1. Publicação nº	2. Versão	3. Data	5. Distribuição
INPE-3053-PRE/474		*Abril, 1984*	☐ Interna ☒ Externa ☐ Restrita
4. Origem *DIN/DPD*	Programa *ATCOMP*		

6. Palavras chaves - selecionadas pelo(s) autor(es)
INDUÇÃO
FUZZY SETS
REGRAS DE DECISÃO

7. C.D.U.: *681.325.6*

8. Título *INPE-3053-PRE/474*

INDUÇÃO DE REGRAS DE DECISÃO NEBULOSAS

10. Páginas: *06*

11. Última página: *05*

12. Revisada por

Flávio R. D. Velasco

9. Autoria *Orion de Oliveira Silva*
Celso de Renna e Souza

Assinatura responsável Orion de Oliveira Silva

13. Autorizada por

Nelson de Jesus Parada
Diretor Geral

14. Resumo/Notas

Neste trabalho são apresentados resultados encontrados na literatura, os quais cobrem tópicos da teoria de conjuntos nebulosos. A apresentação é uniforme, com exemplos e comentários, e são apresentados resultados obtidos no estudo do problema de indução de regras de decisão, bem como algumas aplicações na área de diagnóstico médico. O problema é encarado da maneira apresentada por Larson e Michalski [5], com extensões baseadas na natureza das premissas e conclusões das regras (booleanas ou nebulosas).

15. Observações *Este trabalho será apresentado no IEEE International Conference "LATINCON 84", na cidade do México, México, a realizar-se de 9 a 13 de julho de 1984.*

INDUÇÃO DE REGRAS DE DECISÃO NEBULOSAS

Oríon de Oliveira Silva
Celso de Renna e Souza

Instituto de Pesquisas Espaciais - INPE
Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq
Caixa Postal 515 - 12200 - São José dos Campos - SP - Brazil

Resumo

Neste trabalho são apresentados resultados encontrados na literatura, os quais cobrem tópicos da teoria de conjuntos nebulosos. A apresentação é uniforme, com exemplos e comentários, e são apresentados resultados obtidos no estudo do problema de indução de regras de decisão, bem como algumas aplicações na área de diagnóstico médico. O problema é encarado da maneira apresentada por Larson e Michalski [5], com extensões baseadas na natureza das premissas e conclusões das regras (booleanas ou nebulosas).

Abstract

In this paper results and topics in fuzzy set theory found in the literature are presented. The presentation is uniform, with examples and comments. Results obtained in the study of decision rule induction and applications to medical diagnasis an also presented. This last problem is approached as proposed by Larson and Michalski [5], with extensions based on the types of premisses and conclusions (boolean or fuzzy).

Introdução

Desde o momento que a representação do conhecimento especializado começou a ser apresentada por meio de "Regras de Decisão" ou "Sistemas de Produção", ela tem recebido muita atenção na literatura, principalmente de Sig e Chen [13], Shortliffe [10], Larson e Michalski [5], e tem sido aplicada em inúmeros campos (Elcoke and Donald, [10]; Winston, [16]).

Os dois motivos principais para induzir regras de decisão são: a) reduzir o tamanho do banco de regras;
b) modelar o aprendizado de novas regras por indução.

O MYCIN Shortliffe [10]); é um exemplo bastante conhecido de um sistema que faz diagnóstico médico usando regras de decisão para representação do conhecimento.

Neste trabalho usaram-se, como exemplos, algumas regras práticas na área de diagóstico médico; no entanto, o sistema poderá manipular regras das áreas de prospecção geológica, reconhecimento de padrões, diagnóstico de falhas em sistemas espaciais e ainda em qualquer outra área onde se possa colocar o conhecimento em forma de regras de decisão.

O problema de indução de novas regras de decisão - objeto deste trabalho - a partir de regras dadas é colocado da seguinte maneira: dado um conjunto de regras de decisão com a mesma conclusão (nebulosa ou booleana) e premissas (nebulosas ou booleanas) diferentes, juntamente com "graus de certeza" (número do intervalo [0,1]) sobre as implicações correspondentes, deve-se:

a) transformar o conjunto de regras em um novo conjunto equivalente de regras (booleanas) modificando as funções de pertencimento das premissas e/ou conclusão, bem como o "grau de certeza" da implicação;
b) transformar o conjunto de regras (booleanas) modificado em (a) em um novo conjunto que inclua uma ou mais regras induzidas a partir daquele de acordo com alguns critérios;
c) transformar este novo conjunto em uma regra única que contenha este conjunto;
d) atribuir "grau de certeza" a esta regra única segundo alguns critérios;
e) atribuir distribuições e "grau de certeza" a esta regra única (torná-la nebulosa).

Indução de Regras de Decisão Boolenas

Nesta seção será abordado o problema de indução aplicado a regras de decisão da maneira proposta por Michalski [6] e Larson e Michalski [5]. "Regras de Decisão" são semelhantes às "Regras de Produção" usadas por vários autores, tais como: "Shortliffe [10]", "Capocelli e De Luca [2]", "Waterman [15]", "Rycherier [8]" e "Coulon et alii [3]". O esquema geral é:

CONDIÇÃO ==> DECISÃO

isto é, se uma situação satisfaz uma CONDIÇÃO, então a regra designa para ela uma DECISÃO (que pode ser inválida ou não ter validade prática). A idéia é paralela à do comando "se ... estão" em linguagens de programação e das cláusulas de Horn em PROLOG.

O problema de Regras de Decisão Boolenas vem sendo tratado desde 1979 (Silva, [11]), e são descritos aqui alguns resultados encontrados.

O problema é dividido em três partes, a saber:

a) transformar um dado conjunto de regras de decisão em um novo conjunto que inclua uma nova regra induzida a partir daquele, de acordo com alguns critérios;
b) transformar o novo conjunto de regras em uma regra única que contenha este conjunto;
c) atribuir um "grau de certeza" (um número do intervalo [0,1]) (Shafer, [9]) a esta regra única.

Definição 2.1 - Regra de Produção

Uma regra de produção é definida como um par ordenado de listas de símbolos (LE, LD)*, onde "LE" é o lado esquerdo da regra e "LD" é o lado direito.

Definição 2.2 - Sistema de Produção

Um sistema de produção é definido em função de três componentes básicos: um conjunto de regras de produção, um banco de dados (ou simplesmente uma coleção de símbolos) e um interpretador para estas regras de produção. O interpretador opera, essencialmente, escolhendo o "LE" de cada regra, comparando-o com o banco de dados, incluindo ou não a conclusão desta regra no banco.

---

* *Está subentendido que "LE" implica "LD".*

Definição 2.3 - Regra Consistente

Uma nova regra é "consistente" com um dado conjunto de regras segundo simplificação se, e somente se, para qualquer situação em que a nova regra indicar uma conclusão as regras iniciais indicarem também esta conclusão.

Definição 2.4 - Grau de Certeza

O "grau de certeza" é aqui definido com um número do intervalo [0,1] que determina o grau de confiabilidade (confiança) que se tem em uma "regra de produção" ou "regra de decisão". O grau de certeza de uma regra é dado por um especialista da área e normalmente é inteiramente subjetivo, embora possa em alguns casos ser deduzido a partir de dados estatísticos.

Definição 2.5 - Grau de Certeza Global

Sejam $x_1, x_2, \ldots, x_n$ os graus de certeza das regras $r_1, r_2, \ldots, r_n$ que tenham a mesma decisão; o grau de certeza global de $x_1, \ldots, x_n$ é dado por Shafer, [9]:

$$\mu(X) = 1 - e^{-Z} \quad \text{onde: } Z \stackrel{\Delta}{=} \sum_{i=1}^{n} \log(1 - xi). \qquad (1)$$

Definição 2.6 - Regra Induzida Consistente

Uma nova regra induzida a partir de um conjunto de regras é "consistente" se, e somente se, para qualquer situação em que as regras iniciais designarem uma conclusão, a nova regra induzida designar esta mesma conclusão e esta regra induzida tiver o mesmo grau de certeza global do conjunto de regras que a induz.

O critério para dizer que uma regra qualquer é "boa", "satisfatória", etc. é escolhido de acordo com o problema. Podem-se citar alguns exemplos de critérios utilizados, tais como memória requerida pelo banco de regras, tempo de avaliação da regra, custo para obter a informação necessária à avaliação da regra, etc.; no presente caso, o critério utilizado será o da simplicidade da expressão booleana,no sentido de minimização de dois níveis (Petersons and Hill, [7]).

Suponha-se que o sistema contenha várias regras com a mesma decisão: $A = \{b_1 \to F, b_2 \to F, \ldots bn \to F\}$. A existência simultânea dessas regras no banco de dados tem o sentido considerado: $b = (b_1 \vee b_2 \vee \ldots \vee b_n)$.

A cada regra $bi \to F$ tem-se associado um "grau de certeza" (número do intervalo $[0,1], \mu_A(bi \to F)$).

Agora o problema será abordado por partes:

a) induzir uma nova regra b_{n+1} a partir do conjunto A, formando o conjunto $B = A \cup \{b_{n+1} \to F\}$

b) simplificar o conjunto B e atribuir o "grau de certeza" a esta nova regra.

Postulado 2.1 - A nova regra é induzida de tal maneira que o novo conjunto $B = A \cup \{b_{n+1} \to F\}$ terá o mesmo grau de certeza global de A. Isto se justifica pela noção de que uma regra induzida de um conjunto de regras dado não poderá fornecer maior credibilidade (confiança) do que o conjunto que a gerou.

Postulado 2.2 - Aplicando sucessivamente o Postulado 2.1, cada regra induzida do conjunto anterior não poderá ter menor "grau de certeza" do que o menor "grau de certeza" do conjunto de regras que o gerou. Quando isto acontecer, terminará o processo de indução.

A metodologia será explicada com um exemplo.

Exemplo 2.1 - Processo de indução de uma regra (ponto)

Seja: $b_1 = A \wedge (\neg B \vee C) \to F \quad \mu(b_1) = 0{,}6;$

$b_2 = \neg A \wedge C \to F \quad \mu(b_2) = 0{,}5;$

$b_3 = A \wedge C \to F \quad \mu(b_3) = 0{,}7.$

Tomando os oito possíveis casos de valores booleanos para as proposições A, B e C, apresenta-se a Tabela 2.1, que foi elaborada com base nos métodos de Mill e na Álgebra Nebulosa (Zadeh, [14]).

Tabela 2.1 - Tabela Lógica.

A	B	C	0,6	0,5	0,7	$b = b_1 \vee b_2 \vee b_3$	F	$\mu A(F)$
			b_1	b_2	b_3			
0	0	0	0	0	0	0	?	?
0	0	1	0	1	0	1	1	0,5
0	1	0	0	0	0	0	?	?
0	1	1	0	1	0	1	1	0,5
1	0	0	1	0	0	1	1	0,6
1	0	1	1	0	1	1	1	0,7
1	1	0	0	0	0	0	?	?
1	1	1	1	0	1	1	1	0,7

No caso de haver mais de um bi com o valor 1 (em uma linha), postula-se para F o "grau de certeza" máximo dos bi's $[\max(\mu_A(bi))]$, o que é perfeitamente justificado pela operação de "conjuntos nebulosos" onde $\mu_A(x) \vee \mu_B(x) = \max[\mu_A(x), \mu_B(x)]$.

Na Tabela 2.1 pode-se observar a seguinte situação no cubo-3 (Figura 2.1).

O critério usado aqui para a indução é o seguinte: induzir um ponto do subcubo (Peterson and Hill, [9]) que tiver maior cobertura, ou maior número de pontos, ou maior dimensão, com valor 1(um) para F. Se dois ou mais subcubos tiverem a mesma cobertura, escolher-se-á aquele que tiver o maior "grau de certeza global" calculado sobre os pontos deste subcubo. Repete-se o processo enquanto o novo ponto gerado não tiver menor "grau de certeza global" do que o menor "grau de certeza" do conjunto de regras que o gerou. Não se pode induzir nenhum ponto em "cubos completos", ou seja, em cubos que têm o número de pontos (vértices) igual a uma potência de 2, ou seja quadrado, cubo, etc.

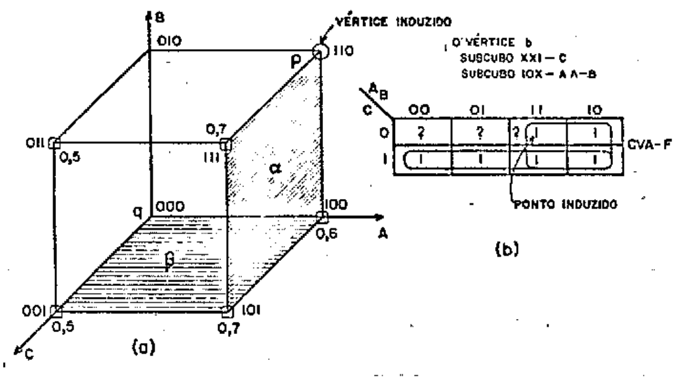

Figura 2.1 - Indução e Minimização.

Por exemplo, no caso da Figura 2.1 escolheu-se o ponto (vértice) "p" pelo fato de o "grau de certeza global" $\mu_{\alpha}(p) = 0,6\ 1134$ ser maior do que o "grau de certeza global" $\mu_{\beta}(q) = 0,608513$, ou seja, induz-se, de preferência, que F é verdadeiro também para 110, uma vez que o "grau de certeza global" no subcubo incompleto α é maior do que o de β.

No caso citado (Exemplo 2.1) é bastante simples produzir uma forma booleana $b' \cdot b$ que seja mínima. No caso, $b' = C \vee A$ cobre exatamente b.

Portanto, as três regras de α e a regra induzida podem ser substituídas por: $C \vee A \to F$.

O "grau de certeza" da nova regra P_4 é calculado usando o Postulado 2.1 e a Definição 2.6 previamente enunciados: $\mu(P_4) = 0,608113$, obtido por: $\mu(P_4) = 1 - e^{-Z} = 1 - 0,3914871 = 0,6081513$, onde $Z = 0,3333 \ldots [-\log(1-0,5) - \log(1-0,7) - \log(1-0,6)]$, sendo os valores 0,5; 0,7 e 0,6 os graus de certeza das regras originais.

Inferência Nebulosa

O tratamento matemático visto em Silva e Souza [12] será usado no problema de inferência de regras de decisão (produção) quanto as premissas e conclusões das regras forem nebulosas.

O problema será dividido em:

a) transformar o conjunto de regras dado em um novo conjunto equivalente de regras (booleanas) modificando as funções de pertencimento das premissas e/ou conclusão, bem como o "grau de certeza" da implicação.

b) usar o caso anterior (Seção 2) de indução.

Definição 31. – Função F (Silva e Souza [12])

$$F(k, x, y) = \begin{cases} y^k & \text{se } x \leq y, \\ x^{(1+x-y)k} & \text{se } x > y, \\ 1 & \text{se } k = x = y = 0. \end{cases} \quad (2)$$

Definição 3.2 – Função g (Silva e Souza, [12])

$$g(k, y) = \begin{cases} y^k, & y \text{ e } k \in [0,1], \\ 0 \ \forall \ y=1 \text{ e } k=0. \end{cases} \quad (3)$$

Sejam dadas as seguintes regras de decisão nebulosas:

$$\begin{cases} B_1 = A_1 \wedge (\sim B_1 \vee C_1) \xrightarrow{0,66} S, \\ B_2 = \sim A_2 \wedge C_2 \xrightarrow{0,71} S, \\ B_3 = A_3 \wedge C_3 \xrightarrow{0,62} S, \end{cases}$$

Onde:

A_1 – a criança é recém-nascida (muito nova);

B_1 – urina normalmente;

C_1 – tem contagem alta de globulos brancos na urina;

A_2 – tem menos de dois meses (mais do que nova);

C_2 – tem contagem muito alta de glóbulos brancos na urina;

A_3 – é nova;

C_3 – tem contagem moderadamente (mais ou menos) alta de glóbulos brancos na urina;

S – tem chance de ITU (Infecção de Trato Urinário).

Os conjuntos nebulosos originais (primitivos) são: $A_2 = A^{1,25}$; $A_3 = A^{2,0}$ e $A_1 = A$ (a criança é recém-nascida); $B = B_1$ = urina normalmente; $C_1 = C^2$; $C_2 = C_1^2 = C^4$; $C_3 = C$ = tem contagem mais ou menos alta de glóbulos brancos na urina.

Então o conjunto primitivo é postulado como:

$A = A_1$; $B = B_1$; $C = C_3$; $D = \sim B_1$ e $E = \sim A_2$.

Tem-se, reescrevendo as regras em termos das distribuições primitivas,

$$\begin{cases} B_1 = A \wedge (D \vee C^2) \xrightarrow{0,66} S, \\ B_2 = E \wedge C^4 \xrightarrow{0,71} S, \\ B_3 = A^2 \wedge C \xrightarrow{0,62} S. \end{cases}$$

Primeiro, transformar-se-ão as regras B_1, B_2 e B_3 em:

$$\begin{cases} B_{11}^1 = A \wedge (D \vee C) \xrightarrow{y_1} S, \\ B_{21} = E \wedge C \xrightarrow{y_2} S, \\ B_{31} = A \wedge C \xrightarrow{y_3} S, \end{cases}$$

e calcular-se-ão y_1, y_2, y_3.

Dar-se-á como exemplo apenas a regra B_{21}:

$A_2 =$

a_{21}	a_{22}	a_{23}
10 dias	40 dias	60 dias
1,0	0,80	0,72

;

$C =$

c_1^1	c_2	c_3
8.000	10.000	100.000
0,84	1,00	0,89

;

$E =$

b_1	b_2	b_3
60 dias	70 dias	100 dias
0,4	0,90	1,0

;

$C_2 = C^4 =$

c_{21}	c_{22}	c_{23}
8.000	10.000	100.000
0,50	1,00	0,62

;

$S =$

S_1	S_2	S_3
0,50	0,75	1,0

.

Então tem-se:

$B_2 =$

(a_{21}, c_{21})	(a_{21}, c_{22})	(a_{21}, c_{23})
0,4	0,4	0,40

(a_{22}, c_{21})	(a_{22}, c_{22})	(a_{22}, c_{23})	(a_{23}, c_{21})
0,50	0,90	0,62	0,50

(a_{23}, c_{22})	(a_{23}, c_{23})
1,00	0,62

;

B_{21} =

(e_1, c_1)	(e_1, c_2)	(e_1, c_3)	(e_2, c_1)
0,40	0,40	0,40	0,84

(e_2, c_2)	(e_2, c_3)	(e_3, c_1)	(e_3, c_2)	(e_3, c_3)
0,90	0,89	0,84	1,0	0,89

.

Calculando $F(1, \overline{B}_2^1, \overline{S^{0,5}}$ e compondo-o com B_{21}, tem-se:

$$B_{21} \circ F(1, \overline{B}_2^1, \overline{S^{0,5}}) = \left|\frac{S_1}{0{,}71}\right|\frac{S_2}{0{,}87}\left|\frac{S_3}{1{,}0}\right| ,$$

onde $0{,}5^{0,5} = 0{,}71 \Longrightarrow y_2 = 0{,}5$,

isto significa que: $B_{21} = E \wedge G \xrightarrow{0,5} S$.

Usando o mesmo critério para todas as regras tem-se:

$$\begin{cases} B_{11} = A \wedge (D \vee C) \xrightarrow{0,6} S \\ B_{21} = E \wedge C \xrightarrow{0,5} S \\ B_{31} = A \wedge C \xrightarrow{0,7} S \end{cases} \equiv \begin{cases} B_{11} = A \wedge (\neg B \vee C) \xrightarrow{0,6} S \\ B_{21} = \neg A \wedge C \xrightarrow{0,5} S \\ B_{31} = A \wedge C \longrightarrow S \end{cases}$$

Já se sabe que a regra induzida é $A \wedge B \wedge \neg C \xrightarrow{0,61} S$, é que a regra que substitui as anteriores é $A \vee C \xrightarrow{0,68} S$, ou seja, a regra induzida é:

"Se a criança é recém-nascida, urina normalmente e não tem contagem mais ou menos alta de glóbulos brancos na urina, então tem chance de ITU em 61%".

Admitindo que esta regra seja aceita como razoável por especialistas, a sua inclusão no conjunto de regras e subsequente simplificação é regra única equivalente:

"Se a criança é recém-nascida ou tem contagem maior ou menor alta de glóbulos brancos na urina, então a chance de ITU é 60,8%".

CONCLUSÃO

O presente documento teve como objetivo descrever alguns resultados encontrados pelos autores no tratamento do problema de indução nebulosa aplicando o "modus ponens modificado".

A maior limitação do presente enfoque é o fato de que a indução é feita somente do lado esquerdo da regra de decisão, ou seja, em suas premissas, embora se tenha permitido, no caso de regras nebulosas, modificadores lingüísticos na conclusão.

Outra dificuldade independente do enfoque tomado é a de como "provar" a validade da indução feita, se é que esta validade é possível de ser "provada".

Os autores acreditam que a forma de indução de regra quando as premissas são nebulosas poderá ser feita diretamente usando o mecanismo de minimização de funções nebulosas sem necessidade de redução feita ao caso booleano. Entretanto, o problema de minimização de funções nebulosas ainda não está totalmente resolvido (Bhat, [1]).

Entre as vantagens do método de indução estão generalidade e simplicidade.

O método pode ser aplicado a várias tarefas indutivas. Também é fácil compreender as informações e a interpretação dos resultados.

REFERÊNCIAS BIBLIOGRÁFICAS

[1] BHAT, K.V.S. Further remarks on Neff-Kandel algorithm for minimization of fuzzy conjuctive forms. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-12(3):430-431, May/June 1982.

[2] CAPOCELLI, R.M.; DE LUCA, A. Fuzzy sets and decision theory. *Information and Control*, 23:446-473, 1973.

[3] COULON, D.; AKMIN, Y.; LANEL, J.M.; MONGILS, M. *A language to describe knowledge by a network of procedures*. Nancy, Universities of Paris, Centre de Recherche in Informatique, 1977.

[4] ELCOCK, E.W.; DONALD, M. *Machine intelligence*. New York, John Wiley, 1977.

[5] LARSON, J.; MICHALSKI, R.S. Inductive inference of VL decision rule. *SIGART Newsletter*, 63:38-44, June 1977.

[6] MICHALSKI, R.S. *A variable-valued logic systems as applied to picture description and recognition*. Urban, Department of Computer Science, University of Illionois, 1972.

[7] PETERSON, C.R.; HILL, F.J. *Introduction to switching theory and logical design*. London, John Wiley, 1968.

[8] RYCHERIER, M.D. *The student production system. A study of encoding knowledge in production systems*. Pittsburg, Carnegie - Mellon University, 1975.

[9] SHAFER, G. *A mathematical theory of evidence*. Princeton, NJ, Princeton University Press, 1976.

[10] SHORTLIFFE, E.H. *Computer - based medical consultation MYCIN*. Ph.D., Amsterdan, Stanford University School of Medicine. Division of Clinical Pharmocology, 1975. (Artificial Intelligence Series, 2).

[11] SILVA, O.O. *Indução de regras de decisão*. São José dos Campos, SP, INPE, maio 1979. (INPE-1492-RPE/036).

[12] SILVA, O.O.; SOUZA, C.R. *Changing the fuzzy rule of Detachment*. 12th International Symposium on Multiple Valued Logic, Paris, France, May 1982.

[13] SIG, P.; CHEN, C.S. Minimization of fuzzy, function. *IEEE Transaction on Computer*, C21(1): 100-102, Jan. 1972, Short notes.

[14] ZADEH, L.A. Fuzzy sets. *Information and Control*, 8:338-353, 1965.

[15] WATERMAN, P.A. *Adaptative production systems.* Pittsburgh, Department of Physchology. Carnegic Mellon University, 1974.

[16] WINSTON, P.H. *Artificial Intelligence.* New York, Addison-Wesley, 1977.