

Lógica Fuzzy e Diagnóstico de Processo Inflamatório a Partir de Medidas de Proteínas: uma Aplicação à Parvovirose Canina

Luciana T. Gomes¹,

Laécio C. Barros²

DMA, IMECC – UNICAMP, 13.083-859, Campinas/SP.

Resumo. Neste trabalho é mostrado um aspecto de como, a partir de dados laboratoriais, é possível utilizar a teoria dos conjuntos fuzzy para a obtenção do diagnóstico de um processo inflamatório, mais especificamente da gastrenterite hemorrágica. É gerado e implementado um algoritmo no software Mathematica e, como aplicação, realiza-se o diagnóstico da parvovirose canina, o qual se dá a partir de dados laboratoriais introduzidos no programa, suficientes para a identificação do processo inflamatório citado.

Palavras-chave: *Conjuntos Fuzzy, Diagnóstico Médico, Proteínas de Fase Aguda, Gastrenterite Hemorrágica, Parvovirose Canina.*

1. Introdução

O diagnóstico médico é uma decisão tomada por especialistas, baseando-se na caracterização dos sintomas do paciente e comparando-a aos sintomas da possível doença. Essa caracterização de sintomas é um procedimento que

¹lucianatakata@gmail.com

²laeciocb@ime.unicamp.br

retorna respostas muitas vezes abstratas, relativas e inexatas (tais como febre alta, muita dor, tosse freqüente ou baixa taxa de determinada proteína no sangue).

No nosso dia-a-dia, estamos constantemente lidando com esse tipo de informação. Recebemos diferentes dados incertos e, a partir de alguma análise realizada em nosso cérebro, se necessário retornamos uma resposta, a qual também pode ser incerta, abstrata e relativa.

A necessidade de transformar esse tipo de informação para a linguagem matemática motivou o desenvolvimento da teoria dos conjuntos fuzzy. Através dela, é possível lidar matematicamente com conceitos abstratos de maneira muito mais eficiente do que utilizando-se apenas a matemática clássica.

É o caso do diagnóstico médico. Os especialistas recebem como informação do paciente a caracterização de sinais, através de adjetivos ditos pelos próprios pacientes, ou o resultado de um exame, com imagens ou números a serem interpretados.

Os sintomas podem ser relativos, recebendo caracterizações diferentes de acordo com o sexo, a altura, o peso, a idade ou outros fatores próprios de cada paciente; podem ser incertos, devido à não exatidão do exame laboratorial, inerente do procedimento de medição de dados; e são abstratos, uma vez que a caracterização final de um sintoma é dado por um adjetivo, o qual pode ser interpretado de maneiras diferentes, de acordo com cada mente.

Portanto, a teoria de conjuntos fuzzy se mostra uma ferramenta eficiente para se lidar com esta situação.

2. Objetivos

- Utilizar a teoria dos conjuntos fuzzy para a obtenção de diagnóstico da gastroenterite hemorrágica a partir de dados laboratoriais;
- Construir e implementar no software Mathematica um algoritmo que receba como entrada os dados de um paciente e forneça um grau de

possibilidade da doença;

- Avaliar os resultados e concluir se o uso da teoria dos conjuntos fuzzy se mostra adequado.

3. Metodologia

3.1 A parvovirose canina, a GEH e as PFA

A parvovirose canina é uma doença altamente contagiosa e letal causada pelo parvovírus canino e que ocorre principalmente em cães jovens não vacinados. A sua transmissão se dá através do contato via oral com fezes de outro animal infectado. A replicação do vírus no estágio avançado da contaminação causa a necrose das criptas do epitélio do intestino delgado, com eventual destruição das vilosidades, e episódios gastroentéricos severos, ocasionando muitas vezes hemorragia. Outros sintomas são apatia, hipertermia, vômitos, desidratação, leucopenia, esgotamento linfóide e, possivelmente, conjuntivite e inflamação no coração.

Como a gastroenterite hemorrágica (GEH) está freqüentemente presente no quadro da parvovirose canina, a identificação precoce deste sintoma pode auxiliar no tratamento da doença de maneira mais eficaz.

Para tanto, podem ser utilizadas as chamadas PFA (proteína de fase aguda), proteínas cujo processo de produção é acelerado em resposta à injúria tecidual. De acordo com Delayte et al. (2003), três dessas proteínas relacionadas à GEH são a haptoglobina, a ceruloplasmina e a α -glicoproteína ácida. Como a concentração desses elementos se eleva no sangue em caso de GEH, podem ser utilizados para se diferenciar cães saudáveis de cães sob o processo inflamatório.

Neste trabalho são utilizados dados das PFA de cães saudáveis e de cães com sinais clínicos de gastroenterite hemorrágica (GEH) aguda e leucopenia (e, portanto, presuntivamente diagnosticados com parvovirose canina).

3.2 Teoria

Um conjunto clássico A pode ser representado pela sua função característica. Ou seja, dados um conjunto U e um subconjunto A de U , A pode ser representado como:

$$\chi_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases} \quad (3.1)$$

Baseando-se neste tipo de caracterização se define um subconjunto fuzzy:

Definição 1 *Seja U um conjunto clássico. Um subconjunto fuzzy F de U é caracterizado por uma função*

$$F : U \rightarrow [0, 1] \quad (3.2)$$

pré-fixada, chamada de função de pertinência de F .

Em analogia à representação dos conjuntos clássicos apresentada, se x é um elemento de U e $F(x) = 1$ dizemos que o elemento pertence completamente ao conjunto F , enquanto que se $F(x) = 0$, sua pertinência é nula, não correspondendo a este conjunto.

Dizemos que uma variável lingüística X no universo U é uma variável cujos valores assumidos por ela são subconjuntos fuzzy de U . Por exemplo, se X é quantidade de haptoglobina, um valor que pode ser assumido por esta variável é “alta”. Se abaixo do valor 1.6 a haptoglobina não é considerada nada alta, mas a partir de 2.2 é certamente alta, o subconjunto fuzzy F haptoglobina alta pode ser dado, por exemplo, por:

$$F(x) = \begin{cases} 0 & \text{se } x \leq 1.6 \\ \frac{x-1.6}{0.6} & \text{se } 1.6 < x \leq 2.2 \\ 1 & \text{se } x > 2.2 \end{cases} \quad (3.3)$$

Ou seja, entre $x = 1.6$ e $x = 2.2$, à medida que aumentam os valores da haptoglobina se tornam cada vez mais pertinentes ao subconjunto “haptoglobina alta”, até serem considerados completamente altos (pertinência 1).

3.2.1 Aplicação dos conjuntos fuzzy

A seguir introduzimos a notação a ser utilizada neste trabalho:

- d : dado coletado do paciente (número);
- D : dado coletado “fuzzyficado” (subconjunto fuzzy);
- Δ : diagnóstico;
- S_i : fator que caracteriza a doença (variável lingüística);
- F_i : valor do sintoma (subconjunto fuzzy).

Como estamos lidando com exames laboratoriais, ao se medir uma substância temos a inexatidão inerente ao exame. É por isso que o dado d do paciente é inicialmente considerado um número e depois é “fuzzyficado”, ou seja, é transformado em um subconjunto fuzzy:

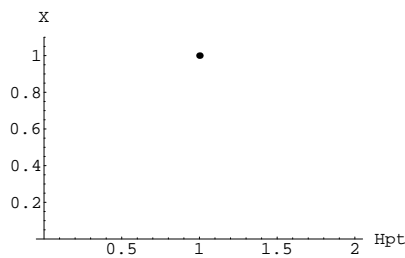


Figura 1: Gráfico da função característica do subconjunto clássico do dado do paciente (com $d = 1$).

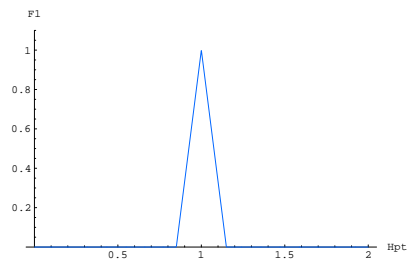


Figura 2: Gráfico da função de pertinência do subconjunto fuzzy do dado do paciente (com $d = 1$).

Portanto, $Hpt = 1$ é o valor que possui pertinência 1 ao subconjunto fuzzy “dado do paciente”. Seus valores vizinhos também possuem uma alta pertinência a este conjunto, e vão diminuindo à medida que vão se afastando.

Também temos a inexatidão devido à caracterização do sintoma. É por isso que são utilizados subconjuntos fuzzy, ao invés de conjuntos clássicos,

que atribuiriam uma forte descontinuidade ao problema. Consideremos, por exemplo, que a partir de 2.0 a haptoglobina seja alta e antes disso não. Teríamos o gráfico da figura 3.

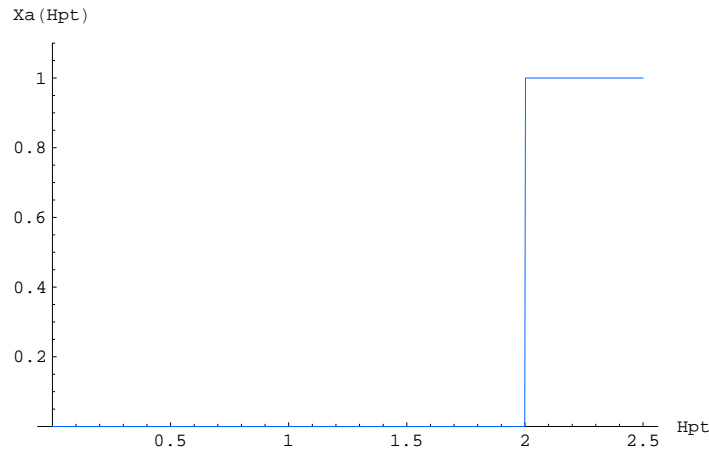


Figura 3: Gráfico da função característica do subconjunto clássico “Haptoglobina Alta”

Essa descontinuidade poderia ser fonte de muitos falsos positivos e falsos negativos, uma vez que o limite entre “normal” e alto não é um número exato. Há uma faixa de transição, como fornece o gráfico da figura 4.

S_i é o fator que caracteriza a doença, ou seja, a quantidade de uma PFA. Neste trabalho S_1 corresponde à concentração de haptoglobina no sangue do animal; S_2 corresponde à concentração de ceruloplasmina e S_3 corresponde à concentração de α -glicoproteína ácida.

F_i é o valor deste fator, isto é, a caracterização do fator para que este seja considerado um sintoma. Assim, F_1 corresponde à concentração alta de haptoglobina; F_2 à concentração alta de ceruloplasmina e F_3 à concentração alta de α -glicoproteína ácida.

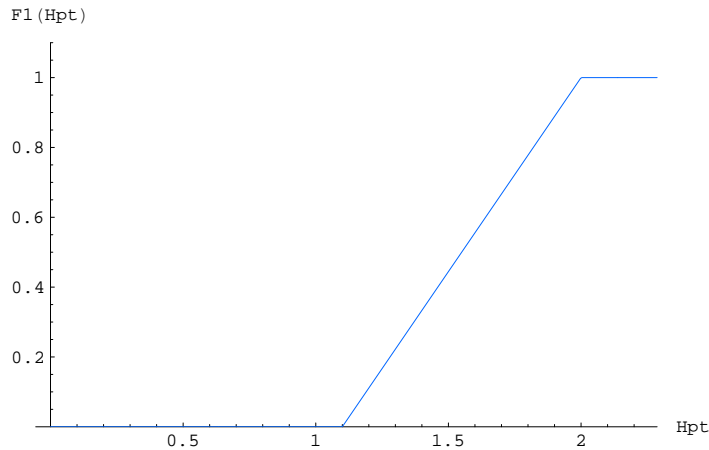


Figura 4: Gráfico da função de pertinência do subconjunto fuzzy “Haptoglobina Alta”

3.2.2 O modelo

De acordo com o modelo de Sanchez Sanchez (1998), a possibilidade P de um paciente p ter a doença Δ é dado por:

$$P\{S_1(p) \text{ ser } D_1 \text{ e } S_2(p) \text{ ser } D_2 \text{ e } S_3(p) \text{ ser } D_3, \text{ dado que } (S_1 \text{ é } F_1) \text{ e } (S_2 \text{ é } F_2) \text{ e } (S_3 \text{ é } F_3)\} = \min\{\Pi(F_1, D_1), \Pi(F_2, D_2), \Pi(F_3, D_3)\}$$

em que

$$\Pi(F_i, D_i) = \sup_x \{F_i(x) \cap D_i(x)\}$$

Esta fórmula retorna um número entre 0 e 1 que corresponde à possibilidade de os fatores caracterizadores da doença (as três PFA) de um paciente terem os valores D_i e F_i ao mesmo tempo. Ou seja, é a possibilidade das concentrações das PFA do paciente serem altas.

Para entender melhor a fórmula, apresentaremos alguns gráficos.

Mas antes é necessário lembrar que, como queremos a ocorrência simultânea de cada PFA ter os valores D_i e F_i , ou, mais especificamente, dois

subconjuntos fuzzy ao mesmo tempo, calcula-se o mínimo em cada ponto. Ou seja, dados os dois conjuntos F_i e D_i , para cada elemento $x \in U$, calcula-se o mínimo da pertinência entre os dois subconjuntos, obtendo-se um terceiro subconjunto fuzzy dado por:

$$(F_i \cap D_i)(x) = \min\{F_i(x), D_i(x)\}$$

Portanto, tendo em mãos os dados de um paciente (função de pertinência triangular do gráfico 5) e os valores para a PFA alta (função de pertinência não triangular do gráfico 5) realizamos a interseção dos conjuntos, obtendo o gráfico 6.

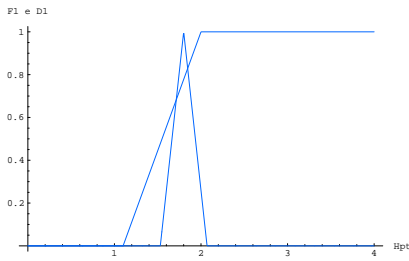


Figura 5: Gráfico das funções de pertinência dos subconjuntos fuzzy “Haptoglobina Alta” e “Dados do paciente”.

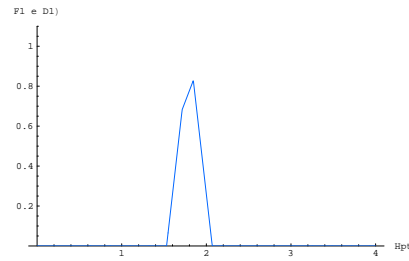


Figura 6: Gráfico da função de pertinência da interseção dos subconjuntos “Haptoglobina Alta” e “Dados do paciente”.

Realiza-se o mesmo procedimento para cada proteína.

Nota-se que quanto mais a i -ésima PFA for baixa, menor será a área resultante da interseção. Isso significa que o dado do paciente não se encaixa no sintoma.

Como a saída é dada por um número e não por conjuntos, para representar cada gráfico é escolhida a pertinência máxima de cada interseção. Como para o diagnóstico positivo é necessário ter os três sintomas ao mesmo tempo, é preciso que $\sup\{F_i \cap D_i\}$ para cada i seja alto. Portanto, escolhe-se o mínimo entre esses três números. Caso algum seja muito baixo (o que

significa que o paciente não tem algum sintoma que é necessário para o diagnóstico), significa que a possibilidade de a doença ocorrer é muito baixa.

3.3 Aplicação

Em seu artigo, Sanchez (1998) ilustrou a aplicação de seu modelo através do diagnóstico de síndromes protéico-inflamatórias a partir de cinco marcadores (substâncias do organismo que são afetadas em quantidade devido à doença).

Neste trabalho foram utilizados, como marcadores, três proteínas de fase aguda: haptoglobina, ceruloplasmina e α -glicoproteína ácida. Essas PFA são características da gastroenterite hemorrágica (GEH), sintoma intimamente ligado à parvovirose canina.

Através dos dados encontrados em Delayte et al. (2003), estabeleceram-se os conceitos de haptoglobina alta, ceruloplasmina alta e α -glicoproteína ácida alta, criando-se os seguintes subconjuntos fuzzy:

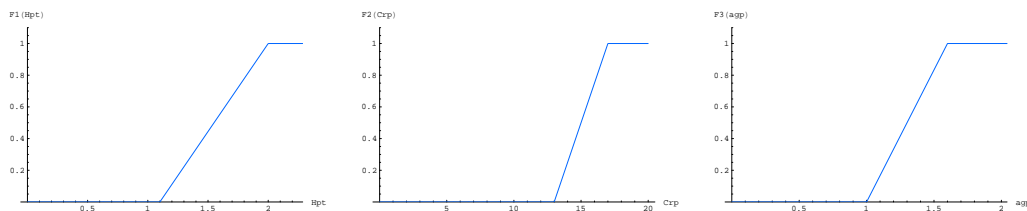


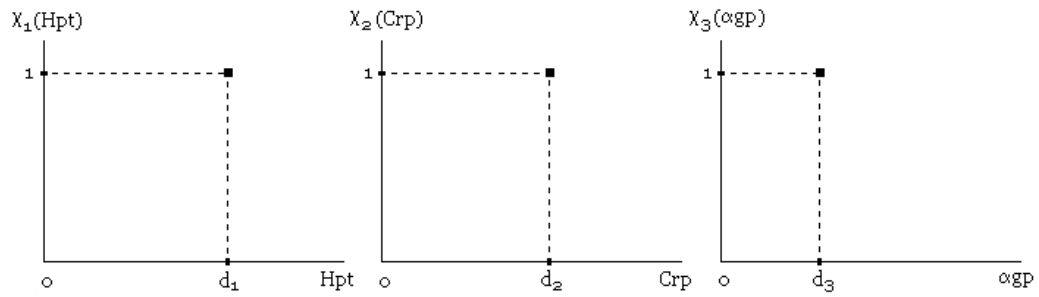
Figura 7: Gráfico da função de pertinência do subconjunto “Haptoglobina Alta”
 Figura 8: Gráfico da função de pertinência do subconjunto “Ceruloplasmina Alta”
 Figura 9: Gráfico da função de pertinência do subconjunto “ α -glicoproteína ácida Alta”

3.4 A rotina

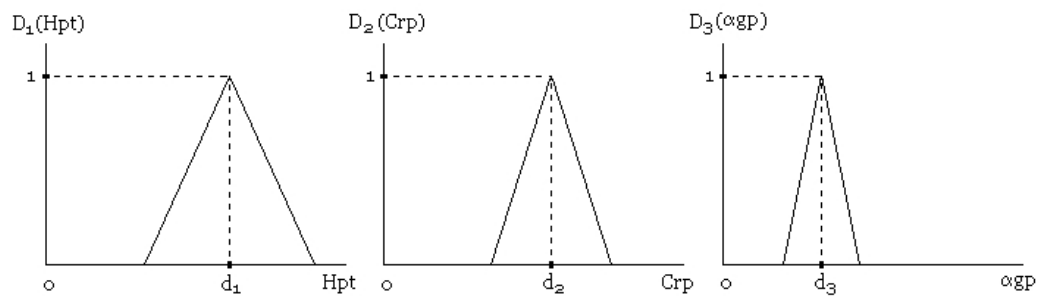
A rotina utilizada para o cálculo do diagnóstico é a seguinte:

Para cada cão:

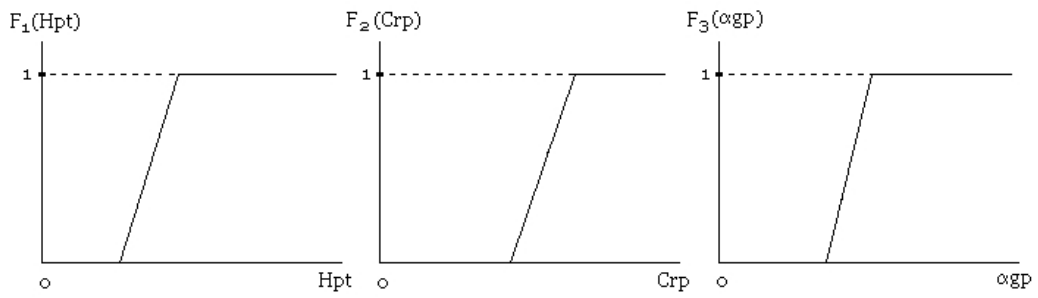
1. São inseridos d_1 , d_2 e d_3 (valores das PFA medidas)



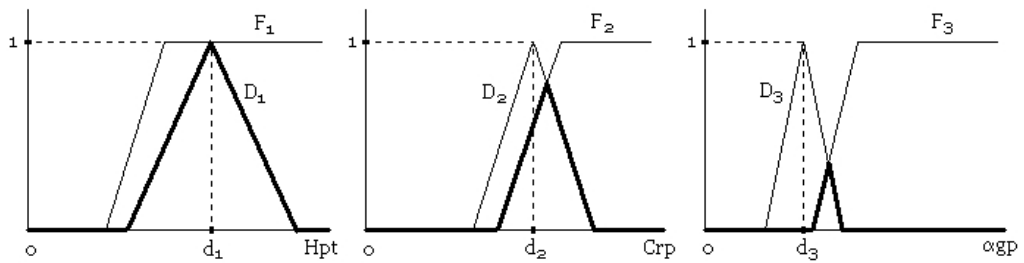
2. Definem-se as funções de pertinência dos subconjuntos fuzzy D_1 , D_2 e D_3



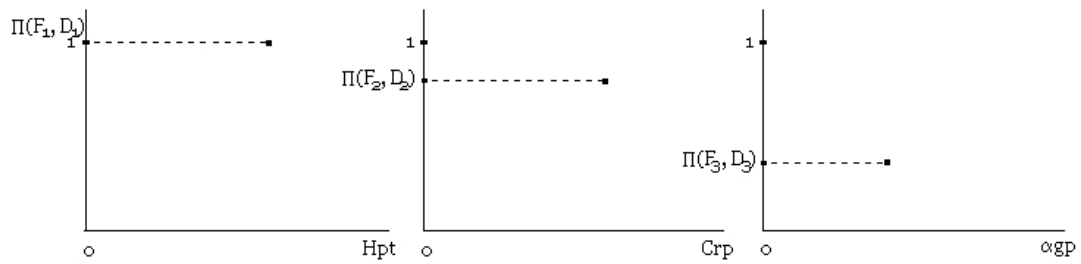
3. Definem-se as funções de pertinência dos subconjuntos fuzzy F_1 , F_2 e F_3



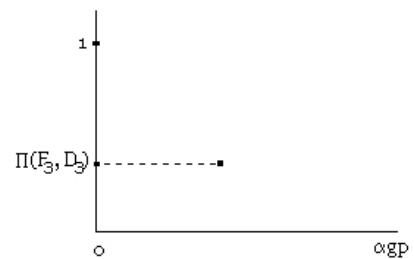
4. São calculados $F_1 \cap D_1$, $F_2 \cap D_2$, $F_3 \cap D_3$



5. São calculados $\Pi(F_1, D_1)$, $\Pi(F_2, D_2)$ e $\Pi(F_3, D_3)$



6. É calculado $\min\{\Pi(F_1, D_1), \Pi(F_2, D_2), \Pi(F_3, D_3)\}$



O valor final corresponde à possibilidade do cão ter a doença.

Poderemos ver posteriormente na sessão de resultados, porém, que algumas possibilidades obtidas não correspondem ao esperado (ver tabela 5). Isso deve-se ao fato de ainda não termos empregado o peso relativo, a ser introduzido a seguir.

3.5 O peso relativo

Como dito anteriormente, as PFA são proteínas que aumentam em quantidade como resposta à injúria tecidual. A parvovirose canina é uma doença que pode apresentar um grande número de sintomas e debilitar muito o animal, deixando-o mais suscetível a infecções, bacterianas ou virais; desequilíbrios no organismo; e outros fatores. Assim, é possível surgirem injúrias que podem estimular o aumento dessas mesmas PFA.

Em Delayte et al. (2003), há dados das PFA de 11 cães saudáveis e 11 cães com parvovirose. Os valores encontrados estão contidos nos intervalos apresentados na tabela abaixo:

Tabela 1: Faixas de valores dos dados laboratoriais.

Cães	Hpt	Crp	α gp
Doentes	2,10 a 2,70	11,8 a 23,1	1,05 a 3,33
Sadios	0,01 a 1,32	8,0 a 16,0	0,54 a 1,39

Nota-se que a quantidade de haptoglobina nos cães doentes se situa em um intervalo completamente diferente do dos sadios, ou seja, não há possíveis valores em comum. Já para a ceruloplasmina e a α -glicoproteína, o quadro é diferente: apesar de ser necessária uma quantidade alta da proteína para o diagnóstico positivo da doença, há cães sadios com quantidades maiores do que alguns cães doentes. Ou seja, os intervalos possuem valores em comum.

Possíveis causas para uma quantidade relativamente alta de PFA em cães sem GEH são outros processos inflamatórios. Também existe a possibilidade de cães doentes com ceruloplasmina ou α -glicoproteína com valores semelhantes ao de sadios. Isso pode ocorrer porque há problemas (como desnutrição, por exemplo) que podem diminuir a quantidade de certas PFA no sangue, ou ainda, diferentes fases do processo inflamatório podem causar variações nesses valores.

Como a haptoglobina se mostra bem marcante e determinante do pro-

cessos de diagnóstico, recebeu peso relativo 1. Para as demais proteínas foram escolhidos pesos menores, como mostra a tabela a seguir:

Tabela 2: Pesos relativos β de cada PFA no diagnóstico.

Hpt	Crp	agp
1	0,4	0,4

em que $\beta_i = 0$ significa que, entre as três proteínas, a i -ésima é completamente insignificante, e $\beta_i = 1$ atribui importância máxima ao marcador em questão.

Foram definidos os subconjuntos fuzzy F_i^β , $i = 1, 2, 3$, aos quais ficam agregadas as informações dos pesos relativos, passando assim a substituir os subconjuntos fuzzy F_i .

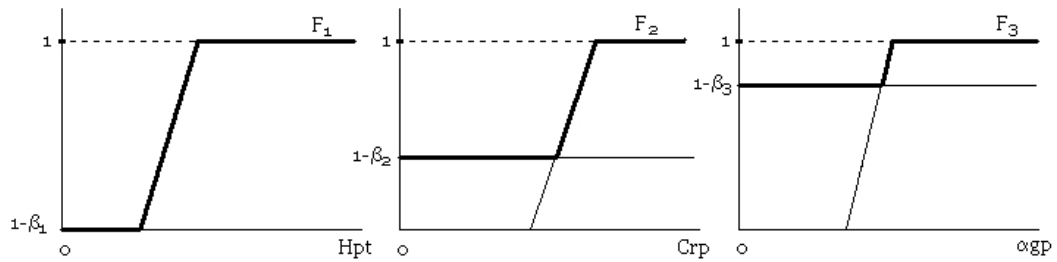
$$F_i^\beta = \max\{1 - \beta, F\} \quad (3.4)$$

Temos uma nova rotina, em que os três primeiros itens são idênticos aos três primeiros da anterior:

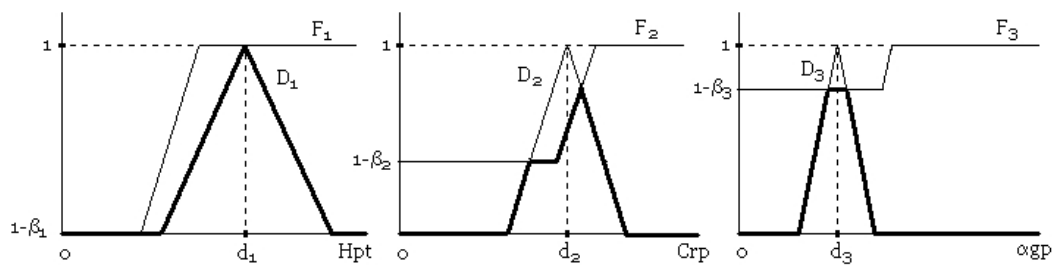
Para cada cão:

1. São inseridos d_1 , d_2 e d_3 (valores das PFA medidas)
2. Definem-se as funções de pertinência dos subconjuntos fuzzy D_1 , D_2 e D_3
3. Definem-se as funções de pertinência dos subconjuntos fuzzy F_1 , F_2 e F_3
4. São inseridos β_1 , β_2 e β_3 (pesos relativos das PFA)

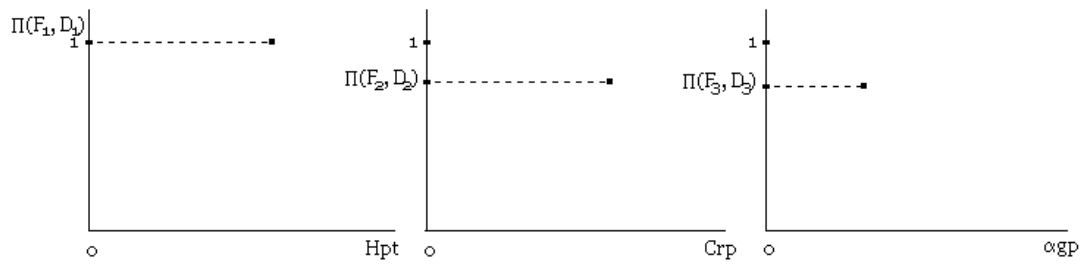
5. Redefinem-se as funções de pertinência de F_1 , F_2 e F_3 , através de (3.4)



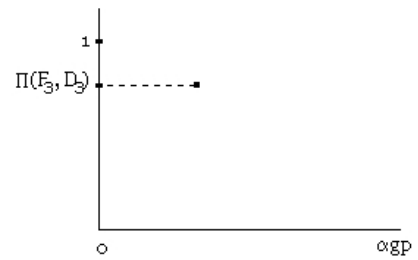
6. São calculados $F_1 \cap D_1$, $F_2 \cap D_2$, $F_3 \cap D_3$



7. São calculados $\Pi(F_1, D_1)$, $\Pi(F_2, D_2)$ e $\Pi(F_3, D_3)$



8. É calculado $\min\{\Pi(F_1, D_1), \Pi(F_2, D_2), \Pi(F_3, D_3)\}$



4. Resultados

Aqui apresentamos como exemplo três amostras de cães saudáveis e três de cães doentes.

Tabela 3: Valores das PFA em três cães doentes.

Cão doente	Hpt(g/L)	Crp(U/L)	αgp(g/L)
1	2,51	19,4	3,33
2	2,50	17,5	2,55
3	2,70	14,6	1,05

Tabela 4: Valores das PFA em três cães sadios.

Cão Sadio	Hpt(g/L)	Crp(U/L)	αgp(g/L)
1	0,07	8,0	0,82
2	0,47	16,0	1,01
3	1,32	13,0	1,39

E calculamos a possibilidade de diagnóstico para cada animal:

Tabela 5: Possibilidade de diagnóstico de parvovirose para cada cão doente.

Cão Doente	$\Pi(F_1, D_1)$	$\Pi(F_2, D_2)$	$\Pi(F_3, D_3)$	Possibilidade
1	1	1	1	1
2	1	1	1	1
3	1	0,33	0,82	0,33

Tabela 6: Possibilidade de diagnóstico de parvovirose para cada cão sadio.

Cão Sadio	$\Pi(F_1, D_1)$	$\Pi(F_2, D_2)$	$\Pi(F_3, D_3)$	Possibilidade
1	1	0	0	0
2	0	0,84	0,21	0
3	0,38	0,33	0,74	0,33

Apesar de a possibilidade de estar doente para o cão saudável 3 ser diferente de zero, o resultado é um número baixo, menor do que 0.5, indicando uma baixa relação entre a doença e os sintomas do animal. É, portanto, aceitável.

Porém, os dados do cão doente 3 fornecem uma possibilidade de valor muito baixo, abaixo de 0.5, o que não é aceitável. Mas há uma explicação para isto: não foram considerados os chamados pesos relativos referentes a cada proteína.

Os resultados obtidos aplicando-se os pesos relativos se apresentam na tabela a seguir:

Tabela 7: Possibilidade de diagnóstico de parvovirose para cada cão doente, levando-se em conta os pesos relativos.

Cão Doente	$\Pi(F_1, D_1)$	$\Pi(F_2, D_2)$	$\Pi(F_3, D_3)$	Possibilidade
1	1	1	1	1
2	1	1	1	1
3	1	0,6	0,82	0,6

Tabela 8: Possibilidade de diagnóstico de parvovirose para cada cão sadio, levando-se em conta os pesos relativos.

Cão Sadio	$\Pi(F_1, D_1)$	$\Pi(F_2, D_2)$	$\Pi(F_3, D_3)$	Possibilidade
1	0	0,6	0,6	0
2	0	0,84	0,6	0
3	0,38	0,6	0,74	0,38

Apesar do pequeno aumento na possibilidade do diagnóstico positivo de parvovirose para o terceiro cão sadio, o aumento da possibilidade para o terceiro cão doente foi significativo. Desta maneira, os cães sadios que eram corretamente classificados como tal continuaram sendo considerados sadios.

Dentre os cães doentes, por sua vez, os que já eram doentes permaneceram com o diagnóstico positivo e o que estava sendo erroneamente considerado sadio passou a ter muito mais possibilidades de ser classificado como doente.

5. Conclusões

Os resultados do programa se mostraram condizentes com os diagnósticos realizados, permitindo-nos concluir que o uso dos conjuntos fuzzy foi bastante adequado. O emprego dos pesos relativos foi determinante em alguns casos, tornando as saídas do modelo mais fiéis à realidade.

No entanto, este não é um método validado. Para tanto, seria necessário testar o modelo em uma grande quantidade de animais, sadios e doentes, obtendo-se um alto número de diagnósticos corretos e/ou consultar um especialista para confirmar se essas três PFA são suficientes para um diagnóstico e criticar os valores dos pesos relativos deduzidos.

Agradecimentos

Agradeço à CAPES pelo financiamento e ao grupo de Seminários de Lógica Fuzzy da USP pelos conselhos e ajuda na finalização deste trabalho.

Referências

- Barros, L. C. e Bassanezi, R. C. (2006). *Tópicos de Lógica Fuzzy e Bio-matemática*. Campinas.
- Delayte, E., Elias, F., Hagiwara, M., Kogika, M., Notomi, M., e Pereira, D. (2003). Determinação sérica de haptoglobina, ceruloplasmina e alfa-glicoproteína ácida em cães com gastrenterite hemorrágica. *Ciência Rural*, 33:513–517.

Sanchez, E. (1998). Fuzzy logic and inflammatory protein variations. *Clinica Chimica Acta*, 270:31–42.