

# REDES BAYESIANAS(CPN) X REDES CAUSAIS NEBULOSAS (CFN)

LEILTON S. LEMOS , KARL H. KIENITZ

*Departamento de Engenharia Elétrica e Computação  
Instituto Tecnológico De Aeronáutica - ITA.*

*Praça Mal. Eduardo Gomes, 50, Vila das Acácias, São José dos Campos, SP, BR, 12.228-900*

*E-mails: leiltonz@zipmail.com.br , kienitz@ele.ita.br*

**Resumo**— Redes Bayesianas (CPN) tem atraído grande atenção de pesquisadores e da indústria em geral; porém ao contrário de outros sistemas especialistas estas redes necessitam de experiência na ferramenta e domínio do problema. Mesmo com estas características, algumas incapacidades de modelamento podem restringir e até impossibilitar o bom uso desta ferramenta para auxílio a decisões. Este trabalho ataca estas dificuldades com a proposição de um novo modelo: Redes Causais Nebulosas, que busca manter a estrutura e qualidades de uma Rede Bayesiana retirando restrições conhecidas.

**Abstract**— Bayesian Networks (CPN) have attracted great attention of researchers and industry; however unlike other expert systems these networks needs tool experience and problem domain expertise. Even with these characteristics, some modeling incapacities may restrict or disable the use of this tool as decisions aid. This work attacks these difficulties with the proposition of a new model: Fuzzy Causal Networks, its objective is to maintain the structure and qualities of a Bayesian Network removing known handicaps.

**Keywords**— Bayesian Networks, Fuzzy, graph, Expert System.

## 1 Introdução

As Redes Causais Probabilísticas (CPN) com sua vertente mais conhecida “Redes Bayesianas”, tem sido usadas para desenvolver sistemas especialistas que possuem domínio e sistema conhecido. CPN provêm uma arquitetura genérica de alto nível para sensoriamento ou entrada de dados de múltiplos tipos. Como tanto, adequam-se bem em sistemas de previsão tipo causa-efeito.

Devido a simplicidade matemática, e sua representação em grafos, são altamente intuitivas, estruturadas, bem como de fácil visualização.

Do ponto de vista da teoria de probabilidade, CPN são arquiteturas que garantem a consistência e coerências do Modelo de Causas Probabilísticas de um domínio e mantêm o equilíbrio probabilístico do modelo caso inseridas novas informações. Estas redes possuem um modelo de inferência muito simples baseado na teoria de Bayes (Finn, 1996) que permite a propagação de novas informações (evidências) para as demais variáveis (nós da rede).

## 2 Rede Bayesiana

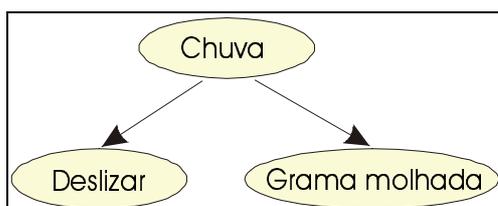
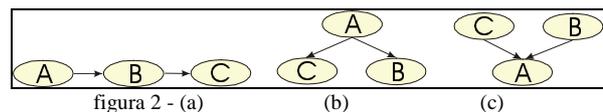


figura 1 - Exemplo de uma rede simples

As Redes Bayesianas consistem de um conjunto de variáveis e conexões direcionadas entre estas variáveis;

sendo este modelo chamado *grafo direcionado (GD)*. Exemplificando: da figura 1, podemos extrair três variáveis ( “Chuva”, “Deslizar” e “Gramma molhada” ) e duas conexões direcionadas: uma saindo de “Chuva” direcionada para “Deslizar” e a outra saindo de “Chuva” direcionada para “Gramma molhada”.

Cada variável refere-se a ocorrência de um evento e possui n estados em que pode se encontrar, esta variável é representada na forma “Variável[estados]”. Exemplificando, uma variável “Doença” pode ter vários estados, como: resfriado, bronquite, tuberculose, câncer, desconhecida, etc. ( representada como: *Doença [resfriado, bronquite, tuberculose, câncer, ...]* .)



As conexões (direcionadas) entre as variáveis podem ter o seguinte formato: *seriais* (figura 2.a), *divergentes* (figura 2.b) e *convergentes* (figura 2.c). Estes três casos cobrem todas as formas nas quais as evidências de ocorrência de um estado em uma variável podem ter seus efeitos (ou impactos) propagados para as demais variáveis da rede.

### 2.1 Separação-d

Dado duas variáveis A e C, estas tem separação-d se todos os caminhos que conduzem de A a C possuem uma variável intermediária X que:

- Tenha conexão serial ou divergente e estado de X conhecido;
- Ou, tenha conexão convergente e nem X nem seus descendentes tenham sido evidenciados (condição para evitar ciclos).

Se para todas as variáveis conectadas por uma variável intermediária tem separação-d, este grafo é intitulado *grafo acíclico direcionado (GAD)*. Ou seja, é um *grafo direcionado* que não possui ciclos (pré-requisito de Redes Bayesianas). Sendo os GADs uma representação não ambígua de dependência.

## 2.2 Axiomas básicos

A probabilidade  $P(A)$  de um evento é um número no intervalo  $[0,1]$  e:

- $P(A) = 1$  se e só se  $A$  é certo;
- Se  $A$  e  $B$  são mutuamente exclusivos, então  $P(A \vee B) = P(A) + P(B)$

Da regra de Bayes:[6,cap.2]

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}$$

Com a adição de  $C$ :

$$P(A | B, C) = \frac{P(B | A, C)P(A | C)}{P(B | C)}$$

Observa-se assim a possibilidade de generalização para permitir que a regra leve em conta todas as variáveis do grafo. Foram assim, definidos vários mecanismos de propagação que mantêm-se consistentes com a teoria da probabilidade mas buscam reduzir a carga computacional/complexidade da propagação das evidências aos demais nós da rede (Pearl, 1988).

## 2.3 Mecanismo básico de propagação em Redes Bayesianas

A propagação funciona em hierarquia com as folhas originais (sem pais) coletando as evidências de seus nós filhos e calculando a influência destas evidências em suas crenças; seus filhos por sua vez executam o mesmo processo (bottom-up), sendo que este processo corre consecutivamente até um nó sem filhos (nó final/folha extremidade). Após o nó pai ter coletado as evidências ele propaga suas crenças para seus filhos e assim por diante na forma de uma cascata (top-down). (Finn, Olesen e Andersen, 1990)

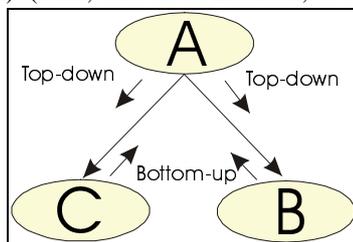


figura 3 - Propagação em uma rede

- 1º passo: Bottom-Up;
- 2º passo: atualiza a crença BEL de A;
- 3º passo: Top-Down
- 4º passo: atualiza a crença BEL de C e B;

## 2.4 Uso de uma Rede Bayesiana e principais problemas

Como mencionado anteriormente, sendo o sistema conhecido pode-se modelá-lo usando Redes Bayesianas. No entanto algumas dificuldades surgem devido a restrições da rede:

- Para determinar a tabela de influências entre dois nós, faz-se necessário o uso de um ou mais especialistas que distribuem os pesos baseados em sua experiência. Caso contrário, é necessária a obtenção um vasto conjunto de dados, que representem o sistema como um todo, para treinar a rede. Sobrando a dúvida, se o conjunto usado foi suficientemente representativo.
- Os especialistas normalmente retornam uma distribuição de possibilidades e não de probabilidades (dificilmente o dado estará em forma de probabilidade). Assim, estes dados se encaixam melhor em conjunto nebulosos. (Zadeh, 1973 e 1979)
- Devido a características do modelo, a entrada de evidências nos nós é sempre do tipo mutuamente exclusiva (se um estado ocorre os demais não ocorrem) ou seja, não importa o número de possíveis estados que um nó possa ter, apenas um dos estados é marcado como ativo (100%) e os demais como inativos (0%). Criando assim uma dificuldade na inserção de evidências conjuntas (em caso de incerteza gerada pelo modelo) ou evidências que não tem seus limites bem definidos. (outra facilidade de conjuntos nebulosos).

Ex. Sensor com defeito, não há certeza da precisão de sua resposta, a variável vazão[10, 20, 30, 40], pode ter seus estados evidenciados como: [10=90%, 20=10%, 30=0%, 40=0%]. Esta variável não pode ser usada diretamente no modelo CPN, necessita ser pré-processada para escolher um dos estados como evidenciado (condição de mutuamente exclusivos).

Este mesmo problema se reflete no resultado final da propagação das evidências inseridas na rede; os nós folhas terão como resposta probabilidades de ocorrência distribuídas em seus diversos estados e, caso deseje-se uma saída contínua ou discreta, porém exclusiva, (apenas um dos estados) há a dificuldade da extração destes dados na forma de um valor único, seja discreto (correspondente aos estados) ou contínuo. (outra facilidade de conjuntos nebulosos). Ex. A variável vazão[10, 20, 30, 40], pode ter sua possibilidade de ocorrência de estados como [ 10=10%, 20=15%, 30=40%, 40=35%] necessitando de um pós-processamento para escolher um dos estados como válido.

### 3 Rede Causal NEBULOSA (CFN)

A proposta deste trabalho busca utilizar a lógica nebulosa para suprir as dificuldades encontradas devido o bom acoplamento da teoria nebulosa com o modelo já conhecido; ou seja:

- Os modelos montados em grafos podem facilmente ser convertidos para a forma de regras nebulosas pois, o grafo em sua essência, é previamente definido na forma de regras se-então em Lógica de Primeira Ordem (LPO). Estas regras nebulosas podem portanto ser apresentadas ao usuário na mesma interface de grafos já conhecida e definida; unindo assim as qualidades do grafo com o poder da lógica nebulosa.
- O uso de codificadores/decodificadores nebulosos nas extremidades da rede, que recebem/amostram valor quantificado e não estados, facilita a entrada e leitura dos dados finais; estes agora não precisando mais de um tratamento adicional (em outro sistema) para adequá-los a interface Bayesiana, possibilitando o uso dos dados sem pré-processamento na entrada bem como obtenção das saídas sem pós-processamento.

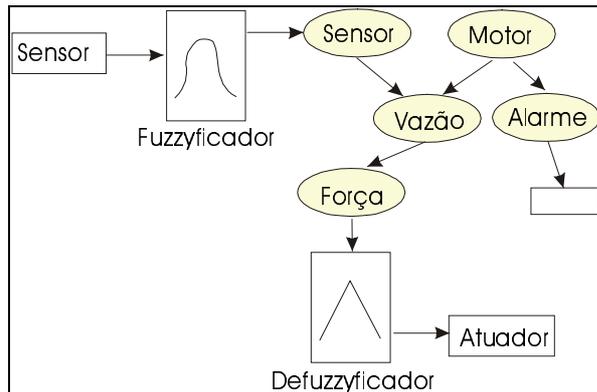


figura 4 - Exemplo de um sistema CFN

Uma das regras retiradas do exemplo (figura 4):  
Vazão → Força.

#### 3.1.1 Sistema da figura 4 como Rede Bayesiana

Seja a entrada “sensor de vazão” um índice entre 0 e 499, este valor pode ser codificado como 500 estados (vazão[0,...,499]) ou em estados representando faixas (vazão[nenhuma, baixa, ..., alta], tendo que passar por um pré-processamento em ambos os casos para ajuste dos estados da variável. A resposta desta variável (caso não evidenciada) também precisa de um pós-processamento.

Se esta entrada contém uma incerteza (o que não é aceito em CPN, pois a evidência não está exclusiva com 100% mas esta sim com uma possibilidade) após a propagação não se recupera o valor inicial com o qual a variável foi evidenciada.

#### 3.1.2 Sistema da figura 4 como Rede Nebulosa

Pela CFN, não há a necessidade do pré/pós-processamento e a resposta mantém-se a mesma que foi evidenciada.

#### 3.2 Mecanismo básico de propagação em CFN

Da LPO, pode-se obter regras do tipo se-então (se Vazão é V então Força é F)(Kienitz, 1990). E por implicação Fuzzy podemos representar as regras como: Vazão → Força.

Seja a implicação fuzzy  $a \rightarrow b$ , esta pode ser definida por uma relação fuzzy R. Esta relação R é representada por uma matriz associada com a implicação e pelo produto cartesiano  $U \times V$  (com  $u \in U$  e  $v \in V$ ) (Watanabe e Symon, 1992 e Mendel, 1995). Usando-se a implicação baseada no operador min, R é definida como:

$$\mu_R(u, v) = \min(\mu_a(u), \mu_b(v)) \quad \text{com } u \in U \text{ e } v \in V$$

Seja R a relação Fuzzy de U para V com  $u \in U$ , então o subconjunto  $v \in V$  induzido por u é dado pela composição de R e u; ou seja,  $v = u \circ R$ . E usando a regra composicional max-min, obtemos:

$$\mu_b(v) = \max_{u \in U} [ \min(\mu_a(u), \mu_R(u, v)) ]$$

#### 3.2.1 Para o caso de conexão em série:

Retirando o  $\mu_b(v)$  (original) da relação, obtem-se a função Top-Down de  $a \rightarrow b$ :

$$\mu_b(v) = \max_{u \in U} [ \min(\mu_a(u), \mu_R(u, v)) ]$$

De forma correlata pode-se obter a função Bottom-Up de  $a \rightarrow b$  será:

$$\mu_a(u) = \max_{v \in V} [ \min(\mu_a(u), \mu_b(v), \mu_R(u, v)) ]$$

#### 3.2.2 Fórmula genérica para $a \rightarrow b_1 \text{ E } b_2 \text{ E } \dots \text{ E } b_n$

Função Top-Down de  $a \rightarrow b_1 \text{ E } b_2 \text{ E } \dots \text{ E } b_n$

$$\min(\mu_{b_1}(v), \dots, \mu_{b_n}(v_n)) = \max_{u \in U} [ \min(\mu_a(u), (\mu_{b_1}(v), \dots, \mu_{b_n}(v_n)), (\mu_{R_1}(u, v), \dots, \mu_{R_n}(u, v_n))) ]$$

Da separação-d,  $\mu_{b_x}(v_x) = \min_{b_x}(\mu_{b_1}(v), \dots, \mu_{b_n}(v_n))$

Função Bottom-Up de  $a \rightarrow b_1 \text{ E } b_2 \text{ E } \dots \text{ E } b_n$  será:

$$\mu_a(u) = \max_{v \in V} [ \min(\mu_a(u), (\mu_{b_1}(v), \dots, \mu_{b_n}(v_n)), (\mu_{R_1}(u, v), \dots, \mu_{R_n}(u, v_n))) ]$$

#### 3.2.3 Fórmula genérica de $a_1 \text{ E } \dots \text{ E } a_m \rightarrow b_1 \text{ E } \dots \text{ E } b_n$

Função Top-Down de  $a_1 \text{ E } \dots \text{ E } a_m \rightarrow b_1 \text{ E } b_2 \text{ E } \dots \text{ E } b_n$

$$\min(\mu_{b_1}(v), \dots, \mu_{b_n}(v_n)) = \max_{u_1 \in U_1 \dots \max_{u_m \in U_m} [ \min((\mu_{a_1}(u_1), \dots, \mu_{a_m}(u_m)), (\mu_{b_1}(v), \dots, \mu_{b_n}(v_n)), (\mu_{R_1}(u, v), \dots, \mu_{R_n}(u, v_n))) ]$$

Função Bottom-Up de  $a \rightarrow b_1 \text{ E } b_2 \text{ E } \dots \text{ E } b_n$  será:

$$\mu_{a_1}(u_1) = \max_{v \in V} [ \min((\mu_{a_1}(u_1), \dots, \mu_{a_m}(u_m)), (\mu_{b_1}(v), \dots, \mu_{b_n}(v_n)), (\mu_{R_1}(u, v), \dots, \mu_{R_n}(u, v_n))) ]$$

Da separação-d,  $\mu_{b_x}(v_x) = \min_{b_x}(\mu_{b_1}(v), \dots, \mu_{b_n}(v_n))$

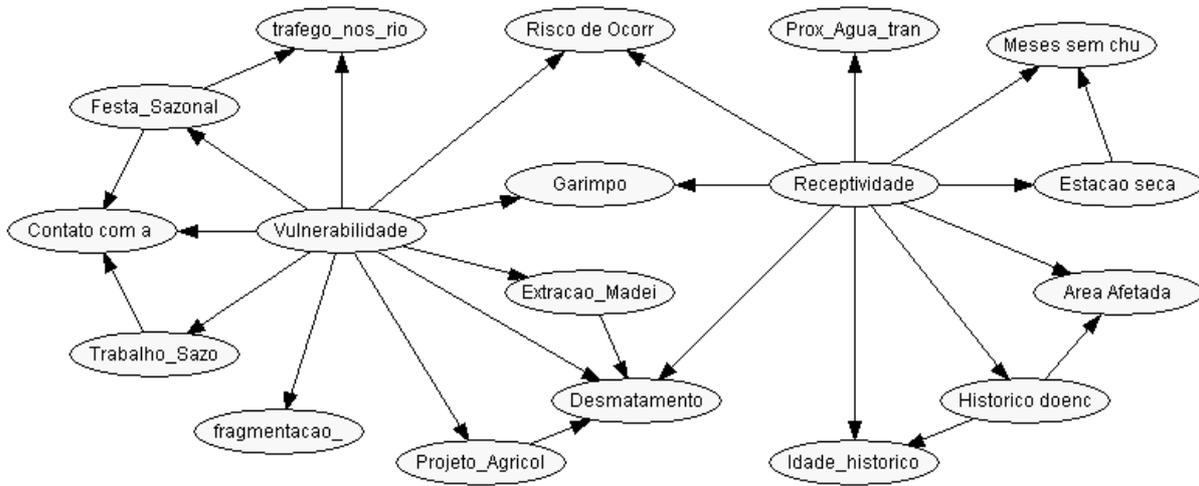


figura 5 - Rede "real" de domínio da saúde

Ressalta-se que o  $\mu$  resultante das operações Bottom-Up e Top-Down são as respectivas crenças da variável (ou o Bel em CPN). Observa-se ainda que ao contrário dos mecanismos tradicionais de propagação de crença em Redes Bayesianas a operação Top-Down não pode ser baseada em:

$$\pi_Y(x) = \alpha \frac{BEL(x)}{\lambda_Y(x)} \quad (\text{Pearl, 1988, pg.168})$$

Pois devido as características da lógica nebulosa, a reversão de operação não é realizável e por tanto devem ser utilizadas as fórmulas aqui dispostas pois estas realizam as operações Bottom-Up e Top-Down com os valores originais de cada variável.

#### 4 Aplicação Em Rede Real

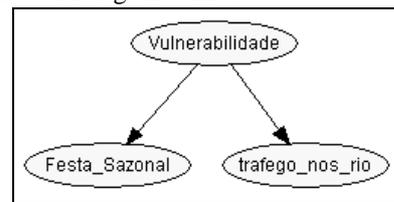
Como aplicação real, buscou-se uma rede validada e que apresenta-se as restrições descritas anteriormente. Em conjunto com o SIVAM, que desenvolve um projeto usando Redes Bayesianas, para o monitoramento de focos de doenças Epidemológicas na região da Amazonia, obteve-se a rede desejada (figura 5). Nesta rede, algumas variáveis sofrem do problema de quantização, necessitando um pré ou pós-processamento para serem utilizadas o que dificulta o uso em tempo real (ex.: fragmentação[não\_fragmentado, fragm., muito fragm.] três estados para o percentual de fragmentação), sofre também, de problemas de atribuição de evidências.(ex.: em "fragmentação" estados nebulosos encaixam-se melhor, pois o percentual de fragmentação tem um índice variando de 0 a 100%, sendo assim, três estados podem não ser suficientemente descritivos à variável crisp porém aumentam sua capacidade caso a variável seja nebulosa).

Nesta rede, deseja-se primordialmente saber o risco de ocorrência de malária de uma dada região. Para tanto, as demais variáveis são inseridas na rede em forma de evidências ou calculadas com dados da rede. Desta forma, pode-se não só extrair o risco

esperado como também o valor esperado de outras variáveis não evidenciadas. A importância de poder extrair dados de uma rede com apenas algumas evidências, tange no fato, de que, pode-se utilizar a rede para estimar uma variável qualquer sem o conhecimento total das demais variáveis do sistema; ou seja, pode-se obter a crença de qualquer uma das variáveis da rede, sem a dependência da posição dos nós que foram evidenciado (independe da direção da conexão).

#### 5 Testes em uma rede simples

Devido o volume de dados da rede real, a rede foi seccionada para facilitar a demonstração dos resultados. Assim, a rede usada na demonstração foi uma rede com três variáveis: "Vulnerabilidade", "Festa Sazonal" e "Tráfego nos Rios".



A primeira variável corresponde a vulnerabilidade da região. *Vulnerabilidade[não vulnerável, vulnerável] leia-se [F,T]*

A variável Festa Sazonal, indica se ocorre uma festa sazonal. *Festa Sazonal[não ocorre, ocorre] leia-se [F,T]*

O Tráfego nos Rios indica se o tráfego no rio é relevante. *Tráfego nos Rios [irrelevante, relevante] leia-se [F,T]*

Tanto "Vulnerabilidade" como "Tráfego nos rios" tem estados insuficientes, evidenciando a facilidade do uso de conjuntos nebulosos. (pois o que é um tráfego relevante? Posso definir uma boa faixa de corte com apenas dois estados?)

Tabela 1 - Crença prévia da vulnerabilidade

Vulnerabilidade	
F	0,7
T	0,3

**Crença prévia** da vulnerabilidade é a possibilidade na qual se acredita que ocorram os estados da vulnerabilidade caso não seja inserida uma evidência nesta variável.

Tabela 2 - Relacionamento entre Vulnerabilidade e Tráfego

Tráfego	Vulnerabilidade	
	F	T
	F	0,7
T	0,3	0,6

Tabela 3 - - Relacionamento entre Vulnerabilidade e Festa Sazonal

Festa	Vulnerabilidade	
	F	T
	F	0,6
T	0,4	0,6

As **tabelas de relacionamento** condicionam as influências entre as variáveis conectadas. (causa→efeito)

### 5.1 Primeira Simulação

Simulando o sistema com seus valores originais; ou seja, sem a entrada de nenhuma evidência:

Resultado obtido na Rede Bayesiana:

	Festa Sazonal	Vulnerabilidade	Tráfego
F	0,54	0,7	0,61
T	0,46	0,3	0,39

Resultado obtido na Rede Nebulosa:

	Festa Sazonal	Vulnerabilidade	Tráfego
F	0,553846	0,6666	0,598425
T	0,446154	0,3333	0,401575

### 5.2 Segunda Simulação

Simulando o sistema com a entrada da evidência “não ocorreu Festa Sazonal” (Festa Sazonal[False=1, True=0]):

Resultado obtido na Rede Bayesiana:

	Festa Sazonal	Vulnerabilidade	Tráfego
F	1	0,7778	0,6333
T	0	0,2222	0,3667

Resultado obtido na Rede Nebulosa:

	Festa Sazonal	Vulnerabilidade	Tráfego
F	1	0,6666	0,6666
T	0	0,3333	0,3333

### 5.3 Terceira Simulação

Caso não realizável em Redes Bayesianas pois deseja-se inserir uma evidência em “Tráfego nos rios” de

[F=0,8 , T=0,2] ou seja, considerando o tráfego mais irrelevante do que relevante:

Resultado não realizável por Redes Bayesianas.

Resultado obtido na Rede Nebulosa:

	Festa Sazonal	Vulnerabilidade	Tráfego
F	1	0,6666	0,8
T	0	0,3333	0,2

Neste último caso, a rede sofreu dos problemas citados anteriormente, e para melhorar a resposta do sistema, a rede original foi modificada pelo SIVAM, com o aumento de estados na variável “Tráfego nos rios” e conseqüentemente o remodelamento dos programas de pré e pós-processamento.

### 5.4 Análise dos resultados

Como esperado os valores obtidos não são os mesmos da Rede Bayesiana, no entanto, eles mantêm relações de grandeza coerentes. Por se tratar de possibilidade e não de probabilidade deve-se encarar o resultado de forma diferente e portanto os valores não devem necessariamente coincidir.

No último exemplo, não pode-se calcular o resultado em Bayes pois a evidência inserida não foi exclusiva a um estado; a teoria e por conseqüência a rede não permite este tipo de cálculo.

## 6 Conclusões

Com este novo modelo de rede, as limitações conhecidas puderam ser retiradas e com a mudança de probabilidade para possibilidade, as qualidades da Lógica Nebulosa foram adicionadas a rede sem originar grande degradação da resposta em comparação com a rede original.

A inclusão de elementos de fuzzyficação/defuzzyficação criaram uma melhoria em relação a utilização em tempo real do sistema; pois, não há mais a necessidade de processar as informações fora do ambiente, melhorando também a integração com o sistema, facilitando a visualização ou alteração dos valores destes módulos.

### Agradecimentos

Os autores gostariam de agradecer o apoio da Fundação Aplicações de Tecnologias Críticas – ATECH que em conjunto com o SIVAM possibilitaram a realização deste projeto.

### Referência Bibliográfica

Finn, J. V. e Olesen, K. G., Andersen, S. K (1990). An Algebra of Bayesian Belief Universes of Knowledge-Based Systems. Networks, Vol 20 637-659. John Wiley&Sons.

- Finn, Jensen V. (1996). An Introduction to Bayesian Networks, Springer.
- Kienitz, Karl H. (1990). A Plausible Approximate Reasoning, Cybernetics and Systems: An International Journal, 21:467-654, Hemisphere Publishing Corporation.
- Mendel, J. M. (1995). Fuzzy Logic Systems for Engeneering: A Tutorial. Proceedings of the IEEE, vol.83, no.3. IEEE.
- Pearl, Judea (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers.
- Watanabe, H. e Symon, J. R (1992). Fuzzy Logic for the Management of Uncertainty. Cap.11. John Wiley & Sons.
- Zadeh, L.A. (1979). A Theory of Approximate Reasoning. Machine Inteligence 9, 9:149-189, Ellis Harwood.
- Zadeh, L.A. (1973). Outline of a New Approach to the Analysis of Complex Systems and Decision Processes, IEEE trans. Syst. Man. Cybern., vol SMC-3, no.1 pp28-400, IEEE.

### **Referência Complementar**

- Hájek, Petr, Godo, Lluís e Esteva, Francesc (1995). Fuzzy Logic and probability. Proceedings of Eleventh Annual Conferene on Uncertainty in Artificial Intelligence (UAI-95).
- Huang, Y. C. and Wang, X. Z. (1999). Application of Fuzzy Causal Networks to Water Treatment Plants. Elsevier Science Ltd.
- Pan, Heping (1998). Fuzzy Bayesian Networks – A General formalism for Representation, Inference and Learning with Hybrid Bayesian Networks. International Journal of Pattern Recognition and Artificial Intelligence.
- Pan, Heping e McMichael, Daniel (1998). Fuzzy Causal Probabilistic Network – a new ideal and practical Inference Engine.
- Sandri, Sandra e Correa, Cláudio (1999). Lógica Nebulosa. V Escola de Redes Neurais: Conselho Nacional de Redes Neurais, pp.c073-c090.