

PUC

Série: Monografias em Ciência da Computação,
No. 28/89

ALGORITMOS PARA COMBINAÇÃO DE EVIDÊNCIAS

Wagner T. Silva
Ruy L. Milidiú

Departamento de Informática

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO
RUA MARQUÊS DE SÃO VICENTE, 225 - CEP-22453
RIO DE JANEIRO - BRASIL

PUC/RJ - DEPARTAMENTO DE INFORMÁTICA

Séries: Monografias em Ciências da Computação, N^o 28/89

Editor: Paulo A. S. Veloso

Novembro, 1989.

ALGORITMOS PARA COMBINAÇÃO DE EVIDÊNCIAS*

Wagner Teixeira da Silva
DI-PUC/RJ e CIC-UnB

Ruy Luiz Milidiú
DI-PUC/RJ e IM-UFRJ

* Este trabalho foi parcialmente financiado pela CAPES e FINEP.

Responsável por publicações:

Rosane Teles Lins Castilho
Assessoria de Biblioteca, Documentação e Informação
FUC R10, Departamento de Informática
Rua Marquês de São Vicente, 225 - Gávea
22453 - Rio de Janeiro, RJ
BRASIL

Tel.: (021) 529-9386
BITNET: user@linc.bitnet

TELEX: 31078

FAX: (021) 274-4546

RESUMO

Este artigo resume os conceitos básicos da teoria de Dempster e Shafer, analisa diferentes modos de representar uma função de crença, e descreve algumas heurísticas para combinar funções de crença. Um algoritmo para combinar uma função de crença bayesiana com uma lista de funções de crença é também descrito. E, finalmente são revistos o Algoritmo de Barnett e o Algoritmo de Shafer & Logan. Uma formulação completa é proposta para este último algoritmo.

PALAVRAS-CHAVE: Teoria de Dempster e Shafer, teoria da evidência, funções de crença, combinação de evidências, raciocínio evidencial, algoritmo.

ABSTRACT

This paper summarizes all Dempster-Shafer Theory key concepts. It also analyses different ways to represent belief functions, and shows some heuristics for belief function combination. An algorithm to combine a bayesian belief function with a list of belief functions is also described. Finally, Barnett's Algorithm and Shafer & Logan Algorithm are reviewed, and some enhancements to this last algorithm are proposed.

KEY-WORDS: Dempster-Shafer theory, theory of evidence, belief functions, combination of evidence, evidential reasoning, algorithm.

SUMÁRIO

	página
1. Introdução	1
2. Teoria de DS	
2.1 Massa Básica de Crença	3
2.2 Função de Crença	6
2.3 Soma Ortogonal	8
2.4 Plausibilidade e Número Comunitário	11
2.5 Tipos de Funções de Crença	13
3. Representação de funções de Crença	
3.1 Massa Básica de Crença	14
3.2 Função de Crença	18
3.3 Função Comunitária	18
3.4 Função de Plausibilidade	19
4. Heurísticas	
4.1 Esforço de Combinação	20
4.2 Usando Heurísticas	22
5. Algoritmos Para combinar evidências	
5.1 Combinando Funções Bayesianas	27
5.2 Algoritmo de Barnett	32
5.3 Algoritmo de Shafer-Logan	35
6. Conclusão	40
Bibliografia	42

1. INTRODUÇÃO

A Teoria da Evidência de Shafer(1976) resultou da reformulação do trabalho de Dempster(1963, 1967a, 1967b, 1967c, 1968a, 1968b, 1969) sobre limites inferior e superior de probabilidades. Esta reformulação foi feita para domínios discretos finitos, comumente encontrados na área de sistemas baseados em conhecimento, e expressa em termos de crença, suporte e evidência. Os conceitos de limites inferior e superior de probabilidade são identificados como graus de crença e graus de plausibilidade, respectivamente. O nome "Teoria da Evidência" é devido ao fato dela lidar com pesos de evidência e graus numéricos de suporte baseados em evidências. Contudo ela é mais conhecida como Teoria de Dempster e Shafer (DS), ainda que outros nomes também lhe sejam atribuído, tais como Função de Crença, Modelo de Transferência de Crença, etc.

O termo crença significa o ato ou efeito de crer; ter fé, convicção, etc. Portanto é algo epistêmico, subjetivo e pode variar de pessoa para pessoa. Alguém pode acreditar muito, pouco ou nada em uma determinada proposição.

O reverendo Thomas Bayes(1702,1761) propôs que a crença obedecesse aos axiomas básicos da probabilidade. Isto é, que a crença fosse dada por uma função $Cb: 2^{\Theta} \rightarrow [0,1]$ e obedecesse as regras:

$$a.1 \quad Cb(\phi) = 0 \quad (1.1)$$

$$a.2 \quad Cb(\Theta) = 1 \quad (1.2)$$

$$a.3 \quad \text{Se } A_1, \dots, A_n \text{ são subconjuntos disjuntos de } \Theta \\ \text{então} \quad (1.3)$$

$$Cb\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n Cb(A_i), \quad A_i \cap A_j = \phi, \quad i \neq j.$$

$$a.4 \quad Cb(A/B) = Cb(A \cap B) / Cb(B) \quad (1.4)$$

Aqui Θ é um conjunto finito e exaustivo de proposições mutuamente exclusivas. A regra a.3 é chamada de regra da adição e a regra a.4 é a regra de condicionamento. Por meio da última pode-se atualizar a crença de alguém face a novas evidências. Na área de inteligência artificial, a maioria das propostas numéricas para lidar com raciocínio impreciso tem enfoque bayesiano. O fato do enfoque bayesiano obedecer aos axiomas da probabilidade lhe

traz todo o instrumental desta teoria, o que lhe é muito vantajoso.

A Teoria da Evidência difere da Teoria Bayesiana. Nela a função de crença $Cr: 2^{\Theta} \rightarrow [0,1]$, obedece as seguintes regras:

$$b.1 \quad Cr(\emptyset) = 0 \quad (1.5)$$

$$b.2 \quad Cr(\Theta) = 1 \quad (1.6)$$

b.3 Se A_1, \dots, A_n São subconjuntos de Θ , então

$$Cr\left(\bigcup_{i=1}^n A_i\right) \geq \sum_{I \subseteq \{1, \dots, n\}} (-1)^{|I|+1} Cr\left(\bigcap_{i \in I} A_i\right) \quad (1.7)$$

Além da ausência da regra de condicionamento, a terceira regra é que traz toda a diferença em relação ao enfoque bayesiano. Por ela temos que $Cr(A) + Cr(\bar{A}) \leq 1$:

Exemplo: Sejam $\Theta = \{a, b, c, d\}$, e a função de crença Cr dada por

$$Cr(A) = \begin{cases} 1/2, & \text{se } A \neq \Theta \text{ e } \{a, b\} \subset A \\ 0, & \text{se } A \neq \Theta \text{ e } \{a, b\} \cap A \neq \{a, b\} \\ 1, & \text{se } A = \Theta \end{cases}$$

Tomando $A = \{a, b\}$, vem que $Cr(\emptyset) = 0$, $Cr(\Theta) = 1$, mas $Cr(A) + Cr(\bar{A}) < 1$.

Este exemplo mostra que há uma massa de crença $m(\Theta) = Cr(\Theta) - Cr(A) - Cr(\bar{A})$ que está no próprio Θ , já que ela não está em A e nem em qualquer subconjunto próprio de Θ . Esta massa de crença é inerente ao próprio Θ e não de seus subconjuntos próprios. Ela representa a nossa ignorância. É a crença que poderia estar em qualquer subconjunto próprio de Θ , mas não se sabe quanto e em quais subconjuntos distribuí-la. Diante disto é alocada em Θ e posteriormente, a vista de mais evidências, ela pode ser transferida para um subconjunto próprio de Θ . Por isto a Teoria da Evidência é também conhecida como "modelo de transferência de crença", Smets(1988).

Na seção 2, apresentamos conceitos gerais, e alguns específicos, da Teoria de DS necessários a explicação dos processos de combinação de evidências tratados neste trabalho.

Na seção 3, analisamos as diversas formas de representação de crença e o custo de conversão de uma representação para outra.

Na seção 4, sugerimos algumas heurísticas que podem minorar o

esforço de combinação de evidências.

Na seção 5, propomos alguns algoritmos para combinação de evidências envolvendo pelo menos uma função bayesiana, e um algoritmo para combinar funções politômicas. Comentamos os algoritmos de Barnett(1981) e o de Shafer & Logan(1987). Como este último foi descrito parcialmente em termos de seis rotinas básicas, nós estamos propondo um algoritmo completo usando as seis rotinas propostas. Este algoritmo, como proposto, tem complexidade de tempo quadrática no fator de ramificação da árvore de hierarquia, e não linear como foi afirmado por Shafer e Logan(1987). Um algoritmo, com projeto baseado no paradigma de orientação a objetos e explorando os conceitos de refinamento e condensação de quadros de discernimento, foi proposto por Silva e Milidiú(1989), apresentando complexidade linear na quantidade de elementos do quadro de discernimento Θ do problema.

Na seção 6, concluimos, comentando as vantagens e desvantagens que, ao nosso ver, a teoria da evidência apresenta e também endereçamos os interessados a tópicos pouco explorados na literatura.

2. TEORIA DE DS

Nesta seção damos uma visão geral dos tópicos da Teoria de DS julgados importantes para a descrição dos algoritmos e heurísticas tratados neste trabalho.

2.1 MASSA BÁSICA DE CRENÇA

O Enfoque de Shafer(1976) enfatiza a representação dos graus de crença com que um corpo de evidências suporta as proposições do quadro de discernimento Θ e o processo de combinação de dois corpos distintos de evidências. Contudo, nada é dito acerca de como quantificar a crença de alguém, nas proposições de Θ , dada uma peça de evidência. Levando em conta isto, e o fato do processo em si ser epistêmico, ficamos livres para usar qualquer artifício que nos aprouver: nossa intuição, ou qualquer ferramental probabilístico. Desde que respeitemos os axiomas básicos da função de crença, podemos usar todo o conhecimento de que disponhamos para quantificar a nossa crença face à evidência disponível.

O conjunto Θ é uma disjunção das suas proposições elementares, e cada subconjunto A de Θ também é uma disjunção das proposições nele contidas. Se $A, B \subset \Theta$, e $A \subset B$, então $A \longrightarrow B$, isto é, a disjunção A implica na disjunção B . Suponha que $a \in \Theta$ seja a proposição verdadeira, então todos os suconjutos de Θ que contém a também são proposições verdadeiras. Entretanto, a única proposição elementar de Θ que é verdadeira é desconhecida.

Podemos representar a nossa incerteza, em relação a proposição verdadeira de Θ , através de uma distribuição de probabilidades de um conjunto aleatório x tomando valores no conjunto das partes de Θ , 2^Θ , definida como

$$\begin{aligned} \Pr[x=\phi] &= 0; \\ 0 \leq \Pr[x=A] &\leq 1, \quad \forall A \subset \Theta \\ \sum_{A \subset \Theta} \Pr[x=A] &= 1 \end{aligned}$$

Essa distribuição de probabilidades é resultante de um mapeamento, epistêmico ou não, do espaço de evidências \mathbb{E} para o conjunto das partes de Θ . No jargão da Teoria de DS, a distribuição de probabilidades do conjunto aleatório x é renomeada por uma função chamada de *probabilidade básica associada*, contudo nós preferimos chama-la de *massa básica de crença* (mbc) por lembrar o caráter epistêmico do seu processo de obtenção. A mbc m associa a cada subconjunto de Θ uma massa de crença que é apenas dele e não dos seus subconjuntos próprios. O valor $m(A)$ é o suporte com que as evidências favorecem a disjunção $A \subset \Theta$. A mbc $m: 2^\Theta \longrightarrow [0,1]$, é definida por

$$m(A) = \Pr[x=A], \quad \forall A \subset \Theta$$

e goza das seguintes propriedades:

- i. $m(\phi) = 0$
- ii. $\sum_{A \subset \Theta} m(A) = 1$

A mbc m pode ser considerada como uma generalização da densidade de probabilidade.

Exemplo: Sejam $\{1,2,3,4,5,6\}$ o espaço amostral associado ao lançamento de um dado D não-honesto, e $\Theta = \{\theta_i \mid 1 \leq i \leq 6\}$, onde cada θ_i é uma proposição do tipo "O resultado do lançamento é i ". Feito o lançamento do dado D , antes de conhecermos o resultado, a nossa

crença na hipótese de que cada proposição de Θ possa ser verdadeira é dada por

$$m(\theta_i) = \begin{cases} 2/9, & \text{se } i \text{ é par} \\ 1/9, & \text{se } i \text{ é ímpar} \end{cases}$$

Neste caso m é uma densidade de probabilidade. Mas se as evidências de que dispomos são tais que só nos permitem avaliar m como

$$m(A) = \begin{cases} 2/3, & \text{se } A = \{\theta_4, \theta_5, \theta_6\} \\ 1/3, & \text{se } A = \Theta \end{cases}$$

e se tomamos $\{\{\theta_1, \theta_2, \theta_3\}, \{\theta_4, \theta_5, \theta_6\}, \Theta\}$ como o espaço de eventos, e a função Cr como a função de distribuição acumulada de m , temos que

$$Cr(\Theta) = 1 \neq m(\{\theta_1, \theta_2, \theta_3\}) + m(\{\theta_4, \theta_5, \theta_6\}) = 2/3$$

Logo m não é um função de densidade, e nem Cr uma distribuição acumulada no sentido convencional. Mas Cr e m estão relacionadas univocamente.

Do exemplo precedente fica clara a idéia de que a mbc m é uma generalização da função de densidade de probabilidade. Na última situação $m(\{\theta_4, \theta_5, \theta_6\}) = 2/3$ e $m(\Theta) = 1/3$, nós estamos representando a nossa ignorância. Isto é, $2/3$ da nossa crença favorece a disjunção $\theta_4 \vee \theta_5 \vee \theta_6$. E, $1/3$ da nossa crença beneficia a disjunção Θ , pois qualquer das proposições elementares pode ocorrer, e nada sabemos que possa favorecer uma particular disjunção, além das especificadas.

Verificamos, do exemplo precedente, que nem todos os subconjuntos de Θ tem massa de crença positiva para uma particular mbc m . Assim, uma maneira compacta de representar uma dada função de crença Cr é através da sua mbc m , guardando apenas os subconjuntos A de Θ que tem $m(A) > 0$. Deste modo, a função de crença Cr pode ser representada por $\mathcal{M} = (\Theta, \mathbb{F}, \mu)$, onde

- Θ é um conjunto finito de proposições mutuamente exclusivas;
- $\mathbb{F} = \{A \mid m(A) > 0\}$ é o conjunto dos elementos focais de Cr , Chamado de conjunto focal; e
- $\mu: \mathbb{F} \rightarrow (0, 1]$ é uma restrição da mbc m ao conjunto \mathbb{F} , isto é, $\mu(A) = m(A)$, $\forall A \in \mathbb{F}$.

O menor subconjunto N de Θ , tal que $Cr(N)=1$, é chamado de núcleo da função de crença Cr e é dado por

$$N = \bigcup_{A \in \mathcal{F}} A$$

2.2 FUNÇÃO DE CRENÇA

Uma função $Cr: 2^{\Theta} \rightarrow [0,1]$ que atribui graus de crença $Cr(A)$, para cada subconjunto A de Θ , é chamada uma função de crença se, e só se, existe um subconjunto aleatório x de Θ , não vazio, tal que

$$Cr(A) = Pr[x \in A], \quad \forall A \subset \Theta$$

Como para cada função de crença Cr existe apenas uma função mbc m , e vice-versa, nada mais natural que possamos, dado o conhecimento de uma, calcular a outra. As fórmulas abaixo estabelecem estas relações:

$$\begin{aligned} Cr(A) &= \sum_{B \subset A} Pr[x \in B] \\ &= \sum_{B \subset A} m(B), \quad \forall A \subset \Theta \end{aligned} \quad (2.1)$$

$$m(A) = \sum_{B \subset A} (-1)^{|A-B|} Cr(B), \quad \forall A \subset \Theta \quad (2.2)$$

A fórmula 2.1 nos diz que a crença em um dado subconjunto A de Θ é obtida pela soma de toda a mbc distribuída entre os subconjuntos próprios de A , mais a mbc atribuída ao próprio A . Ou seja, o nosso grau de crença em A é igual a soma de toda a massa de crença que foi distribuída entre os subconjuntos de A .

Exemplo: Sejam $\Theta = \{a,b,c\}$ e $A = \{a,b\}$. A tabela 2.1 ilustra os valores da mbc m e da sua correspondente função de crença Cr em cada subconjunto de Θ .

	m	Cr
a	1/6	1/6
b	2/6	2/6
A	3/6	1
c	0	0
Θ	0	1

Tabela 2.1

No exemplo precedente, os únicos subconjuntos de Θ que possuem massa são $\{a\}$, $\{b\}$ e $A = \{a,b\}$. A estes subconjuntos dá-se o nome de

elementos focais. A união dos elementos focais de uma dada função de crença constituem o seu núcleo Ω , e a crença $Cr(\Omega)$ é igual a 1. Na função de crença dada pela tabela 2.1, a proposição A é também o núcleo de Cr .

Para calcular a crença usando a fórmula 2.1 é necessário, e suficiente, considerar apenas os elementos focais que sejam subconjuntos do conjunto para o qual se quer quantificar a crença Cr . Da tabela 2.1 temos $Cr(A) = m(A) + m(a) + m(b) = 1$.

A fórmula 2.2 é muito útil em situações em que se tem apenas os valores da função Cr , porém deseja-se a função m . Ainda da tabela 2.1 temos:

$$\begin{aligned} m(a) &= Cr(a) = 1/6 \\ m(b) &= Cr(b) = 2/6 \\ m(A) &= Cr(A) - Cr(a) - Cr(b) = 1 - 1/6 - 2/6 \\ &= 3/6 \end{aligned}$$

A idéia da fórmula é tomar a crença do conjunto A e subtrair as massas de crenças dos subconjuntos próprios de A e aí se obtém a massa de crença que é específica de A .

Voltando aos axiomas da função de crença, fórmulas 1.5 a 1.7, podemos verificar que a função Cr dada pela tabela 2.1 os atende. Vejamos:

1. $Cr(\phi) = 0$
2. $Cr(\Theta) = m(A) + m(a) + m(b)$
 $3/6 + 1/6 + 2/6 = 1$
3. Seja a coleção A , $\{a\}$ e $\{b\}$, então
 $Cr(A \cup \{a\} \cup \{b\}) \geq Cr(A) + Cr(a) + Cr(b) - Cr(a) - Cr(b)$
 $Cr(A) \geq Cr(A)$
 Seja agora a coleção $\{a\}$ e $\{b\}$, então
 $Cr(\{a\} \cup \{b\}) \geq Cr(a) + Cr(b)$
 $Cr(A) \geq Cr(a) + Cr(b)$
 $1 \geq Cr(a) + Cr(b)$
 $1 \geq 1/6 + 2/6 = 1/2$

Mais uma vez constatamos que a terceira regra é que permite a representação da ignorância, pela introdução da desigualdade. Ela é que efetivamente torna a Teoria de DS diferente do enfoque Bayesianista.

2.3 SOMA ORTOGONAL

Sejam x_1, x_2 subconjuntos aleatórios não vazios do quadro de discernimento Θ . Suponha que x_1 e x_2 sejam probabilisticamente independentes, e que $\Pr[x_1 \cap x_2 \neq \emptyset] > 0$. Sejam Cr_1 e Cr_2 funções de crença com mbc m_1 e m_2 , respectivamente, dadas por

$$m_1(A_1) = \Pr[x_1 = A_1], \quad \forall A_1 \subset \Theta, \text{ e}$$

$$m_2(A_2) = \Pr[x_2 = A_2], \quad \forall A_2 \subset \Theta.$$

Então a crença $Cr = Cr_1 \oplus Cr_2$, resultante da combinação de Cr_1 e Cr_2 , tem sua mbc m dada por

$$m(A) = \Pr[x_1 \cap x_2 = A \mid x_1 \cap x_2 \neq \emptyset]$$

A combinação de duas funções de crença pode ser formulada em termos evidenciais do seguinte modo. Sejam Cr_1 e Cr_2 duas funções de crença, sobre o quadro de discernimento Θ , representadas por \mathcal{M}_1 e \mathcal{M}_2 , respectivamente. Se

$$K^{-1} = \sum_{\substack{A_1 \cap A_2 \neq \emptyset \\ A_1 \in \mathbb{F}_1 \text{ e } A_2 \in \mathbb{F}_2}} \mu_1(A_1) \cdot \mu_2(A_2) > 0$$

Então

$$m(A) = K \sum_{\substack{A_1 \cap A_2 = A \\ A_1 \in \mathbb{F}_1 \text{ e } A_2 \in \mathbb{F}_2}} \mu_1(A_1) \cdot \mu_2(A_2), \quad \forall A \subset \Theta, A \neq \emptyset$$

Onde m é a função de mbc associada a Cr , e $Cr = Cr_1 \oplus Cr_2$ é a soma ortogonal de Cr_1 e Cr_2 , cuja representação é dada por $\mathcal{M} = \{\Theta, \mathbb{F}, \mu\}$. A soma ortogonal de Cr_1 e Cr_2 está definida somente se o fator de renormalização $K(\infty$, isto é, $K^{-1} > 0$). Shafer (1976) define o peso do conflito entre Cr_1 e Cr_2 como sendo o $\log k$, e o denota por $\text{Con}(Cr_1, Cr_2)$. Este conflito é zero quando K é igual a 1, e infinito (∞) quando K é infinito. Neste último caso Cr_1 e Cr_2 não são combináveis.

A regra para formação da soma ortogonal é chamada de regra de Dempster. A soma ortogonal tem um papel similar ao da quarta regra de condicionamento de Bayes, formula 1.4. Através dela podemos atualizar nossa crença a vista de novas evidências. O nome ortogonal lembra independência. O pressuposto básico para se usar esta regra de combinação é que as crenças a serem combinadas sejam

suportadas por corpos distintos de evidências, sem o que, o processo de combinação leva a enviesamento. Isto é garantido pela independência dos conjuntos aleatórios x_1 e x_2 . O exemplo a seguir ilustra o processo de combinação de evidências.

Exemplo: Suponhamos que em uma dada região existam apenas 3 tribos indígenas: Selva, Dia e Sá. Estas tribos são arredias. Inicialmente nada se conhece sobre as características e quantidade de indivíduos de cada tribo. Mas sabe-se que há índios que são antropófagos e outros não. Uma pesquisa foi realizada para se conhecer mais sobre estas populações. Após a pesquisa passou-se a conhecer as características dominantes dos membros de cada tribo. Inclusive identificou-se a tribo Selva como sendo a única que pratica a antropofagia.

Um índio foi visto nas cercanias de uma fazenda. Devido a distância nenhuma das suas características foram observadas. A pergunta que se formula é "a qual tribo o índio pertence?". O conjunto de respostas possíveis é dado pelo quadro de discernimento $\Theta = \{\text{"pertence a tribo } x" \mid x \in \{\text{Selva, Dia, Sá}\}\}$. Como há uma correspondência uma a um entre o conjunto Θ e o conjunto $\{\text{Selva, Dia, Sá}\}$, nós usaremos este último por facilidade de notação. A função de crença que representa o estado de ignorância inicial e sua mbc são dadas abaixo:

$$m(\Theta) = 1, \text{ e } Cr(\Theta) = 1$$

Isto significa que a crença de que o índio pertença a uma particular tribo é zero. Mas seguramente ele pertence a uma das tribos.

Posteriormente as seguintes características foram observadas: *alto, alourado e olhos pretos*. As características são evidências que levam as seguintes crenças:

Alto: A evidência altura nos dá a seguinte crença

$$Cr_1(\{\text{selva, Dia}\}) = 0.8, \text{ e } Cr_1(\Theta) = 1$$

E a função de mbc m_1 correspondente é

$$m_1(\{\text{Selva, Dia}\}) = 0.8 \text{ e } m_1(\Theta) = 0.2$$

Alourado: A evidência Cor dos cabelos nos dá a seguinte distribuição de mbc:

$$m_2(\{\text{Dia, Sá}\}) = 0.7 \text{ e } m_2(\Theta) = 0.3$$

Olhos pretos: A cor dos olhos nos dá a seguinte distribuição de mbc:

$$m_3(\{Selva, Sá\}) = 0.9 \quad \text{e} \quad m_3(\ominus) = 0.1$$

Como combinar estas evidências de modo a atualizar a nossa crença inicial? A combinação se dá com duas funções por vez, a ordem é irrelevante, pois a soma ortogonal é associativa e comutativa. A figura 2.1 ilustra a combinação da função de crença vacuosa Cr_0 com a função de crença Cr_1 , dada pela evidência *Altura*.

$m_1 \backslash m_0$	1
	\ominus
0.8	0.8
$\langle Selva, Dia \rangle$	$\langle Selva, Dia \rangle$
0.2	0.2
\ominus	\ominus

Figura 2.1 $Cr_0 \oplus Cr_1$

Nos combinamos as crenças Cr_0 e Cr_1 , via soma ortogonal, fazendo a intersecção de cada elemento focal de Cr_0 com todos os elementos focais de Cr_1 e a mbc dos elementos da intersecção é obtida pelo produto das mbc dos elementos intercedidos. Nos indicamos esta combinação pela fórmula $Cr_{01} = Cr_0 \oplus Cr_1$. A nossa crença face a evidência *Altura* é dada agora por Cr_{01} , cuja mbc é dada por

$$m_{01}(\{Selva, Dia\}) = 0.8 \quad \text{e} \quad m_{01}(\ominus) = 0.2$$

Com este resultado combinamos a evidência *Alourado*, que gerou a função de crença Cr_2 , cujas operações são ilustradas na figura 2.2. Esta combinação gera a função de crença Cr_{012} e sua função de mbc, m_{012} . Combinando, agora, a crença Cr

$m_2 \backslash m_{01}$	0.8	0.2
	$\langle Selva, Dia \rangle$	\ominus
0.7	0.56	0.14
$\langle Dia, Sa \rangle$	$\langle Dia \rangle$	$\langle Dia, Sa \rangle$
0.3	0.24	0.06
\ominus	$\langle Selva, Dia \rangle$	\ominus

Figura 2.2 $Cr_{01} \oplus Cr_2$

$m_3 \setminus m_{012}$	0.24 <Selva, Dia>	0.14 <Dia, Sa>	0.56 <Dia>	0.06 \emptyset
0.9 <Selva, Sa>	0.216 <Selva>	0.126 <Sa>	0.504 ϕ	0.054 <Selva, Sa>
0.1 \emptyset	0.024 <Selva, Dia>	0.014 <Dia, Sa>	0.056 <Dia>	0.006 \emptyset

Figura 2.3 $Cr_{012} \oplus Cr_3$

Nesta última combinação apareceu uma dificuldade: a intersecção de {Selva, Sa} e {Dia} é vazia, e há uma massa associada ao conjunto vazio. Mas uma das propriedades da mbc m é que $m(\phi)=0$. Quando isto acontece nos ignoramos o vazio e renormalizamos a função de mbc de modo a que ela continue atendendo aos seus axiomas, conforme prevê a soma ortogonal. Assim a função de crença resultante é dada pela tabela 2.2.

	m	Cr
<Selva>	0.436	0.436
<Dia>	0.113	0.113
<Sa>	0.254	0.254
<Selva, Sa>	0.109	0.799
<Selva, Dia>	0.048	0.597
<Dia, Sa>	0.028	0.395
\emptyset	0.012	1

Tabela 2.2 $Cr_0 \oplus Cr_1 \oplus Cr_2 \oplus Cr_3$

2.4 PLAUSIBILIDADE E NÚMERO COMUNITÁRIO

Duas outras importantes medidas, assim como a função de crença, são macropropriedades da função de massa básica de crença, mbc. A

$$\begin{aligned}
 Pl(A) &= 1 - Cr(\bar{A}) \\
 &= \sum_{B \subset \emptyset} m(B) - \sum_{B \subset A} m(B) \\
 &= \sum_{B \mid A \neq \emptyset} m(B), \quad \forall A \subset \emptyset \quad (2.3)
 \end{aligned}$$

logo, de 2.1 e 2.3, temos:

$$Cr(A) \leq Pl(A)$$

primeira é a *plausibilidade*, que é identificada no trabalho de Dempster como limite superior de probabilidade. Com esta medida nos podemos estabelecer um intervalo de crença para cada proposição ou subconjunto de proposições. Ela se relaciona com a função de crença via a fórmula (2.3).

Assim a crença em uma dada proposição A pode ser melhor caracterizada pelo intervalo $[Cr(A), Pl(A)]$. A plausibilidade pode ser interpretada como sendo a extensão do erro que se comete por refutar A . Conforme Pearl(1988), $Pl(A)$ representa a probabilidade de A ser compatível com o corpo de evidências que suporta a função de crença Cr . Isto é, a proposição A não pode ser negada e é possível com chance $Pl(A)$. O intervalo

$$Pl(A) - Cr(A) = 1 - [Cr(A) + Cr(\bar{A})] \geq 0$$

representa a probabilidade de ambas as proposições A e \bar{A} serem compatíveis com as evidências que suportam Cr .

Na tabela 2.3 reproduzimos a mbc e suas macropropriedades para o problema das tribos, exemplo da seção 2.3 e extensão da figura 2.2.

	m	Cr	Pl	Q
<Selva>	0.436	0.436	0.605	0.605
<Dia>	0.113	0.113	0.201	0.201
<Sa>	0.254	0.254	0.403	0.403
<Selva,Sa>	0.109	0.799	0.887	0.121
<Selva,Dia>	0.048	0.597	0.746	0.060
<Dia,Sa>	0.028	0.395	0.564	0.040
⊕	0.012	1	1	0.012

Tabela 2.3 Macro propriedades de Cr 0123

A segunda medida de crença, macropropriedade de mbc, é a *função comunitária*, representada pela letra Q . Ela aparece na tabela 2.3. Esta função mede toda a massa de crença que está livre para migrar para os elementos de uma dada proposição. Na tabela 2.3 $Q(\{Selva,Sa\})$ diz que há uma massa de crença igual a 0.121 livre para migrar para os elementos Selva ou Sá, dependendo das novas evidências que possam surgir.

2.5 TIPOS DE FUNÇÕES DE CRENÇA

A seguir são definidos alguns tipos de funções de crença que apresentam propriedades relevantes para a construção de algoritmos de combinação de evidências. Todos os exemplos serão dados por funções de crença Cr , com mbc m , definidas em $\Theta = \{a, b, c\}$.

Função de suporte

Sejam Cr uma função de crença, m sua mbc, \mathbb{F} seu conjunto focal, e \mathbb{N} o seu núcleo. Se $\mathbb{N} = \mathbb{F}$, isto é, $m(\mathbb{N}) > 0$, então Cr é chamada de função de suporte.

Exemplo: A função Cr com $m(\{a\}) = 1/4$, $m(\{a, b\}) = 2/4$ e $m(\Theta) = 1/4$.

Função vacuosa

Sejam Cr uma função de suporte, e \mathbb{F} seu conjunto focal. Se $\mathbb{F} = \{\Theta\}$, isto é, o único elemento focal de Cr é Θ , então Cr é uma função de crença Vacuosa.

Exemplo: A função Cr com $m(\Theta) = 1$.

Função de suporte simples (fss)

Seja Cr uma função de suporte, e \mathbb{F} seu conjunto focal. Se $\mathbb{F} \subset \{A, \Theta\}$, isto é, Cr tem no máximo um elemento focal $A \neq \Theta$, e $A \neq \phi$, então Cr é chamada de função de suporte simples (fss). O elemento focal A é conhecido como o foco de Cr .

Exemplo: A função Cr com $m(\{b, c\}) = 2/3$, e $m(\Theta) = 1/3$.

Função dicotômica

Sejam Cr uma função de crença, e \mathbb{F} seu conjunto focal. Se $\mathbb{F} \subset \{A, \bar{A}, \Theta\}$, Isto é, qualquer elemento focal de Cr é um membro de $\{A, \bar{A}, \Theta\}$, então Cr é Chamada de função de crença dicotômica com dicotomia $\{A, \bar{A}\}$. As funções vacuosas e de suporte simples são casos particulares desta função.

Exemplo: A função Cr com $m(\{a, b\}) = 1/4$, $m(\{c\}) = 2/4$ e $m(\Theta) = 1/4$.

Função de crença politômica

Sejam Cr uma função de crença, \mathbb{F} seu conjunto focal, e $\Omega = \{A_1, A_2, \dots, A_f\}$ uma partição de Θ . Se $\mathbb{F} \subset \{A_1, A_2, \dots, A_f, \Theta\}$, então Cr é dita uma função de crença politômica definida em Ω . As funções de crença vacuosas, de suporte simples e dicotômica são casos particulares desta função.

Exemplo: A função Cr com $m(\{a\}) = 2/5$, $m(\{b, c\}) = 1/5$ e $m(\Theta) = 2/5$.

Função de crença bayesiana

Seja Cr uma função de crença na representação $\mathcal{M} = (\Theta, \mathbb{F}, \mu)$. A função de crença Cr é dita ser bayesiana se para cada $A \in \mathbb{F}$, $A \in \Theta$. Isto é, todos os elementos focais de Cr são atômicos. Outra maneira de identificar uma função bayesiana é ver se ela tem as três propriedades dadas por (1.1), (1.2) e (1.3).

Se Cr é bayesiana, então todas as asserções a seguir são equivalentes:

1. Cr é bayesiana
2. Todos os elementos focais de Cr são atômicos
($\forall A \subset \Theta, m(A) > 0 \rightarrow |A|=1$)
3. $Cr(A) = P(A), \forall A \subset \Theta$
4. $Cr(A) + Cr(\bar{A}) = 1, \forall A \subset \Theta$

Definição: Uma função Cr é bayesiana se, e somente se, existe uma função $p: \Theta \rightarrow [0,1]$ tal que

$$\sum_{\theta \in \Theta} p(\theta) = 1 \quad \text{e} \quad Cr(A) = \sum_{\theta \in A} p(\theta), \quad \forall A \subset \Theta$$

A função de crença Bayesiana atende aos axiomas 1.1 a 1.4. O esquema de atualização de crença usa a regra 1.4. Assim, se temos Θ , um quadro de discernimento, e uma distribuição de probabilidade p definida em Θ , tal que $0 \leq p(x) \leq 1, \forall x \in \Theta$, e

$$\sum_{x \in \Theta} p(x) = 1;$$

e sabemos que a resposta a nossa pergunta está num certo $B \subset \Theta$, então devemos atualizar nossa crença em cada elemento de Θ dado o conhecimento da ocorrência de B . Pelo teorema de Bayes temos:

$$p(x/B) = p(\{x\} \cap B) / p(B), \quad \forall x \in \Theta$$

Isto equivale a atribuir probabilidade zero para os elementos fora de B , e renormalizar os elementos em B pelo fator

$$p(B) = \sum_{x \in B} p(x)$$

O exemplo a seguir mostra que a soma ortogonal é equivalente a regra de condicionamento de Bayes, quando a função de crença Cr é bayesiana.

Exemplo: Seja Cr_1 uma função de crença bayesiana definida em $\Theta = \{a, b, c\}$, com mbc p dada por

$$p(a)=1/3, p(b)=1/3, \text{ e } p(c)=1/3$$

Uma nova evidência surge dando origem a função de crença Cr_2 , com mbc p_2 , dada por

$$P_2(\{a, b\}) = 1$$

A crença atualizada face a nova evidência é dada no caso (a) pela aplicação da regra de condicionamento (1.4), e no caso (b) pela soma ortogonal, e o resultado em ambos os casos são idênticos.

(a) Aplicando a regra de condicionamento

$$p(a/\{a, b\}) = p(a)/p(\{a, b\}) = 1/3/2/3 = 1/2;$$

$$p(b) = 1/2; \text{ e}$$

$$p(c) = 0..$$

(b) Aplicando a soma ortogonal

$P_2 \backslash p$		1/3	1/3	1/3
		a	b	c
1		1/3	1/3	1/3
$\{a, b\}$		a	b	ϕ

normalizando vem

$$p(a) = 1/2, p(b)=1/2 \text{ e } p(c)=0$$

Mas quando uma função de crença tem elementos focais não atômicos ela deixa de ser bayesiana, e a correspondente função de mbc não é mais uma densidade de probabilidade. Por exemplo, se $\Theta = \{a, b\}$, e $m(\Theta) = 1/4$, $m(a) = 2/4$ e $m(b) = 1/4$, obviamente isto não é uma densidade de Probabilidade pois $Pr(\Theta) = m(a) + m(b) \neq 1$. Contudo m é uma autêntica função de massa básica de crença.

3. REPRESENTAÇÃO DE FUNÇÕES DE CRENÇAS

Há diversas formas de representar uma função de crença, todas conversíveis entre si. A seguir detalhamos as representações e o esforço de conversão entre elas.

Na análise de complexidade, nesta e em outras seções, levamos em conta que a lista das proposições elementares de Θ , $\theta_1, \dots, \theta_n$, estão em correspondência com os naturais de 1 a n. E as operações sobre os subconjuntos de Θ serão em termos destes números e de um vetor V com n posições. Assim a intersecção união e continência têm complexidade de tempo $O(|\Theta|)$.

3.1 MASSA BÁSICA DE CRENÇA

A representação de uma função de crença, em termos dos seus elementos focais e da mbc a eles associada, é dada por $\mathcal{M} = (\Theta, \mathbb{F}, \mu)$ e está especificada na seção 2.1. Via de regra, tem-se poucos elementos focais para uma dada função de crença. As fórmulas de conversão de μ para a função de crença Cr , função de plausibilidade Pl , e função comunitária Q , são dadas a seguir.

$$Cr(A) = \sum_{\substack{B \subseteq A, \\ B \in \mathbb{F}}} \mu(B), \quad \forall A \subseteq \Theta \quad (3.1)$$

$$Pl(A) = \sum_{\substack{A \cap B \neq \emptyset, \\ B \in \mathbb{F}}} \mu(B), \quad \forall A \subseteq \Theta \quad (3.2)$$

$$Q(A) = \sum_{\substack{B \subseteq \mathbb{F}, \\ A \subseteq B}} \mu(B), \quad \forall A \subseteq \Theta \quad (3.3)$$

O esforço de conversão, apenas para uma dada proposição $A \subseteq \Theta$, é estudado abaixo:

Crença

A complexidade de tempo para se obter $Cr(A)$, a partir de \mathcal{M} , é $O(|\mathbb{F}| \cdot |A|)$, se $|A| > 1$ e é igual a 1, se $|A| = 1$.

Prova: Supondo $|A| > 1$, então (a) O número de somas na fórmula 3.1. é dado por

$$k \cdot \mathcal{S}_c = k \cdot |\{x \in \mathbb{F} \mid x \subseteq A\}| \leq k \cdot |\mathbb{F}|$$

(b) O esforço para se obter os \mathcal{S}_c elementos focais, contidos em A é

dado por

$$k \cdot \mathbb{D}_c = k(|\Theta| + |A| + |\mathbb{D}| |A| + |F|) \leq k(|\Theta| + |A| + |F| |A| + |F|)$$

Onde

- $\mathbb{D} = \{x \in F \mid |x| \leq |A|\}$ é o conjunto de elementos focais com cardinalidade menor ou igual a $|A|$;
- $|\Theta| + |A| + |\mathbb{D}| |A|$ é o número de comparações e atribuições necessários para se obter os \mathbb{S}_c elementos focais contidos em A , já que apenas os $x \in \mathbb{D}$ são verificados se estão ou não contidos em A ;
- São necessários $|F|$ comparações para se determinar os elementos do conjunto \mathbb{D} ; e,
- k é o esforço maior entre as operações de *soma*, *comparação* e *atribuição* de inteiros.

Assim de (a) e (b) temos $k(\mathbb{S}_c + \mathbb{D}_c) \leq 2k|F| + k|F| |A| + k|\Theta| + k|A|$ o que dá uma complexidade $O(|F| \cdot |A|)$.

Plausibilidade

A complexidade para obtenção de $P1(A)$, a partir de \mathcal{M} , é $O(|F| \cdot |\Theta|)$.

Prova: (a) O número de somas na fórmula 3.2, é dado por

$$k|\{x \in F \mid x \cap A \neq \emptyset\}| \leq k|F|$$

(b) O esforço necessário para se obter os elementos focais com intersecções não vazias com A é dado por

$$k(|\Theta| + |A| + \sum_{x \in F} |x|) \leq k(|\Theta| + |A| + |F| |\Theta|)$$

Assim, de (a) e (b) vem o esforço $k(|F| + |\Theta| + |A| + |F| |\Theta|)$ o que implica em uma complexidade $O(|F| \cdot |\Theta|)$.

Comunitário

A complexidade para se obter o número comunitário $Q(A)$ para uma dada proposição A , a partir de \mathcal{M} , é $O(|F| \cdot |\Theta|)$.

Prova: (a) O número de somas na fórmula 3.3 é dado por

$$k|\mathbb{D}| = k|\{x \in F \mid A \subset x\}| \leq k|F|$$

(b) O esforço necessário para se obter os elementos focais que contêm A é dado por

$$k(|\Theta| + |A| + \sum_{x \in \mathbb{D}} (|x| + |A|)) \leq k(|\Theta| + |A| + |F| (|\Theta| + |A|))$$

Assim, de (a) e (b) temos um esforço de $k(|\Theta| + |A| +$

$|F|(|\Theta|+|A|)+|F|)$ que implica em complexidade $O(|F| \cdot |\Theta|)$.

Ainda que não pareça razoável, podem haver aplicações em que se façam necessário as medidas acima para cada subconjunto de Θ . Neste caso, temos que a complexidade é $O(|F| \cdot |\Theta| \cdot |2^\Theta|)$.

3.2. FUNÇÃO DE CRENÇA

A representação de crença, pela função de crença Cr , é dada por $\mathcal{E} = (\Theta, \mathbb{C}, \zeta)$, onde

- Θ é o quadro de discernimento do problema.
- $\mathbb{C} = \{A \subset \Theta \mid Cr(A) > 0\}$, isto é, \mathbb{C} é o conjunto de todos os subconjuntos de Θ com crença maior que zero.
- $\zeta: \mathbb{C} \rightarrow (0, 1]$ é uma restrição da função de crença Cr ao conjunto \mathbb{C} , isto é, $\zeta(A) = Cr(A)$, $\forall A \in \mathbb{C}$.

Podemos afirmar que a representação \mathcal{M} é mais compacta que a representação \mathcal{E} , pois todo elemento focal tem crença positiva e a união de quaisquer dois elementos focais também tem crença positiva. E mais, qualquer subconjunto de Θ , que contenha pelo menos um elemento focal, também, tem crença positiva, logo $|F| \leq |\mathbb{C}|$.

As conversões desta representação para as outras medidas são dadas pelas fórmulas abaixo.

$$m(A) = \sum_{\substack{B \subset A, \\ B \in \mathbb{C}}} (-1)^{|A-B|} \zeta(B), \quad \forall A \subset \Theta$$

$$Pl(A) = 1 - \zeta(\bar{A}), \quad \forall A \subset \Theta \text{ e } \bar{A} \in \mathbb{C}$$

$$Q(A) = \sum_{\substack{B \subset A, \\ \bar{B} \in \mathbb{C}}} (-1)^{|B|} \zeta(\bar{B}), \quad \forall A \subset \Theta$$

A complexidade de conversão para a mbc $m(A)$ ou para número comunitário $Q(A)$ é $O(|\mathbb{C}| \cdot |\Theta|)$. Enquanto que para calcular $Pl(A)$ o esforço é constante. As provas pode ser dada de maneira similar àquelas das conversões da representação \mathcal{M} .

3.3 FUNÇÃO COMUNITÁRIA

A representação da crença usando a função comunitária é dada pela tripla $\mathcal{Q} = (\Theta, \mathbb{Q}, \vartheta)$. Onde

- Θ é o quadro de discernimento

- $\mathcal{Q} = \{x \subset \Theta \mid Q(x) > 0, x \neq \emptyset\}$ é o conjunto de todos os subconjuntos de Θ com número comunitário maior que zero.
- $\mathcal{Q}: \mathcal{Q} \rightarrow (0, 1]$, é uma restrição de Q ao conjunto \mathcal{Q} , isto é, $\mathcal{Q}(A) = Q(A), \forall A \in \mathcal{Q}$.

Da representação Q podemos converter para as outras medidas conforme as fórmulas abaixo:

$$Cr(A) = \sum_{\substack{B \subset A, \\ B \in \mathcal{Q}}} (-1)^{|B|} \mathcal{Q}(B), \quad \forall A \subset \Theta$$

$$Pl(A) = \sum_{\substack{B \subset A, \\ B \neq \emptyset, \\ B \in \mathcal{Q}}} (-1)^{|B|+1} \mathcal{Q}(B), \quad \forall A \subset \Theta$$

Nesta representação o esforço de conversão, para cada $A \subset \Theta$, é $O(|\mathcal{Q}| \cdot |\Theta|)$, tanto para $Cr(A)$ quanto para $Pl(A)$, respectivamente.

3.4 FUNÇÃO DE PLAUSIBILIDADE

A representação de crença usando a plausibilidade é dada por $\mathcal{P} = (\Theta, \mathcal{P}, \rho)$, onde

- Θ é o quadro de discernimento,
- $\mathcal{P} = \{x \subset \Theta \mid Pl(x) > 0\}$ é o conjunto de todos os subconjuntos de Θ que têm plausibilidade maior que zero.
- $\rho: \mathcal{P} \rightarrow (0, 1]$, é uma restrição de Pl ao conjunto \mathcal{P} , isto é, $\rho(A) = Pl(A), \forall A \in \mathcal{P}$.

Da representação \mathcal{P} podemos calcular as outras medidas pelas fórmulas abaixo:

$$Cr(A) = 1 - \rho(\bar{A}), \quad \forall A \subset \Theta$$

$$Q(A) = \sum_{\substack{B \subset A, \\ B \in \mathcal{P}}} (-1)^{|B|+1} \rho(B), \quad \forall A \subset \Theta$$

A complexidade para se determinar $Cr(A)$, é constante, e para se determinar $Q(A)$ é $O(|\mathcal{Q}| \cdot |\Theta|)$.

O esforço de conversão de uma dada representação para uma medida de outra representação, apenas para uma dada proposição, tem complexidade, via regra, igual $O(|\Theta| \cdot |\lambda|)$, onde $\lambda \in \{F, C, Q, P\}$,

dependendo da atual representação. Entretanto, para atualizar a crença via soma ortogonal é necessária a representação \mathcal{M} , ou então através da função comunitária. Portanto se a representação está em termos de \mathcal{M} ou \mathcal{Q} nenhuma conversão se faz necessário para fins de atualização da crença. Contudo a representação \mathcal{M} é a mais recomendada, por ser a mais compacta.

4. HEURÍSTICAS

4.1 ESFORÇO DE COMBINAÇÃO

Sejam Cr_1 e Cr_2 duas funções de crença, F_1 e F_2 , respectivamente, seus conjuntos focais. Sejam $n=|F_1|$ e $k=|F_2|$, então a soma ortogonal $Cr_1 \oplus Cr_2$ tem $n.k$ intersecções, tem no máximo $2n.k$ produtos e $n.k$ somas, em função dos vazios possivelmente obtidos nas intersecções. Supondo $n=k$, então o esforço de somas e produtos é da ordem de n^2 e de intersecções é da ordem de $n^2 |\Theta|$. O que leva a combinação de duas funções de crenças ter complexidade $O(n^2 |\Theta|)$. Como consequência, se combinarmos várias funções de crença, Cr_1, \dots, Cr_k , durante o processo de combinação pode haver uma fragmentação de elementos focais muito grande, em função das características das funções de crenças envolvidas. A soma ortogonal $Cr_1 \oplus \dots \oplus Cr_k$, onde cada Cr_i tem um número de elementos focais $N_i = |F_i|$, pode chegar ao máximo de $\min \left(\prod_{i=1}^k N_i, |2^\Theta| \right)$ elementos focais. Supondo $N_i = N, \forall i$, então este número de elementos focais, $N^k (\leq |2^\Theta|)$ pode crescer exponencialmente com o número de funções combinadas. Por exemplo, todas as funções de crenças, geradas pelas evidências disponíveis, do problema das tribos, seção 2, tinha cada uma apenas 2 elementos focais. Entretanto a combinação delas resultou em 7 elementos focais. Isto é, todos os subconjuntos do conjunto das partes de Θ é um elemento focal, exceto ϕ . Por isto o processo de combinação de crenças tem, em geral, complexidade exponencial.

A ordem em que combinamos as funções de crença não altera o resultado final, pois a soma ortogonal é associativa e comutativa. Assim, se temos a soma $Cr_1 \oplus Cr_2 \oplus \dots \oplus Cr_n$, onde cada Cr_i é uma função de crença, podemos fazer uma permutação qualquer entre as funções Cr_i e combiná-las que o resultado final será o mesmo. Entretanto, deve haver uma ordem em que a fragmentação dos elementos focais,

durante os passos de combinação das funções Cr_i com as funções resultantes intermediárias, seja minimizado, trazendo uma redução do esforço total.

Exemplo: As combinações ilustradas nas figuras 4.1, 4.2 e 4.3 são três ordens diferentes de combinar as funções Cr_1 , Cr_2 e Cr_3 , definidas em $\Theta = \{a, b, c\}$. A ordem $(Cr_1 \oplus Cr_3) \oplus Cr_2$, ilustrada na figura 4.1, tem 22 atribuições/comparações, 4 produtos, e 1 soma. O que lhe dá um esforço menor ou igual a $23c$. Onde c é o maior esforço dentre as operações elementares de soma, atribuição, comparação e produto (de inteiros). A ordem $(Cr_2 \oplus Cr_3) \oplus Cr_1$, ilustrada na figura 4.2, tem 24 atribuições/comparações, 6 produtos, e 2 somas. O que lhe dá um esforço menor ou igual a $32c$. A ordem de combinação $(Cr_1 \oplus Cr_2) \oplus Cr_3$, ilustrada na figura 4.3, tem esforços 29 atribuições/comparações, 8 produtos e 2 somas. O que lhe dá um esforço menor ou igual a $39c$.

$m_3 \setminus m_1$	$2/3$	$1/3$
1	$\{a, b\}$	Θ
$\{a, b\}$	$\{a, b\}$	$\{a, b\}$

 \longrightarrow

$m_2 \setminus m_{13}$	1
$2/3$	$\{a, b\}$
$\{b, c\}$	$\{b\}$
$1/3$	$1/3$
Θ	$\{a, b\}$

Figura 4.1 $(Cr_1 \oplus Cr_3) \oplus Cr_2$

$m_3 \setminus m_2$	$2/3$	$1/3$
1	$\{b, c\}$	Θ
$\{a, b\}$	$\{b\}$	$\{a, b\}$

 \longrightarrow

$m_1 \setminus m_{23}$	$2/3$	$1/3$
1	$\{b\}$	$\{a, b\}$
$2/3$	$4/9$	$2/9$
$\{a, b\}$	$\{b\}$	$\{a, b\}$
$1/3$	$2/9$	$1/9$
Θ	$\{b\}$	$\{a, b\}$

Figura 4.2 $(Cr_2 \oplus Cr_3) \oplus Cr_1$

$m_2 \setminus m_1$	$2/3$	$1/3$
1	$\{a, b\}$	Θ
$2/3$	$4/9$	$2/9$
$\{b, c\}$	$\{b\}$	$\{b, c\}$
$1/3$	$2/9$	$1/9$
Θ	$\{a, b\}$	Θ

 \longrightarrow

$m_3 \setminus m_{12}$	$2/9$	$4/9$	$2/9$	$1/9$
1	$\{a, b\}$	$\{b\}$	$\{b, c\}$	Θ
$2/9$	$4/9$	$2/9$	$1/9$	
$\{a, b\}$	$\{a, b\}$	$\{b\}$	$\{b\}$	$\{a, b\}$

Figura 4.3 $(Cr_1 \oplus Cr_2) \oplus Cr_3$

Mas a escolha da ordem só terá importância se conhecermos as características das distribuições de mbc de cada função Cr_i .

4.2 USANDO HEURÍSTICAS

A seguir exploramos a estrutura das funções de crença de modo a minimizar o esforço de combinação das mesmas. Isto é feito em termos de um conjunto de premissas que ajudam na redução do esforço de combinação de evidências.

Função identidade: Se Cr_1 e Cr_2 são duas funções de crença e Cr_1 é uma função vacuosa, então $Cr_1 \oplus Cr_2 = Cr_2$.

Prova: Segue-se da soma ortogonal.

Nenhum esforço precisa ser gasto nesta combinação pois o resultado é conhecido a priori.

Funções com foco comum: Se Cr_1 e Cr_2 são funções de suporte simples (fss) com mesmo foco A , então $Cr = Cr_1 \oplus Cr_2$ também é uma função de suporte simples com foco em A .

Prova: Segue-se da soma ortogonal.

Por esta assertiva, a combinação de funções com foco comum não gera fragmentação dos elementos focais. Assim, dada uma lista de funções de crença a combinar, deve-se substituir todas as fss de foco comum pela soma ortogonal delas. Isto minimiza o esforço de combinação. Contudo, há um esforço adicional para se determinar quais são as fss, e, dentre estas, quais as que têm foco comum. O esforço para determinar se uma função de crença é uma fss é constante, isto é, tem ordem de complexidade $O(1)$. E, dado duas fss, com focos A_1 e A_2 , determinar se elas têm ou não focos em comum, $A_1 = A_2$ ou $A_1 \neq A_2$, envolve um tempo de ordem linear no tamanho do maior foco entre elas. Isto é, $O(|A|)$, onde $|A| = \max(|A_1|, |A_2|)$. O exemplo a seguir ilustra a combinação de duas fss.

Exemplo: Sejam $\Theta = \{a, b\}$, Cr_1 e Cr_2 fss com mbc m_1 e m_2 , respectivamente. Se $Cr = Cr_1 \oplus Cr_2$ é uma função de crença com mbc m , então a tabela 4.1 ilustra os valores das mbc das funções combinadas e da função resultante.

	m1	m2	m
a	s1	s2	s1+s2-s1s2
⊙	1-s1	1-s2	(1-s1)(1-s2)

Tabela 4.1: $Cr = Cr_1 \oplus Cr_2$

Funções com dicotomia comum: Se Cr_1 e Cr_2 são funções de crença dicotômicas com dicotomia comum (A, \bar{A}) , então $Cr = Cr_1 \oplus Cr_2$ é também dicotômica com a mesma dicotomia (A, \bar{A}) .

Prova: Segue-se da soma ortogonal.

Por esta assertiva, a combinação de funções de crença dicotômicas com dicotomia comum não gera fragmentação dos elementos focais. Assim, dada uma lista de funções de crença a combinar, deve-se substituir todas as funções dicotômicas com dicotomia comum pela soma ortogonal delas. Isto minimiza o esforço de combinação. Contudo, há um esforço adicional para determinar quais às funções da lista são dicotômicas e dentre estas quais as que têm dicotomia comum. O esforço para se determinar se uma dada função de crença é dicotômica é de ordem $O(|\Theta|)$. E, dada duas funções de crença dicotômicas, o esforço para determinar se elas têm dicotomia comum ou não é $O(|A|)$, onde A é o maior elemento focal entre elas, exceto Θ . O exemplo a seguir ilustra a combinação de duas funções de crença dicotômicas.

Exemplo: Sejam $\Theta = \{a, b, c, d\}$, Cr_1 e Cr_2 funções de crença com dicotomia comum $\{(a, b), \{c, d\}\}$, e mbc dadas por

$$m_1(\{a, b\})=1/3, m_1(\{c, d\})=1/3 \text{ e } m_1(\Theta)=1/3; \text{ e,}$$

$$m_2(\{a, b\})=1/4, m_2(\{c, d\})=2/4 \text{ e } m_2(\Theta)=1/4.$$

Então a mbc m de $Cr = Cr_1 \oplus Cr_2$ é dado por

$$m(\{a, b\})=3/9, m(\{c, d\})=5/9 \text{ e } m(\Theta)=1/9.$$

Funções politômicas sobre uma partição comum: Se Cr_1 e Cr_2 são funções de crença politômicas definidas em uma partição $\Omega = \{A_1, A_2, \dots, A_f\}$ de Θ , então $Cr = Cr_1 \oplus Cr_2$ é também politômica em Ω .

Prova: Sejam F, F_1 e F_2 os conjuntos focais de Cr, Cr_1 e Cr_2 , respectivamente. Pela soma ortogonal tem-se:

- (1) Se $\Theta \in F_1$ e $\Theta \in F_2$, então $F = F_1 \cup F_2$.
- (2) Se $\Theta \in F_1$ e $\Theta \notin F_2$, então $F = F_1$.

- (3) Se $\Theta \in F_1$ e $\Theta \in F_2$, então $F = F_1$.
- (4) Se $\Theta \in F_1$ e $\Theta \notin F_2$, então $F = F_1 \cap F_2$.

Os quatro itens exaurem todas as possibilidades, e em cada deles tem-se que $F \in \mathcal{C}(A_1, A_2, \dots, A_f, \Theta)$. Logo Cr é politômica em Ω .

```

Função Politomia ( $\mathcal{M}$ ): booleana ;
(*  $\mathcal{M} = (\Theta, F, \mu)$  representa a função  $Cr$  *)
Var  $V$ : array [ 1.. $|\Theta|$  ] de booleanos ;
(* supoe-se que o compilador aloque o array
   com todos os elementos setados para falso *)
Início
  Para cada  $Ac(F \setminus \Theta)$ , faça
    Para cada  $i \in A$ , faça
      Se  $V_i = \text{Verdadeiro}$ , então Retorna(falso)
      Senão  $V_i \leftarrow \text{Verdadeiro}$  ;
  Retorna(Verdadeiro) ;
Fim ;

```

Figura 4.4 Algoritmo Politomia

Sejam $f = |\Omega|$, pela assertiva de funções politômicas sobre uma partição comum Ω de Θ , combinar um número qualquer dessas funções não altera o limite máximo de elementos focais da função resultante que é $f+1$. Assim dada uma lista de funções de crença a combinar, deve-se substituir todas as funções politômicas sobre uma partição comum de Θ pela soma ortogonal delas. Isto minimiza o esforço de combinação, e mantém a fragmentação de elementos focais da função resultante sobre controle. Mas há um custo adicional para identificar aquelas funções que são politômicas e dentre as politômicas identificar aquelas que são definidas sobre uma mesma partição. O Algoritmo Politomia, ilustrado na figura 4.4, determina se uma função de crença, na representação \mathcal{M} , é politômica ou não a um custo de ordem linear $O(|\Theta|)$. E, Dado duas funções politômicas, O algoritmo PartiçãoPolitômica, ilustrado na figura 4.5, determina se elas estão definidas sobre uma mesma partição de Θ ou não, em tempo de ordem linear $O(|\Theta|)$.

Observa-se que este procedimento pode dar verdadeiro para Cr_1 e

Cr_2 , Cr_1 e Cr_3 e falso para Cr_2 e Cr_3 . Assim, se Cr_1 e Cr_2 são definidas na mesma partição, o próximo teste deverá ser feito entre $Cr_1 \oplus Cr_2$, e Cr_3 , evitando falsas inferências.

```

Função PartiçãoPolitômica ( $\mathcal{M}_1, \mathcal{M}_2$ ): booleana ;
(*  $\mathcal{M}_i = (\Theta, \mathbb{F}_i, \mu_i)$  representa a função Politômica  $Cr_i$  *)
Var  $\mathbb{V}, \mathbb{T}$ : array [ 1.. $|\Theta|$  ] de inteiros ;
(* supoe-se que o compilador aloque o array
   com todos os elementos setados para zero *)
Início
  nef  $\leftarrow$  0 ;
  Para cada  $AC(\mathbb{F}_1 \setminus \Theta)$ , faça
  Início
    ne  $\leftarrow$  0 ;
    nef  $\leftarrow$  nef + 1 ;
    Para cada  $i \in A$ , faça
    Início
      ne  $\leftarrow$  ne + 1 ;
       $\mathbb{T}_i \leftarrow$  nef ;
    Fim ;
     $\mathbb{V}_i \leftarrow$  ne ;
  Fim ;
  Para cada  $AC(\mathbb{F}_2 \setminus \Theta)$ , faça
  Início
    ne  $\leftarrow$  0 ;
    nef  $\leftarrow$  0 ;
    Para cada  $i \in A$ , faça
    Início
      ne  $\leftarrow$  ne + 1 ;
      Se  $\mathbb{T}_i > ne = 1$ , então nef  $\leftarrow$   $\mathbb{T}_i$  ;
      Se  $\mathbb{T}_i \neq nef$ , então Retorna(Falso) ;
    Fim ;
    i  $\leftarrow$  nef ;
    Se  $i > 0$   $\mathbb{V}_i \neq ne$ , Retorna(Falso) ;
  Fim ;
  Retorna(Verdadeiro) ;
Fim ;

```

Figura 4.5 Algoritmo PartiçãoPolitômica

Redução de núcleo: Se Cr_1 e Cr_2 são funções de crença com núcleo N_1 e N_2 respectivamente, e $N_1 \cap N_2 \neq \emptyset$, então $Cr = Cr_1 \oplus Cr_2$ terá núcleo $|N| \leq \min(|N_1|, |N_2|)$.

Prova: Sejam F , F_1 e F_2 os conjuntos focais de Cr , Cr_1 e Cr_2 , respectivamente. Como cada elemento $A \in F$ é o resultado da interseção de $B \in F_1$ e $D \in F_2$, se $B \cap D \neq \emptyset$. Então, se $A \in F$, existe um elemento B de F_1 e um elemento D de F_2 , onde $A \subset B$ e $A \subset D$. Mas N é a união dos elementos de F , assim $N \subset N_1$ e $N \subset N_2$. Logo $|N| \leq \min(|N_1|, |N_2|)$.

Como o número de elementos focais de uma função de crença é limitado pela cardinalidade do conjunto das partes de seu núcleo, então quanto menor for o núcleo menor será o limite superior do número de elementos focais da função. Pela assertiva de redução de núcleos podemos obter uma provável redução do esforço de combinação se dada uma lista $L = \mathcal{M}_1, \dots, \mathcal{M}_k$ de funções de crença combinarmos sempre as duas funções de menor e maior núcleos respectivamente.

Exemplo: Sejam as funções de crença Cr_1 , Cr_2 e Cr_3 definidas sobre $\Theta = \{a, b, c, d\}$, e m_1 , m_2 , m_3 sejam suas mbc, respectivamente. Onde $m_1(\{a\}) > 0$, $m_1(\Theta) > 0$ e zero para os demais casos; $m_2(\{a, b\}) > 0$, $m_2(\{b, d\}) > 0$ e zero para os demais casos; $m_3(\{b, c\}) > 0$, $m_3(\{c\}) > 0$ e zero para os demais casos. Portanto uma ordenação crescente pela cardinalidade dos núcleos produz Cr_3, Cr_2, Cr_1 . Os esquemas nas figuras 4.6 e 4.7 mostram que se combinarmos $(Cr_3 \oplus Cr_1) \oplus Cr_2$ teremos 8 intersecções e 8 produtos, entretanto se combinarmos na ordem $(Cr_1 \oplus Cr_2) \oplus Cr_3$ teremos 10 intersecções e 10 produtos, o que respeitada a escala há um substancial aumento de esforço. Ainda que tenhamos o esforço de ordenação, vale a pena procurar uma melhor ordem de combinação.

$m_3 \setminus m_1$	$\{a\}$	Θ	\Rightarrow	$m_1 \setminus m_2$	$\{a, b\}$	$\{b, d\}$
$\{b, c\}$	\emptyset	$\{b, c\}$		$\{b, c\}$	$\{b\}$	$\{b\}$
$\{c\}$	\emptyset	$\{c\}$		$\{c\}$	\emptyset	\emptyset

Figura 4.6 $(Cr_1 \oplus Cr_3) \oplus Cr_2$

$m_2 \setminus m_1$	$\{a\}$	\emptyset	\Rightarrow	$m_1 \setminus m_2$	$\{b, c\}$	$\{c\}$
$\{a, b\}$	$\{a\}$	$\{a, b\}$		$\{a\}$	\emptyset	\emptyset
$\{b, d\}$	\emptyset	$\{b, d\}$		$\{a, b\}$	$\{b\}$	\emptyset
				$\{b, d\}$	$\{b\}$	\emptyset

Figura 4.7 $(Cr_1 \oplus Cr_2) \oplus Cr_3$

Em geral, a operação de maior custo no processo de combinação de evidências é a interseção de elementos focais. Há casos, contudo, em que a aplicação de uma ou mais das heurísticas apresentadas não envolve muita interseção para identificar as funções que podem ser combinadas por uma dada heurística. Em outros casos o custo de aplicação de uma heurística é muito alto. Assim a aplicação de qualquer heurística envolve um estudo de custos, a priori, para cada instância do problema.

5 ALGORITMOS PARA COMBINAR EVIDÊNCIAS

5.1 COMBINANDO FUNÇÕES BAYESIANAS

Lema 1: Seja Cr_0 uma função de crença bayesiana e Cr_1 uma função de crença qualquer, se elas forem combináveis ($\mathbb{N}_0 \cap \mathbb{N}_1 \neq \emptyset$), então $Cr_0 \oplus Cr_1$ é uma função de crença Bayesiana.

Prova: Sejam \mathbb{F} , \mathbb{F}_0 e \mathbb{F}_1 os conjuntos focais de Cr , Cr_0 , e Cr_1 respectivamente, onde $Cr = Cr_0 \oplus Cr_1$. Como $\mathbb{F} = \{x \mid x = A \cap B, A \in \mathbb{F}_0 \text{ e } B \in \mathbb{F}_1\}$, se $x \in \mathbb{F}$, então $x \in \mathbb{F}_0$, pois Cr_0 é Bayesiana. Assim todos os elementos focais de \mathbb{F} são atômicos. Logo Cr também é Bayesiana.

Teorema 1: Seja $\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_k$ uma lista de funções de crença, onde ao menos uma delas é bayesiana, e seja $\mathcal{M} = \bigoplus_{i=0}^k \mathcal{M}_i$ a soma ortogonal delas, então \mathcal{M} é bayesiana.

Prova: Segue-se do lema 1 e da propriedade associativa e comutativa da soma ortogonal, por indução sobre k .

Teorema 2: Sejam Cr , Cr_0 e Cr_1 funções de crença definidas em Θ , e Pl_1 a função de plausibilidade associada a Cr_1 . Onde $Cr = Cr_0 \oplus Cr_1$, Cr_0 é uma função de crença bayesiana e Cr_1 é uma função de crença qualquer. Então mbc m de Cr é dada por

$$m(\theta) = m_0(\theta) \cdot Pl_1(\theta), \quad \theta \in \Theta$$

Prova: Pela soma ortogonal temos que a mbc m de Cr é dada por

$$\begin{aligned}
 m(\theta) &= \sum_{\substack{A \cap B = \theta \\ A \in \mathcal{F}_0, B \in \mathcal{F}_1}} m_0(A) \cdot m_1(B), \\
 &= m_0(\theta) \cdot \sum_{\substack{\{\theta\} \cap B \neq \emptyset \\ B \in \mathcal{F}_1}} m_1(B), \\
 &= m_0(\theta) \cdot Pl_1(\theta), \quad \forall \theta \in \Theta
 \end{aligned}$$

Passamos da primeira para a segunda igualdade pelo fato dos elementos de Cr_0 serem atômicos. E, da segunda para a terceira igualdade pela fórmula (3.2) de conversão de mbc m para Pl .

Procedimento CombinaBayes ;

(* Entrada: Lista de Funções de Crença $\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_k$.
 Saída: Função de crença resultante \mathcal{M} *)

*)

- 1 $P_0: \forall i \in I, \text{ faça } V_i \leftarrow 0 ; \quad (* I = \{1, 2, \dots, n\} *)$
- 2 $\mathcal{M} \leftarrow \mathcal{M}_0 ; \quad (* \mathcal{M} = (\mathcal{F}, \mu) \text{ é a soma final} *)$
- 3 $P_1: \forall j \in K, \text{ faça } \quad (* K = \{1, 2, \dots, k\} *)$
- 4 Início
- 5 $\forall i \in \mathcal{F}_0, \text{ faça } V_i \leftarrow 1 ;$
- 6 $\mathcal{M} \leftarrow \emptyset ; \quad (* \text{Zera } \mathcal{M} *)$
- 7 $\forall i \in \mathcal{F}_j, \text{ faça}$
- 8 Se $V_i = 1, \text{ então}$
- 9 Início
- 10 $\mathcal{F} \leftarrow \mathcal{F} \cup \{i\} ;$
- 11 $\mu(\{\theta_i\}) \leftarrow \mu_0(\{\theta_i\}) * \mu_j(\{\theta_i\}) ;$
- 12 Fim ;
- 13 $\forall i \in \mathcal{F}_0, \text{ faça } V_i \leftarrow 0 ;$
- 14 $\mathcal{M}_0 \leftarrow \mathcal{M} ;$
- 15 Fim ;
- 16 $P_2: S \leftarrow 0 ; \quad (* \text{Renormaliza} *)$
- 17 $\forall i \in \mathcal{F}, \text{ faça } S \leftarrow S + \mu(\{\theta_i\}) ;$
- 18 $\forall i \in \mathcal{F}, \text{ faça } \mu(\{\theta_i\}) \leftarrow \mu(\{\theta_i\}) / S ;$

Figura 5.1 Algoritmo: CombinaBayes

5.1.1 BAYESIANA x BAYESIANA

Teorema: Dado uma função de crença bayesiana \mathcal{M}_0 , e uma lista $L = \mathcal{M}_1, \dots, \mathcal{M}_k$ de funções bayesianas, o algoritmo CombinaBayes, ilustrado na figura 5.1, combina-as de forma eficiente e com complexidade linear em k e $n = |\Theta|$, isto é, $O(k.n)$.

Prova: No passo P_0 , é feito n atribuições. No passo P_1 , para cada uma das k funções de crença combinadas são feitos no máximo $4n$ atribuições, n produtos e n comparações. No Passo P_2 , são feitas $2n$ atribuições. Seja c uma constante representando o esforço da operação de maior esforço dentre *atribuir*, *comparar*, *multiplicar* e *somar*. Neste caso o esforço total do algoritmo é dado por $3nc + 4ckn$. Isto é, o tempo de computação do algoritmo é linear em $|\Theta|$ e no número de funções combinadas k , cuja complexidade é expressa por $O(k.n)$.

5.1.2 SOMA ORTOGONAL COMO CONDICIONAMENTO

Teorema: Dado uma função de crença bayesiana \mathcal{M}_0 e uma lista $L = \mathcal{M}_1, \dots, \mathcal{M}_k$ de funções de suporte simples, onde cada \mathcal{M}_i tem o grau de crença do foco A_i igual a 1. O algoritmo ilustrado na figura 5.1, com a linha 7 substituída pelas linhas 7.1 e 7.2, e a linha 11 pela linha 11.1, dadas a seguir

- 7.1 $\forall A \in \mathbb{F}_j$, faça (* $|\mathbb{F}|=1$ *)
 7.2 $\forall i \in A$, faça
 11.1 $\mu(\{\theta_i\}) \leftarrow \mu_0(\{\theta_i\}) * \mu_j(A)$

combina estas $k+1$ funções com complexidade de tempo $O(k.n)$. A este novo algoritmo chamaremos de algoritmo de Condicionamento tendo em vista sua equivalência com a fórmula de condicionamento de Bayes. Prova: A sua complexidade e análise é idêntica ao algoritmo da figura 5.1, feita na subseção 5.1.1. Isto fica claro pelo fato do conjunto focal \mathbb{F}_j , da função de crença \mathcal{M}_j , ter apenas um elemento focal.

5.1.3 BAYESIANA x FUNÇÃO DE SUPORTE SIMPLES

Teorema: Dado uma função de crença bayesiana \mathcal{M}_0 e uma lista $L = \mathcal{M}_1, \dots, \mathcal{M}_k$ de funções de suporte simples, o algoritmo CombinaBayesXCrença, ilustrado na figura 5.2, combina essas

Procedimento CombinaBayesXCrença ;

(* Entrada: Lista de Funções de Crença $\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_k$.
 Saída: Função de crença resultante \mathcal{M} *)

- 1 P₀: $\forall i \in \mathbb{I}$, faça $V_i \leftarrow 0$; (* $\mathbb{I} = \{1, 2, \dots, n\}$ *)
- 2 $\mathcal{M} \leftarrow \mathcal{M}_0$; (* $\mathcal{M} = \langle \mathbb{F}, \mu \rangle$ é a soma final *)
- 3 P₁: $\forall j \in \mathbb{K}$, faça (* $\mathbb{K} = \{1, 2, \dots, k\}$ *)
- 4 Início
- 5 $\forall i \in \mathbb{F}_0$, faça $V_i \leftarrow 1$;
- 6 $\mathcal{M} \leftarrow \phi$; (* Zera \mathcal{M} *)
- 7.1 $\forall A \in \mathbb{F}_j$, faça
- 7.2 $\forall i \in A$, faça
- 8 Se $V_i = 1$, então
- 9 Início
- 10.1 $\mathbb{F} \leftarrow \mathbb{F} \cup \{i\}$; $V_i \leftarrow 2$;
- 11.1 $\mu(\{\theta_i\}) \leftarrow \mu_0(\{\theta_i\}) * \mu_j(A)$
- 12 Fim
- 12.1 Senão
- 12.2 Se $V_i = 2$, então
- 12.3 $\mu(\{\theta_i\}) \leftarrow \mu(\{\theta_i\}) + \mu_0(\{\theta_i\}) * \mu_j(A)$;
- 13 $\forall i \in \mathbb{F}_0$, faça $V_i \leftarrow 0$;
- 14 $\mathcal{M}_0 \leftarrow \mathcal{M}$;
- 15 Fim ;
- 16 P₂: $S \leftarrow 0$; (* Renormaliza *)
- 17 $\forall i \in \mathbb{F}$, faça $S \leftarrow S + \mu(\{\theta_i\})$;
- 18 $\forall i \in \mathbb{F}$, faça $\mu(\{\theta_i\}) \leftarrow \mu(\{\theta_i\}) / S$;

Figura 5.2 Algoritmo: bayesiana X função de crença

funções de forma eficiente e em tempo polinomial de ordem $O(k.n)$.
 Prova: Em relação ao algoritmo CombinaBayes, ilustrado na figura 5.1, CombinaBayesXfss tem um esforço adicional máximo de $2n$ atribuições, n somas e n produtos, devido à possibilidade de que cada função de crença possa ter até dois elementos focais. Mas este esforço não altera a ordem de grandeza do tempo de computação que continua sendo $O(k.n)$.

5.1.4 BAYESIANA x FUNÇÃO DICOTÔMICA

Teorema: Dado uma função de crença bayesiana \mathcal{M}_0 e uma lista $L = \mathcal{M}_1, \dots, \mathcal{M}_k$ de funções dicotômicas (\mathcal{M} com $\mathbb{F} \subset \{A, \bar{A}, \emptyset\}$), o algoritmo **CombinaBayesXCrença**, ilustrado na figura 5.2, neste caso chamado **CombinaBayesXDicotômico**, combina essas funções de forma eficiente e em tempo polinomial de ordem $O(k.n)$.

Prova: Aqui, o acréscimo de esforço, em relação ao algoritmo **CombinaBayes**, ilustrado na figura 5.1, é de $4n$ atribuições, $2n$ somas e $2n$ produtos, devido possibilidade de que cada função de crença possa ter até três elementos focais. Contudo, isto não altera a ordem de grandeza da complexidade de tempo que continua sendo $O(k.n)$.

5.1.5 BAYESIANA x FUNÇÃO DE CRENÇA

Teorema: Dado uma função de crença bayesiana \mathcal{M}_0 e uma lista $L = \mathcal{M}_1, \dots, \mathcal{M}_k$ de funções de crença quaisquer, o algoritmo **CombinaBayesXCrença**, ilustrado na figura 5.2, combina essas funções de forma eficiente e em tempo polinomial de ordem $O(k.f.n)$.

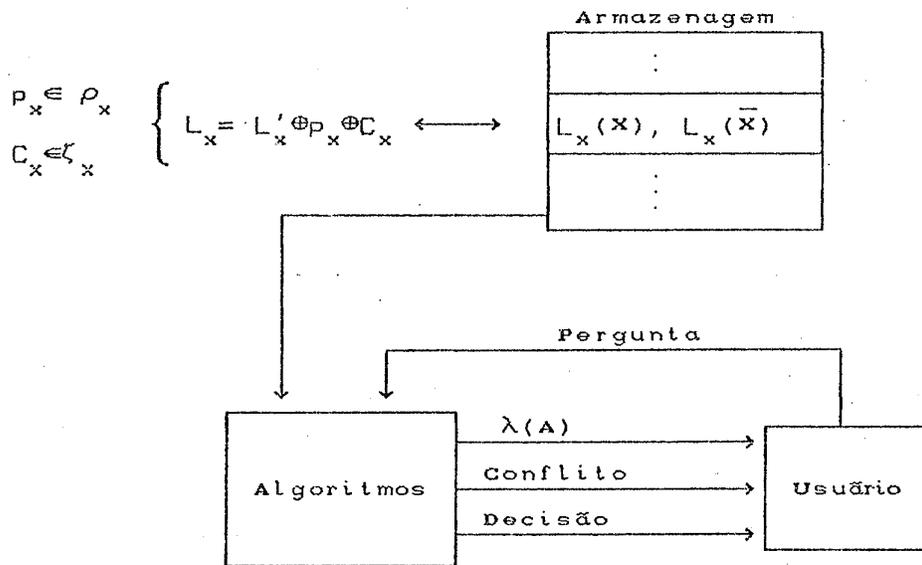
Prova: Agora, o conjunto focal de cada função de crença da lista tem um número variável de elementos focais. E, a complexidade reflete este fato. Uma análise das iterações do algoritmo **CombinaBayesXCrença**, seguida de majorações sobre o tamanho máximo de cada elemento focal, e tomando f como o número máximo de elementos focais entre as funções da lista L revela uma complexidade de tempo $O(k.f.n)$.

Observamos que o uso da regra de Dempster para combinar funções de crença, onde pelo menos uma das funções é bayesiana, é tratável se conseguirmos que a primeira soma ortogonal, da série de somas, seja com uma função bayesiana. Isto não parece muito difícil, já que na maioria das aplicações envolvendo incerteza, normalmente, o especialista dispõe de algum conhecimento prévio, a ignorância quase nunca é absoluta.

5.2 ALGORITMO DE BARNETT

O algoritmo para combinação de evidências, proposto por Barnett(1981), está baseado em funções dicotômicas, isto é, funções simples que têm como elementos focais A , \bar{A} e \emptyset , onde A é um elemento de uma dada partição Ω de Θ . Os problemas que permitem disciplinar a coleta de evidências de modo a que cada peça de evidência confirme ou negue uma proposição elementar de Ω podem usufruir da proposta de Barnett. Este algoritmo faz a combinação de evidências em tempo polinomial na cardinalidade de Ω e calcula qualquer das medidas m , Cr , Pl , Q e K em tempo também polinomial. Em seguida tentamos delinear este algoritmo.

Seja Ω uma partição de Θ com n elementos, $n=|\Omega|$. Para cada proposição $x \in \Omega$ há uma lista ρ_x de funções de suporte simples (fss) definidas em Ω e com foco em x , e uma outra lista ζ_x de fss definidas em Ω e com foco \bar{x} . A combinação das fss das listas ρ_x e ζ_x geram a função dicotômica L_x com dicotomia (x, \bar{x}) . As funções dicotômicas L_x são funções locais a cada proposição elementar $x \in \Omega$. Elas são mantidas em uma base de dados, e a cada vez que nova fss seja incluída em ρ_x ou em ζ_x a função L_x é atualizada. A manutenção das funções L_x em uma base de dados é necessária também para atender as consultas do usuário, conforme sugere a figura 5.3.



5.3 Algoritmo de Barnett

A figura acima define o fluxo de dados entre as diversas fontes de evidências, uma armazenagem central e o usuário. A cada consulta do usuário, um dos algoritmos propostos por Barnett acessa a armazenagem, procede os cálculos e devolve a resposta à consulta formulada. Na figura λ representa uma das medidas Cr, m, Q, Pl ou K. O fator de normalização é dado pelo logaritmo de K. A decisão é a resposta que o usuário espera obter sobre qual proposição tem maior massa básica de crença.

O algoritmo RespondeCrença, ilustrado na figura 5.4, mostra como se calcula cada uma das medidas de crença, conforme proposto por Barnett(1981). Como fica claro pelo contexto, usamos a notação $L(x)$ e $L(\bar{x})$ ao invés de $L_x(x)$ e $L_x(\bar{x})$, sem prejuízo do entendimento. Também usamos $t \in \Omega$, representando $\{t\} \subset \Omega$, de modo a simplificar a notação.

Function RespondeCrença(λ :MedidadeCrença; A:Subconjuto): Real ;
 (* $L(x)$ e $L(\bar{x})$, Funções dicotômicas definidas em memória.

Uma para cada $x \in \Omega$. *)

Início

(* Calcula o Fator de renormalização *)

$$K^{-1} = \left[\prod_{t \in \Omega} (1-L(t)) \right] \left[\sum_{t \in A} \frac{L(t)}{1-L(t)} + 1 \right] - \prod_{t \in \Omega} L(\bar{t}) ;$$

Caso λ Seja

Início

m: Caso $|A|$ Seja (* calcula a mbc $m(A)$ *)

Início

0: Início $m(\phi) := 0$; Retorna($m(\phi)$) Fim ;

1: Início

$$m(A) = K \left[\prod_{\substack{t \in \Omega \\ t \neq A}} (1-L(t)) + (1-L(A)-L(\bar{A})) \cdot \prod_{\substack{t \in \Omega \\ t \neq A}} L(\bar{t}) \right] ;$$

Retorna($m(A)$) ;

Fim ;

2: $|\Omega|$: Início

$$m(A) = K \left[\prod_{t \in A} (1-L(t) - L(\bar{t})) \right] \left[\prod_{t \notin A} L(\bar{t}) \right] ;$$

Retorna($m(A)$) ;

Fim ;

Fim ; (* Fim de Caso de m *)

Cr: Caso $|A|$ Seja

Início

0: Retorna(0) ; (* $Cr(\phi) = 0$ *)

1: Início (* $Cr(A) = m(A)$ *)

$$m(A) = K \left[\prod_{\substack{t \in \Omega \\ t \neq A}} (1-L(t)) + (1-L(A)-L(\bar{A})) \cdot \prod_{\substack{t \in \Omega \\ t \neq A}} L(\bar{t}) \right] ;$$

Retorna($m(A)$) ;

Fim ;

```

2 .. |\Omega|:
  Início
    Cr(A) = K \left[ \left[ \prod_{t \in \Omega} (1-L(t)) \right] \left[ \sum_{t \in A} \frac{L(t)}{1-L(t)} \right] + \right.
      \left. \left[ \prod_{t \in A} (1-L(t)) \right] \left[ \prod_{t \notin A} L(\bar{t}) \right] - \prod_{t \in \Omega} L(\bar{t}) \right];
    Retorna( Cr(A) );
  Fim ;
Fim ; (* Fim de Caso de Cr *)
P1: Caso |A| Seja
  Início
    0: Retorna(0) ; (* P1(\phi) = 0 *)
    1 .. |\Omega|:
      Início (* P1(A) = 1 - Cr(\bar{A}) *)
        P1(A) = 1 - RespondeCrença(Cr, \bar{A} ) ;
        Retorna( P1(A) ) ;
      Fim ;
    Fim ; (* Fim de Caso P1 *)
Q: Caso |A| Seja
  Início
    0: Retorna( 1.0 ); (* Q(\phi) = 1 *)
    1: Início
      Q(A) = k \left( \frac{1-L(\bar{A})}{1-L(A)} \right) \left[ \prod_{t \in \Omega} (1-L(t)) \right];
      Retorna( Q(A) );
    Fim ;
  2 .. |\Omega|:
    Início:
      Q(A) = k \left[ \prod_{t \notin A} (1-L(t)-L(\bar{t})) \right] \left[ \prod_{t \in A} (1-L(t)) \right];
      Retorna( Q(A) );
    Fim ;
  Fim ; (* Fim d Caso Q *)
Fim ; (* fim de Caso \lambda *)
Fim ; (* Fim do Procedimento RespondeCrença *)

```

Figura 5.4 Algoritmo RespondeCrença

A despeito do esforço para se calcular uma medida de crença λ ser polinomial no tamanho de Ω , se o usuário pedir informações sobre todos os $A \subset \Omega$, obviamente que isto teria um esforço que cresceria exponencialmente com o aumento de Ω .

O algoritmo RespondeCrença supõe que todas as funções L_x estejam definidas (ainda que sejam vacuosas). Contudo, a atualização das funções $L_x = L'_x \oplus p_x$ ou $L_x = L'_x \oplus c_x$, $\forall x \in \Omega$, é feita diretamente pela *Soma Ortogonal*.

5.3 ALGORITMO DE SHAFER-LOGAN

Gordon & Shortlife (1984,1985) Trabalhando com hipóteses hierarquizadas em árvore, onde as funções de suporte simples focalizam a favor ou contra as proposições representadas pelos nós da árvore, concluíram que a atualização do conhecimento usando a soma ortogonal de Dempster não é viável em tempo polinomial. Isto os motivou a propor uma versão simplificada daquela regra que tornou o esforço de cálculo polinomial. Mas com isto perderam o suporte teórico e passaram a ter aproximações, em certos casos, completamente insatisfatórias.

Shafer & Logan(1987) propuseram um algoritmo que resolve o problema de Gordon e Shortlife em tempo polinomial, usando todo o suporte teórico e sem aproximações. Silva & Milidiú (1989a) propõem uma versão mais eficiente deste algoritmo, apresentando complexidade $O(|\Theta|)$, enquanto o de Shafer e Logan tem uma complexidade $O(r^2 |\Theta|)$, onde r é o fator de ramificação da árvore. Contudo Shafer & Logan afirmam que a complexidade de seu algoritmo é linear em r , a despeito da rotina 5 por eles proposta apresentar complexidade quadrática em relação a r . A seguir são apresentadas as principais características e os procedimentos deste algoritmo.

5.3.1 CARACTERÍSTICAS

Nas hipóteses hierarquizadas em árvore, cada nó folha representa uma proposição elementar do quadro de discernimento do problema e os nós internos representam uma disjunção dos seus nós filhos. Em termos de conjunto, cada nó interno é a união dos seus nós filhos.

Um fundamento básico é que os filhos de cada nó A unidos ao complemento de A constiuem uma partição $\Omega_A (=F_A \cup \bar{A})$ do quadro de discernimento Θ do problema. Onde F_A é o conjunto dos filhos de A . Assim, as funções de suporte dicotômica a favor ou contra os filhos de A (definidas em Ω_A) podem ser somadas usando o algoritmo de Barnett(1981), já que todo $B \in F_A$ é um átomo da partição Ω_A . Uma vez feita a soma podemos obter a crença em A e \bar{A} , usando o mesmo algoritmo. Com isto, as funções de crenças focalizadas sobre os filhos podem ser combinadas, o seu impacto sobre o nó pai avaliado, e o resultado combinado com a função de suporte de

crença sobre o próprio pai. Este processo permite a propagação das evidências de baixo para cima na árvore. O processo de propagação de baixo para cima termina quando, para todos os filhos da raiz, foram calculados os impactos das evidências que suportam fss sobre os nós das subárvores que eles representam. Agora, a propagação de cima para baixo tem lugar. Ela consiste em calcular a influência que as evidências fora da subárvore de cada nó tem sobre este mesmo nó. Assim, tendo as funções de crença L_x , suportada pelas evidências que impactaram diretamente o nó x ; F_x , suportada pelas evidências que impactaram diretamente os nós da subárvore com raiz x , exceto x ; e Π_x , suportada pelas evidências que impactaram diretamente nós fora da subárvore x podemos combinar L_x , F_x e Π_x para obter a função de crença final C_x , a qual é suportada por todas as evidências que impactaram diretamente algum nó da árvore de hierarquia.

Baseados no algoritmo de Barnett, nos axiomas da Teoria de Dempster-Shafer e numa dosagem alta de criatividade Shafer e Logan delinearam o algoritmo para evidências hierarquizadas. Com ele nós podemos obter medidas de crença para proposições que não estão na árvore, o que não ocorre com o algoritmo proposto por Gordon e Shortlife(1984,1985).

5.3.2 ALGORITMO

5.3.2.1 NOTAÇÃO E ESTRUTURA

A estrutura básica é uma árvore onde Θ é a raiz, e cada nó representa a união de todos o seus nós filhos. Usaremos uma notação similar a de Silva e Milidiú(1989a).

- \mathcal{A} - Coleção de todos os nós da árvore, exceto a raiz.
- D_A - Coleção de todos os filhos do nó A .
- L_A - Função de crença dicotômica, resultante da soma ortogonal das funções de suporte simples a favor e contra o nó $A \in \mathcal{A}$.
- F_A - Função de crença dicotômica, resultante da soma das funções dicotômicas dos filhos de A pelo algoritmo de Barnett(1981). Em verdade F_A representa a soma ortogonal de todas as crenças dos nós abaixo de A .
- Π_A - Função de crença dicotômica, representa a soma ortogonal

de todas as crenças dicotômicas de nós que nem estão abaixo de A e nem são A .

C_A - Representa a soma ortogonal de todas as funções de crença suportadas pelas evidências coletadas.

Nós estamos supondo que cada nó $A \in \mathcal{A}$ já tenha sua função L_A . Isto é, já se encerrou a fase de coleta de evidências e as funções de suporte simples, a favor ou contra A , foram combinadas gerando as funções dicotômicas L_A , cujos dois valores $L_A(A)$ e $L_A(\bar{A})$ estão armazenados no nó A .

O objetivo é calcular $C_A = L_A \oplus F_A \oplus \Pi_A$ para todo $A \in \mathcal{A}$.

Para facilitar a notação, usaremos os símbolos abaixo na descrição do algoritmo:

$$\begin{array}{lll} L(A) = L_A(A) & e & L(\bar{A}) = L_A(\bar{A}) \\ F(A) = F_A(A) & e & F(\bar{A}) = F_A(\bar{A}) \\ \lambda(A) = (L_A \oplus F_A)(A) & e & \lambda(\bar{A}) = (L_A \oplus F_A)(\bar{A}) \\ LP(A) = (L_A \oplus \Pi_A)(A) & e & LP(\bar{A}) = (L_A \oplus \Pi_A)(\bar{A}) \\ B^+ = F_A(B) & e & B^- = F_A(\bar{B}) \end{array}$$

Onde $B \in F_A$, $\forall A \in \mathcal{A}$

$$B^* = F_A(B \cup \bar{A})$$

5.3.2.2 DEFINIÇÃO

Calcula_λ: calcula a soma das crenças dos filhos de cada nó A e este resultado é somado com a crença do próprio nó A . Isto é, $\lambda = F_A \oplus L_A$, calculado para as proposições A e \bar{A} . A partir das folhas vamos subindo na árvore até chegarmos aos filhos da raiz. Em cada nó A , da árvore, calculamos $F(A)$ e $F(\bar{A})$ a partir de $\lambda(B)$ e $\lambda(\bar{B})$, para todo $B \in F_A$. Em seguida calculamos $\lambda(A)$ e $\lambda(\bar{A})$ a partir de $L(A)$, $L(\bar{A})$, $F(A)$ e $F(\bar{A})$.

Calcula_C: Todas as informações captadas a nível dos nós inferiores da árvore estão agora condensadas nos filhos diretos da raiz Θ . **Calcula_C** desce a árvore calculando a função de crença Π_A e combinando-a com as funções L_A e F_A para obter a crença de crença global C_A , para todo nó $A \in \mathcal{A}$.

Calcula_λ faz o percorrimento da árvore em profundidade. Enquanto o procedimento **Calcula_C** faz o percorrimento em largura.

```

F: Fila ; (* F é global *)
Procedimento Calcula_λ(A:nó) ;
  Se folha(A), então
    λ(A) ← L(A) ;
    λ(Ā) ← L(Ā)
  Senão
    Para todo B ∈ FA, faça
      Calcula_λ(B) ;
    FimPara ;
    Se A ≠ r, então (* r = ⊙, raiz da árvore *)
      (* Rotina 1 *)
      
$$K_A^{-1} \leftarrow 1 + \sum_{B \in F_A} \frac{\lambda(B)}{1-\lambda(B)} ;$$

      
$$P_A \leftarrow \prod_{B \in F_A} \frac{\lambda(\bar{B})}{1-\lambda(B)} ;$$

      F(A) ← 1 - KA ; F(Ā) ← KA · PA ;
      (* Rotina 2 *)
      
$$K^{-1} \leftarrow 1 - L(\bar{A})F(A) - L(A)F(\bar{A}) ;$$

      
$$\lambda(A) \leftarrow 1 - K \left[ 1 - L(A) \right] \left[ 1 - F(A) \right] ;$$

      
$$\lambda(\bar{A}) \leftarrow 1 - K \left[ 1 - L(\bar{A}) \right] \left[ 1 - F(\bar{A}) \right]$$

    FimSe ;
  FimSe ;
FimProcedimento Calcula_λ;

```

Figura 5.3 Procedimento Calcula_λ

No corpo do pseudo-código, colocamos, como documentação, a identificação das 6 subrotinas propostas por Shafer e Logan(1987).

O Procedimento Calcula_λ está ilustrado na figura 5.5 e o procedimento Calcula-C esta ilustrado na figur 5.6. Um corpo de programa, ilustrado na figura 5.7, é sugerido. A vantagem é que ele agrupa os dois procedimentos dando uma idéia da relação de antecendência e subsequência entre eles.

A rotina 5, proposta por Shafer & Logan, está inserida no procedimento Calcula_C. O Cálculo do fator de normalização, tanto da rotina 5 quanto da rotina 6 podem ser colocado fora da iteração varrendo todos os filhos de A. Nós as colocamos dentro da iteração para manter a integridade original destas rotinas. Contudo, é fácil ver que complexidade da rotina 5 é quadrática em $|F_A|$.

Procedimento Calcula_C ;

Enquanto $F \neq \phi$, faça

$A \leftarrow$ Primeiro(F) ;

 Se não Folha(A), então (* Expande o nó A *)

 Se $A \neq r$, então (* Calcule o Π para A , subrotina 4 *)

$$K^{-1} = \frac{1 - C(A)}{1 - F(A)} + \frac{1 - C(\bar{A})}{1 - F(\bar{A})} - \frac{1 - C(A) - C(\bar{A})}{1 - F(A) - F(\bar{A})} ;$$

$$\Pi(A) = 1 - K \frac{1 - C(A)}{1 - F(A)} ; \quad \Pi(\bar{A}) = 1 - K \frac{1 - C(\bar{A})}{1 - F(\bar{A})} ;$$

 Para todo $b \in F_A$, faça (* Visita todos os filhos de A *)

$$K^{-1} = 1 + \sum_{\substack{c \in F_A \\ c \neq A}} \frac{\lambda(c)}{1 - \lambda(c)} ; \quad (* \text{ subrotina 5 } *)$$

$$B^+ = 1 - K \left[1 + \sum_{\substack{c \in F_A \\ c \neq B}} \frac{\lambda(c)}{1 - \lambda(c)} \right] ; \quad B^- = 1 - K \frac{1 - \lambda(B)}{1 - \lambda(\bar{B})} ;$$

$$B^* = 1 - K \left[1 + \sum_{\substack{c \in F_A \\ c \neq B}} \frac{\lambda(c)}{1 - \lambda(c)} - \prod_{\substack{c \in F_A \\ c \neq B}} \frac{\lambda(\bar{c})}{1 - \lambda(\bar{c})} \right]$$

$$K^{-1} = 1 - F(A)\Pi(\bar{A}) - F(\bar{A})\Pi(A) ; \quad (* \text{ Subrotina 6 } *)$$

$$C(\bar{B}) = 1 - K(1 - \Pi(\bar{A}))(1 - B^-) ;$$

$$C(B) = K(\Pi(A) \cdot (B^* - F(\bar{A})) + (1 - \Pi(A) - \Pi(\bar{A}))B^+) ;$$

$F \leftarrow$ Enfileira(B, F) ;

 FimPara (* para todo $b \in F_A$ *)

Senão (* $A=r$, Subrotina 3, Calcula crença filhos raiz *)

$$K_A^{-1} \leftarrow 1 + \sum_{B \in F_A} \frac{\lambda(B)}{1 - \lambda(B)} - \prod_{B \in F_A} \frac{\lambda(\bar{B})}{1 - \lambda(\bar{B})} ;$$

 Para todo $b \in F_A$, faça

$$C(B) = 1 - K \left[1 + \sum_{\substack{B \in F_A \\ B \neq A}} \frac{\lambda(B)}{1 - \lambda(B)} \right] - \prod_{\substack{B \in F_A \\ B \neq A}} \frac{\lambda(\bar{B})}{1 - \lambda(\bar{B})}$$

$$C(\bar{B}) = 1 - K \frac{1 - \lambda(\bar{A})}{1 - \lambda(A)} ; \quad F \leftarrow \text{Enfileira}(B, F) ;$$

 FimPara ;

 FimSe ; (* Se $A \neq r$ *)

 FimSe (* Se não Folha(A) *)

 FimEnquanto (* $F \neq \phi$ *)

FimProcedimento ; (* Calcula_C *)

Figura 5.6 Procedimento Calcula_C

```

(* ..... Corpo Principal ..... *)
  Calcula_λ(r) ; (* Calcula as mensagens λ *)
  Esvazia(F) ; (* torna a fila Vazia *)
  Enfileira(r, F) ; (* Põe a raiz r de A na fila *)
  Calcula_C ; (* Calcula a crença final C *)

```

Figura 5.7 Corpo do Programa Shafer & Logan

6. CONCLUSÃO

Neste trabalho é apresetado um resumo da Teoria de Dempster-Shafer. Procuramos ilustrar com exemplos as idéias mais abstratas. Descrevemos uma estrutura para cada uma das diversas representações de crença: em termos de massa básica de crença (m), de função de crença (Cr), de plausibilidade (Pl) e de função comunitária (Q). Avaliamos os esforço de conversão entre estas representações. Mostramos como usar algumas propriedades das funções de crenças para minimizar o esforço de cálculo no processo de combiná-las. Mensuramos a complexidade de combinar funções de crenças das quais pelo menos uma é bayesiana, e finalizamos descrevendo os algoritmos de Barnett(1981) e Shafer & Logan (1987).

Todos os avanços que tornaram o esforço de combinação de evidências tratável em tempo polinomial foram feitos para casos particulares, ainda que importantes, tais como:

O algoritmo de Barnett(1981) que trabalha com uma partição do quadro de discernimento Θ e combina funções de crenças dicotômicas sobre esta partição, onde um dos elementos focais da dicotomia é um átomo da partição.

O algoritmo de Shafer-Logan(1987) que trabalha com uma hierarquia de proposições do quadro de discernimento, usa o algoritmo de Barnett para propagar as evidências dos níveis inferiores para os níveis superiores da hierarquia e a regra pura de Dempster para propagar as evidências coletadas nos níveis superiores para os níveis inferiores da árvore.

Os algoritmos propostos por Silva & Milidiú (1989a), ambos trabalhando sobre uma hierarquia de proposições, projetados segundo o paradigma de orientação a objetos e usando conceitos de família de quadros de discernimento compatíveis. Sendo que um é sequencial e o outro tem enfoque paralelo.

O algoritmo proposto por Silva & Milidiú (1989b), e também no presente texto, que combina uma lista de $k+1$ funções de crença, definidas em Θ , em um tempo de computação de ordem $O(k \cdot |\Theta|)$, desde que ao menos uma das funções seja Bayesiana.

E, por último, o algoritmo proposto por Silva & Milidiú (1989c) que combina uma lista de k funções de crença politômicas, definidas sobre uma partição comum Ω do quadro de discernimento, em tempo de ordem $O(k \cdot |\Omega|)$.

Ficam ainda em aberto propostas tratáveis mais gerais, e mesmo propostas restritivas que cubram as soluções das propostas acima, mas com melhor desempenho, além de outras específicas para determinadas aplicações.

A Teoria de Dempster-Shafer é muito nova, poucas aplicações reais foram feitas, o seu potencial ainda não foi convenientemente explorado. Um tópico de muito interesse, por exemplo, é a idéia de refinamento de quadros de discernimento, e implicações entre proposições de quadros condensados e refinados. Isto permitiria, por exemplo, resolver com menor esforço que o algoritmo de Shafer-Logan o problema de Gordon e Shortlife. Silva e Milidiú (1989a) propõem um outro algoritmo, explorando mais profundamente os conceitos de refinamento e condensação, para resolver o problema de hipóteses hierarquizadas, cuja complexidade é menor que o algoritmo de Shafer e Logan (1987). Uma conexão entre as teorias de Dempster-Shafer e Lógica Nebulosa poderia ser feita através do estudo de funções de suporte consonantes, sucintamente mencionadas neste texto.

BIBLIOGRAFIA

- BARNETT, J.A. Computational methods for a mathematical Theory of evidence. in: *Proceeding of IJCAI'81*:868-875, 1981.
- BHATNAGAR, R. K. & KANAL, L. N. Handling uncertain information: A review of numeric and Non-numeric Methods. *Uncertain in Artificial Intelligewnce*, L. N. Kanal and J.F.Lemmer(eds.):3-26, Elsevier Science, North-Holland, 1986.
- DEMPSTER, A. P. On direct probabilities. *J. Royal Statis. Soc, Ser. B* 25:102-107, 1963.
- DEMPSTER, A. P. New methods for reasoning toward posterior distributions based on sample data. *Ann. Math. Statis.*, 37:355-74, 1967a.
- DEMPSTER, A. P. Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Statis.* 38:325-39, 1967b.
- DEMPSTER, A. P. Upper and lower probability inferences based on a sample from a finite univariant population. *Biometrika*, 54:515-28, 1967c.
- DEMPSTER, A. P. Upper and lower probabilities generated by a random closed interval. *Ann. Math. Statis.* 39(3):957-56, 1968a.
- DEMPSTER, A. P. A generalization of bayesian inference. *J. Roy. Statis. Soc. Ser B* 30:205-47, 1968b.
- DEMPSTER, A. P. Upper and lower probability inferences for families of hypoteses with monotone density ratios. *Ann Math. Statis.* 40:953-69, 1969.
- DUBOIS, D. & PRADE, H. Combination and propagation of uncertainty with belief functions. *Proceedings of IJCAI'85*: 111-13, 1985.
- GARVEY, T.D. et alli. An inference technique for integrating knowledge from disparete sources. *Proceeding of IJCAI'81*:319-25, 1981.
- GORDOM, J. & SHORTLIFE, E.H. The Dempster-Shafer theory of evidence. *Rule-based expert systems*, B.G.Buchanan and E.H. Shortlife(eds.):272-94, California, Addison-Wesley, 1984.
- GORDOM, J. & SHORTLIFE, E.H. A method for managin evidencial reasoning in hierarchical hypotesis space. *Artificial Intelligence*, 26(3):323-57, Jul 1985.
- LEE, Chia-Hoang. A Comparison of two evidencial reasoning schemes. *Artificial Intelligence*, 35(1):127-34, May 1988.
- LEMMER, J.F. Confidence factors, empirism and Dempster-Shafer theory of evidence. *Uncertainty in Artificial Intelligence*, L.N.Kanal and J.F.Lemmer(eds.):117-26, Elsevier Science,

- North-Holland, 1986.
- PEARL, Judea. Probabilistic Reasoning in Intelligent Systems: Network of Plausible Inference. Morgan Kaufmann Publisher, San Mateo, CA, 1988.
- SENNE, E.L.F. Uma abordagem para o tratamento de incerteza em sistemas especialistas. Anais do CNMAC'87:846-51, Set 1987.
- SHAFER, G. A mathematical theory of evidence. Princeton University Press, 1976.
- SHAFER, G. Probability judgment in artificial intelligence. Uncertainty in artificial intelligence. L.N.Kannal and J.F.Lemmer(eds.):127-36, Elsevier Science, North-Holland, 1986.
- SHAFER, G. & LOGAN, R. Implementing Dempster's Rule for hierarchical evidence. Artificial Intelligence, 33:271-98, 1987.
- SILVA, W.T.da & MILIDIÚ, R.L. Algoritmos para Combinação de Evidências em Espaço de Proposições Hierarquizado. Relatório de Pesquisa, DI-PUC/RJ, Rio 1989a.
- SILVA, W.T.da & MILIDIÚ, R.L. Combinando Funções de Crença com uma Função Bayesiana. Sumetido ao XXII SBPO, Fortaleza, Out. 1989b.
- SILVA, W.T.da & MILIDIÚ, R.L. Combinado Funções de Crença Politômicas. Submetido ao SBIA, xxxxx, Rio, 1989c.
- SILVA, Wagner T.da & SCHWABE, D. Metodos de propagar incerteza em sistemas especialistas. Relatório técnico, INF/PUC-Rio, Rio, 1988.
- SMETS, Philippe. Belief functions versus probability functions. Lecture Notes in Computer Science, vol 313:17-24, 1988.
- THOMPSON, T.R. Parallel formulation of evidential-reasoning theories. Proceedings of IJCAI'85:321-27, 1985.
- ZADEH, L.A. A simple view of the Dempster-Shafer theory of evidence and its implication. The AI Magazine, Summer:85-90, 1986.