevem-se resumidamente em seguida:

9.7.3 CADIAG-2: Inferência

Composições 1,2 e 3: **Sintomas / Diagnósticos**

Com estes sintomas **deve ser** a doença x

1. Composição S_iD_j (**hipótese** pelos **sintomas presentes**)

Obtém-se compondo a relação doentes/sintomas com a relação sintomas/doenças usando $\mu_{confirmabilidade}$ (daí o ***c*** **que figura como sobrescrito de** R_{SD})

$$R^1_{PD} = R_{PS} \circ R^c_{SD}$$

$$\mu_{R^1_{PD}}(P_q, D_j) = \max_{S_i} \min\left[\mu_{R_{PS}}(P_q, S_i); \mu_{R^c_{SD}}(S_i, D_j)\right]$$

Com estes sintomas **não** **deve ser** a doença x

2. Composição S_iD_j (**exclusão** pelos **sintomas presentes**)

$$R^2_{PD} = R_{PS} \circ \left(1 - R^c_{SD}\right)$$

$$\mu_{R^2_{PD}}(P_q, D_j) = \max_{S_i} \min\left[\mu_{R_{PS}}(P_q, S_i); 1 - \mu_{R^c_{SD}}(S_i, D_j)\right]$$

Porquê 1-μ ?

Porquê μ_c (e não μ_o) ? Na composição das relações R1 e R2 é usado μ_c e não μ_o uma vez que se está a compor **doentes/sintomas** com **sintomas/doenças** o que significa que se trata de um possibilidade *a posteriori*: ela representa a hipótese ou exclusão de diagnósticos a partir dos sintomas observados e esse grau de confirmação é representado por μ_c

9.7.3 CADIAG-2: Inferência

Porquê $1-\mu_c$ na composição de R2 ?

1. O grau de verdade de *not A* é dado por

$$v(\text{not "}\mathit{u\ is\ A}\text{"}) = v(\neg A) = \neg v(A) = 1\text{-}v(A)$$

2. A definição clássica da implicação é

$$(p \Rightarrow q) = (\neg p \vee q)$$

3. Relação R2 = Exclusão de diagnósticos pelos sintomas presentes: inferir quais os diagnósticos que não são possíveis devido aos sintomas que se observam. Portanto, pretende-se calcular o grau de verdade de

$$\mathbf{v(A) \Rightarrow \neg v(B)} =$$

Segundo 2.

$$= \neg v(A) \vee \neg v(B) =$$

Segundo 1.

$$= \neg v(A) \vee v(\neg B) =$$

Segundo 2.

$$= v(A) \Rightarrow v(\neg B) =$$

Segundo 1.

$$= \mathbf{v(A) \Rightarrow 1\text{-}v(B)} =$$

$$= v(\text{"u is A"}) \Rightarrow v(\text{"z is not B"})$$

9.7.3 CADIAG-2: Inferência

Se fosse a doença x **deveriam existir** certos sintomas ...

3. Composição S_iD_j (**exclusão** pelos **sintomas ausentes**)

Obtém-se compondo a relação doentes/sintomas com a relação sintomas/doenças usando $\boldsymbol{\mu}_{\mathbf{ocorrência}}$ (daí o ***o*** **que figura como sobrescrito de $\mathbf{R_{SD}}$)**

$$R^3_{PD} = (1 - R_{PS}) \circ R^o_{SD}$$

$$\mu_{R^3_{PD}}(P_q, D_j) = \max_{S_i} \min \left[1 - \mu_{R_{PS}}(P_q, S_i); \mu_{R^o_{SD}}(S_i, D_j) \right]$$

R3 resulta de doentes-sintomas e sintomas-doenças mas usando $1\text{-}\mu_{PS}$ e μ_o.

Porquê $\boldsymbol{\mu_o}$? Pretendem excluir-se diagnósticos que não são possíveis porque há sintomas que não se observam. Portanto, pretende calcular-se o grau de verdade de

$$\neg v(A) \Rightarrow \neg v(B) =$$

$$= \neg v(\neg A) \vee \neg v(B) \;=\; v(A) \vee v(\neg B) \;=\; v(\neg B) \vee v(A) \;=\; v(B) \Rightarrow v(A)$$

Em suma: como $(\neg v(A) \Rightarrow \neg v(B)) = (v(B) \Rightarrow v(A))$ deve usar-se $\boldsymbol{\mu_o}$ porque é ele que expressa as ligações no sentido doença → sintomas (i.e. B→A).

Porquê $\mathbf{1\text{-}\boldsymbol{\mu}_{PS}}$? Em que grau é que um sintoma está ausente ? Está presente no grau μ_{PS}, logo está ausente no grau $1\text{-}\mu_{PS}$ ... e portanto usa-se $1\text{-}\mu_{PS}$ para expressar um grau de ausência em vez de presença.

9.7.3 CADIAG-2: Inferência

A composição de relações 4 a 12 baseia-se em processos idênticos aos analisados:

- 4, 5 e 6 - é definida de forma idêntica mas a partir de R_{PqSCl} e de R_{SCiDj} (combinações de sintomas / diagnósticos)
- 7, 8 e 9 é definida de forma idêntica mas a partir de R_{PqSl} e de R_{SCiSCj} (sintomas / sintomas). Destina-se a completar o quadro (*pattern*) de sintomas de cada paciente
- A composição de relações 10, 11 e 12 é definida de forma idêntica mas a partir de R_{PqDl} e de R_{DiDj} (diagnósticos / diagnósticos) (tem a ver com o processo de diagnóstico diferencial)

Diagnóstico confirmado:

- μ_{PD} obtido pela inferência 1 ou 4, tem o valor 1 ou ...
- μ_{PD} obtido pela inferência 10 tem o valor 1

Diagnóstico excluído:

- μ_{PD} obtido pelas inferências 2 ou 3 ou 5 ou 6 tem o valor 1 ou...
- μ_{PD} obtido pelas inferências 11 ou 12 tem o valor 1

Para mais pormenores ver bibliografia indicada no início da secção 6.7

9.8 Match no Fuzzy CLIPS

9.8 *Fuzzy CLIPS*

Uma versão difusa do CLIPS chamada ***Fuzzy CLIPS*** foi desenvolvida no NRC - *National Research Center* - Canadá (*freeware* para fins académicos)

- No Fuzzy CLIPS uma regra pode lidar com valores crespos, difusos e CFs

Porém como determinar o *match* entre um facto e a premissa de uma regra quando ambos são representados por conjuntos difusos ?

Exemplo

SE "rendimento baixo" ENTÃO "crédito baixo"

Qual o crédito a conceder se o indivíduo tiver um "rendimento médio"?

Há que calcular uma medida da semelhança entre baixo e médio !!!

No Fuzzy CLIPS esta semelhança é determinada com base no trabalho de Cayrol, Farenni e Prade, baseando-se nas medidas difusas Possibilidade e Necessidade, como se segue:

9.8 Match no Fuzzy CLIPS

As expressões base são as seguintes:

Se $N(F\beta \mid F\alpha) > 0.5$	$S=P(F\beta \mid F\alpha)$	(1)
Senão	$S=(N(F\beta \mid F\alpha)+0.5)*P(F\beta \mid F\alpha)$	(2)

com S= Semelhança entre $F\alpha$ e $F\alpha$, P= Possibilidade, N= Necessariedade

e P e N são definidas por

$$P(F\beta \mid F\alpha) = \max(\min(\mu F\beta(u), \mu F\alpha(u)) \quad \forall u \in U \qquad (3)$$

$$N(F\beta \mid F\alpha) = 1-P(F\not\subset\beta \mid F\alpha) \qquad (4)$$

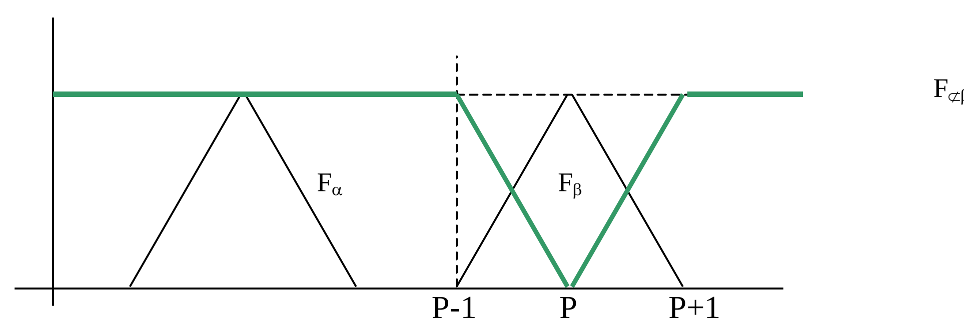
$$\mu\not\subset F\alpha = 1- \mu F\alpha \qquad (5)$$

Exemplos

Caso 1

Os conjuntos α e β não se sobrepõem

É o caso de se querer comparar p.e. "água quente" (facto) com "água fria" (premissa de uma regra)



9.8 Match no Fuzzy CLIPS

De (4) temos que $N(F_\beta|F_\alpha) = 1\text{-}P(F_{\not\subset\beta}|F_\alpha)$
Na figura representou-se a verde $F_{\not\subset\beta}$ atendendo a que através de (5), $\mu_{\not\subset\beta} = 1\text{-}\ \mu_{F\beta}$.

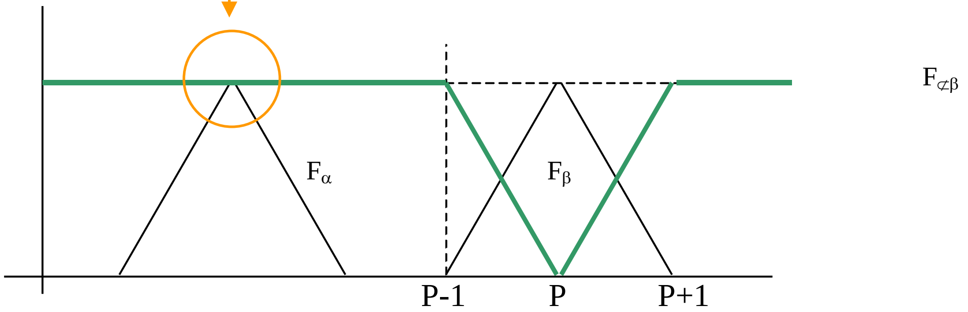
De (3) temos que $P(F_\beta|F_\alpha) = \max(\min(\mu_{F\beta}(u),\mu_{F\alpha}(u))$
que aplicado ao caso presente se transforma em $P(F_{\not\subset\beta}|F_\alpha) = \max(\min(\mu_{F\not\subset\beta}(u),\mu_{F\alpha}(u))$.
Este máximo encontra-se assinalado na figura e tem o valor 1.
Logo, $P(F_{\not\subset\beta}|F_\alpha)=1$.
Donde, de (4), $N(F_\beta|F_\alpha)=1\text{-}1=0$.

Como N=0 é <0.5, a semelhança entre F_α e F_β é dada por (2):
$S=(N(F_\beta|F_\alpha)+0.5)*P(F_\beta|F_\alpha)$
Segundo (3), $P(F_\beta|F_\alpha) = \max(\min(\mu_{F\beta}(u),\mu_{F\alpha}(u))$. Mas como F_α e F_β não têm pontos comuns, $P(F_\beta|F_\alpha)=0$

Então S=0 (conforme esperado) por se tratar de um produto.

Caso 1

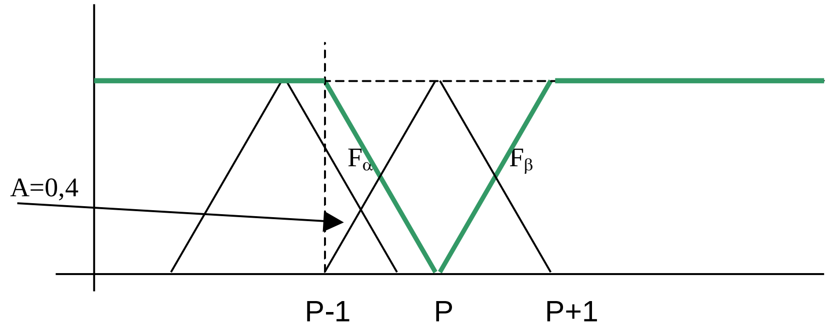
Os conjuntos α e β não se sobrepõem



9.8 Match no Fuzzy CLIPS

Caso 2

Os conjuntos α e β têm bases que se sobrepõem em menos de 50%



De (3) $P(F_{\not\subset\beta}|F_\alpha) = \max(\min(\mu_{F\not\subset\beta}(u),\mu_{F\alpha}(u))$ e vemos que agora os *min* continuam a seguir μ de $F_{\not\subset\beta}$ donde o *max* é 1 e portanto $N(F_\beta|F_\alpha)=0$.

Porém, agora α e β cruzam-se, donde, supondo que o cruzamento se dá em A=0,4, resulta $P(F_\beta|F_\alpha) = \max(\min(\mu_{F\beta}(u),\mu_{F\alpha}(u)) = 0.4$

Como N<0.5, a fórmula a aplicar é a mesma do exemplo anterior:
$S=(N(F_\beta|F_\alpha)+0.5)*P(F_\beta|F_\alpha)$

Donde S = (0+0,5)* 0,4 = 0.2

Para mais pormenores sobre o Fuzzy CLIPS, em particular os métodos de combinação de incerteza (CFs e valores difusos) ver o User's Guide