

Integração de evidências em redes credais e a regra de Jeffrey

José Carlos Ferreira da Rocha ¹

Alaine Margarete Guimarães ¹

Cassio Polpo de Campos ²

Abstract: As redes credais provêm um esquema para a representação de modelos probabilísticos imprecisos. Os algoritmos de inferência usualmente empregados em redes credais computam o intervalo da probabilidade posterior de um evento de interesse dadas evidências do tipo específica - evidências que descrevem o estado atual de um conjunto de variáveis. Estes algoritmos não realizam raciocínio evidencial no caso em que as evidências devem ser processadas segundo a regra de condicionamento proposta por R.C. Jeffrey. Considerando isto este artigo descreve um procedimento para integrar evidências com a regra de Jeffrey quando da realização de inferências em redes credais.

Abstract: Credal networks provide a scheme for dealing with imprecise probabilistic models. The inference algorithms often used in credal networks compute the interval of the posterior probability of an event of interest given evidence of the specific kind - evidence that describe the current state of a set of variables. These algorithms do not perform evidential reasoning in case of the evidence must be processed according to the conditioning rule proposed by RC Jeffrey. This paper describes a procedure to integrate evidence with Jeffrey's rule when performing inferences with credal nets.

1 Introdução

A eficácia de um sistema baseado em conhecimento depende do procedimento de inferência empregado para derivar decisões a partir das informações contidas na base de conhecimento - o módulo do sistema que armazena um modelo formal do conhecimento do domínio da aplicação [18]. Infelizmente, todo modelo simplifica alguns aspectos da realidade; seja porque a complexidade do domínio da aplicação impossibilita o desenvolvimento de um modelo completo, seja porque o conhecimento do domínio é limitado. Estas restrições podem resultar em um modelo cuja a estrutura ou os parâmetros são incertos, uma condição que demanda o emprego de procedimentos capazes de tratar a incerteza durante a execução de inferências [21].

¹Departamento de Informática, UEPG, Campus de Uvaranas
{jrocha@uepg.br, alainemg@hotmail.com}

²IDISIA, Lugano, Suiça

Neste contexto, formalismos baseados na teoria de probabilidades têm sido frequentemente empregados no desenvolvimento de sistemas que precisam raciocinar sob condições de incerteza, em particular, o formalismo das redes bayesianas [19]. Uma das razões para o sucesso deste formalismo é o fato dele permitir codificar modelos multivariados usando um esquema de representação que, ao explorar as relações de independência entre variáveis, facilita a especificação da base de conhecimento. Dada uma rede bayesiana é possível computar a probabilidade posterior de qualquer evento definido sobre as variáveis da rede - este procedimento é chamado de *atualização de crença*. Uma vantagem da atualização de crença em redes bayesianas é que os algoritmos que a implementam são capazes de explorar as relações de independência expressas na rede para reduzir o esforço computacional do processo.

As atuais implementações dos algoritmos de atualização de crença manipula dois tipos de evidência: *específica* e *virtual* [3, 7, 1]. Uma evidência específica informa o estado de uma variável ao passo que uma evidência virtual associa os estados da variável com medidas de verossimilhança. Estes tipos de evidência não capturam a situação em que um agente *A* deseja atualizar suas crenças finais de forma a torná-las consistentes com a crença relatada por um agente *B*. Neste caso, uma possibilidade é usar a regra de condicionamento de Jeffrey. Esta regra assume que a evidência contém uma informação probabilística (distribuição marginal ou condicional) que descreve a crença de *B* a respeito do estado de uma variável ou de um conjunto de variáveis *E* e que o agente *A* pretende ajustar seu modelo interno de forma que as probabilidades relativas a *E* sejam iguais àsquelas reportadas por *B*. Assim, o procedimento de inferência deve: (a) ajustar o modelo probabilístico de forma que ele concorde com a evidência e (b) executar a atualização de crença para calcular as probabilidades de interesse no modelo ajustado.

As redes credais objetivam estender o formalismo das redes bayesianas de maneira a possibilitar a representação de modelos probabilísticos imprecisos [9] [20]. Basicamente, uma rede credal é um grafo acíclico e dirigido cuja topologia destaca relações de dependência condicional direta entre as variáveis do modelo - os nós indicam variáveis e os arcos os relacionamentos. Uma rede credal também é composta por uma coleção de conjuntos de distribuições de probabilidades condicionais - *conjuntos credais* [15] - os quais podem ser combinados em um único conjunto credal de distribuições conjuntas denominado *extensão da rede*. A partir da extensão, que é representada de forma implícita pela rede credal, é possível calcular intervalos de probabilidade para eventos definidos sobre as variáveis da rede.

Os algoritmos usados atualmente no cômputo de inferências em redes credais não exploram a regra de condicionamento de Jeffrey. Considerando isto, este artigo apresenta um procedimento que executa este tipo de inferência em redes credais. O algoritmo proposto estende o procedimento de programação multilinear apresentado por Campos e Cozman [11] para inferência em redes credais. Para tanto, o procedimento descrito aqui emprega a transformação de Chan e Darwiche [4] para converter evidências que devem ser processadas segundo

a regra de Jeffrey em equações multilíneas que são adicionadas àquelas do programa multilíneo.

Este documento é organizado como segue. A Seção 2 descreve o formalismo das redes bayesianas e alguns conceitos sobre o raciocínio baseado em evidências. A Seção 3 introduz o formalismo das redes credais. A Seção 4 apresenta o algoritmo de inferência proposto. A Seção 5 ilustra o funcionamento deste algoritmo com dois exemplos. A última seção contém as considerações finais do trabalho.

2 Redes bayesianas, raciocínio sob incerteza e a regra de Jeffrey

Uma rede Bayesiana \mathcal{N} é uma tripla (G, X, P) na qual $X = \{1, \dots, n\}$ é um conjunto de variáveis aleatórias e G é um grafo acíclico e dirigido que expressa dependências entre variáveis [19]. Cada nó de G representa uma variável aleatória em X e cada arco um relacionamento de dependência condicional direta entre as variáveis conectadas. Seja i os nós pais de i no grafo e r_i o número de eventos atômicos que podem ser definidos sobre i , o componente P é uma coleção de funções de distribuição locais da forma $i|i$. As funções contidas em P são armazenadas em seus respectivos nós. Assim, se i é uma raiz da rede, ele armazena a distribuição i , caso contrário, ele armazena as distribuições condicionais $i|i_1 \dots i_{r_i}$, em que i_j simboliza o j -ésimo evento atômico dos pais de i .

Seja i o conjunto dos descendentes de i no grafo. O formalismo das redes bayesianas assume a seguinte condição de Markov: i é condicionalmente independente de $X \setminus i \cup \{i\}$ dados os estados das variáveis em i , para todo $i \in X$. Dada esta condição, a estrutura de uma rede bayesiana codifica uma distribuição de probabilidades conjuntas X que pode ser recuperada pela expressão:

$$X = \prod_{i=1}^n X_i | X_i.$$

Uma consulta sobre uma variável q de uma rede bayesiana objetiva determinar a distribuição marginal ou posterior de q . Os procedimentos que computam este tipo de informação são ditos métodos de atualização de crença e neste trabalho considera-se o emprego do algoritmo de eliminação de variáveis [24]. Outros métodos podem ser consultados em [19].

O método de eliminação de variáveis pode ser abstraído em duas etapas básicas: a inicialização de fatores e a eliminação de variáveis propriamente dita. A inicialização de fatores constrói uma estrutura auxiliar que facilita a execução do processo de eliminação. Ela começa com o emprego de um procedimento que ordena as variáveis não consultadas e não observadas - a ordem de eliminação. Em seguida cada variável i é associada a uma estrutura de dados chamada *fator* e que é denotada por B_i . Inicialmente cada B_i armazena as funções

probabilísticas $j|j$ tal que $j \in \{i\} \cup i$, sendo i são os filhos de i no grafo. Deve ser observado que:

- a sequência de inicialização dos fatores segue a ordem de eliminação;
- cada função é armazenada no primeiro fator B_i tal que i aparece entre as variáveis que definem o domínio de .

A etapa de eliminação inicia com o processamento do fator B_{i_1} , associado à variável i_1 - a primeira na ordem de eliminação. A eliminação é obtida pelo cômputo da expressão a seguir [6]:

$$i_1|i_1, i_1 = \sum_{i_1} \prod_{j \in \{i_1\} \cup i_1} j|j. \quad (1)$$

O lado esquerdo da Equação 1 é chamado de separador de B_{i_1} e denota uma função não normalizada da qual i_1 foi eliminada. Aqui, i_1 indica os nós cônjuges de i_1 no grafo da rede. Após esta operação o separador é armazenado no fator B_{i_w} . A escolha do fator B_{i_w} segue a ordem de eliminação, isto é, o índice i_w vem depois de i_1 na ordenação dos fatores e B_{i_w} é o primeiro fator em que uma das variáveis de i_1 aparece.

Concluída a atualização das funções que compõe B_{i_w} , o procedimento segue eliminando variáveis até que o último fator seja atingido. Quando isto acontece o último separador é armazenado em um novo fator B_q que é então criado e associado a q . Finalmente, a probabilidade desejada pode ser calculada pela inspeção direta de B_q . Deve ser observado que a eliminação de variáveis é aplicada ao subconjunto das variáveis da rede que são relevantes para a consulta (obtidas com o uso de algoritmos exploram a topologia do grafo da rede para determinar quais variáveis da rede independem de q [19, 17]). O Algoritmo 1, descreve o método de eliminação de variáveis como descrito por Cozman (2000,ref. [4]).

2.1 Tratamento de evidências

Pearl [19] considera duas formas de evidências no processamento de inferências em redes bayesianas. A evidência específica, que descreve com certeza o estado observado de uma variável k e a evidência virtual que expressa incerteza sobre o valor observado. Assim, se a variável k tem m_k categorias e seu valor observado é kl , uma evidência específica reporta este fato como um vetor $\phi(k) = [\phi_1, \dots, \phi_m]$ no qual $\phi_j = 1$ se $j = l$ e 0 caso contrário. Em um relatório de evidência virtual, o vetor $\phi(k)$ especifica uma função de probabilidades tal que $\phi_j = kj|kj$, para cada $j = 1 \dots m$. Aqui k é uma variável auxiliar que codifica a observação dos valores de k . Ambas as formas de evidência podem ser processadas pelo procedimento de eliminação de variável inserindo $\phi(k)$ em um fator que contenha k .

• **Algoritmo 1** : Eliminação de variáveis

- entrada: \mathcal{N}, q ;
 - definições: sejam $X^r \subset X$ o conjunto das variáveis irrelevantes e $X^e \subset X$ o conjunto das variáveis observadas;
1. defina uma ordem entre as variáveis de $X \setminus (\{q\} \cup X^e \cup X^r)$;
 2. para cada variável i crie um fator B_i - processe as variáveis segundo a ordem de eliminação;
 3. para cada fator B_i ; segundo a ordem de eliminação;
 - (a) insira todas as distribuições de P em que i aparece no fator B_i ;
 - (b) multiplique as distribuições de B_i ;
 - (c) calcule o separador de B_i ; para isto execute uma somatória com respeito à variável i no resultado obtido no passo anterior;
 - (d) se B_i não é o último fator segundo a ordem de eliminação, armazene o seu separador no fator B_{i_w} ;
 4. armazene o último separador calculado em um fator B_q ;
 5. normalize as funções em B_q ; o resultado obtido é a distribuição a priori ou posterior de q dependendo se $X^e \equiv \emptyset$ ou não.

A regra de Jeffrey estabelece uma outra maneira para abordar a integração de evidências em modelos probabilísticos [22]. Esta regra assume que a informação não está diretamente relacionada com a observação, como a evidência específica, nem com medidas de verossimilhança, como a evidência virtual, mas se refere às crenças finais do agente de decisão. Por exemplo, seja o fato de que o agente dispõe de uma distribuição conjunta X a qual implica que a distribuição marginal de i é i . Além disso, o agente dispõe de um relatório fornecido por um agente externo que é um especialista em fazer predições sobre i . O relatório informa que a distribuição correta para i é i . Em tal situação o primeiro agente poderia estar interessado em ajustar seu modelo interno de modo que suas crenças sobre i concordassem com i . A regra de Jeffrey estabelece que a distribuição ajustada pode ser calculada com a expressão:

$$X = X - \{i\}|i \quad (2)$$

Deve ser observado que a Equação 2 assume que a distribuição $X - \{i\}|i$ não é afetada pela evidência.

É interessante destacar que é possível codificar um relatório de evidência na forma proposta por Jeffrey como uma evidência virtual [4]. Basta computar cada ϕ_j como

$$k \cdot i_j / i_j,$$

sendo i_j a probabilidade original de i_j em X e k uma constante.

Apesar da existência de uma regra de conversão deve ser enfatizado que esses tipos de evidência não são equivalentes. A evidência virtual é uma informação probabilística sobre uma observação enquanto aquela tratada por Jeffrey estabelece limites sobre a crença final do agente. Elas também diferem em termos de comutatividade [12]. Evidência virtual é comutativa e, portanto, a atualização da crença com a evidência virtual independe da ordem em que observações são adquiridas. O mesmo não acontece quando se emprega a regra de Jeffrey.

3 Redes credais

Um conjunto credal i definido sobre uma variável aleatória discreta i é um conjunto de distribuições de probabilidades definidas sobre i [15]. Neste trabalho é assumido que um conjunto credal é um polítopo representado por seus vértices (distribuições extremas) ou por uma coleção de restrições lineares sobre as medidas de probabilidade. Dado o conjunto de variáveis X , um conjunto credal X é composto por distribuições conjuntas X . As operações de marginalização e condicionamento de conjuntos credais podem ser realizadas da seguinte maneira:

- marginalização - seja $i \in X$, calcule i para cada distribuição extrema de X ; determine o casco convexo das funções resultantes [14]; o resultado obtido é o conjunto credal marginal i ;
- condicionamento - dado o evento y definido sobre $Y \subset X$ calcule a distribuição condicional³ $X \setminus Y|y$ em cada vértice de X ; compute o casco convexo das distribuições resultantes; o conjunto credal condicional é indicado por $X \setminus Y|y$.

Sejam Y e Z dois subconjuntos disjuntos de X , a informação condicional de Y dado Z pode ser organizada em diferentes maneiras [16]. Este trabalho considera que este tipo de informação é dada na forma de uma coleção de conjuntos credais especificados separadamente $YZ = \{Y|z_0, \dots, Y|z_t\}$ em que z_k é um evento conjunto sobre Z . Isto é, YZ é uma coleção que contém um conjunto credal condicional definido sobre Y , para cada conjunção das variáveis em Z .

³Aquí supõe-se que a probabilidade dos eventos condicionantes é positiva.

A complexidade de espaço e a complexidade de tempo para realização de inferências em modelos multivariados dificultam o uso da teoria dos conjuntos credais em aplicações reais [20]. Para atenuar essas dificuldades, o formalismo de redes credais [9] explora relações de independência para representar de forma implícita um conjunto credal multivariado. Basicamente, uma rede credal \mathcal{C} é um grafo acíclico e direcionado em que cada nó i simboliza uma variável e armazena uma lista de conjuntos credais locais que associam i com seus pais. Este trabalho considera redes credais com conjuntos credais especificados separadamente, assim o nó i está associado a ii .

A Figura 1 mostra a topologia de uma rede credal multiconectada. As coleções de conjuntos credais desta rede são 1, 21, 31, 42, 3 e 54.

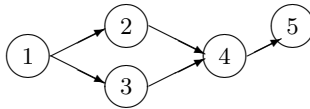


Figure 1. Uma rede credal simples.

É importante notar que existem várias interpretações para o conceito de independência condicional na teoria dos conjuntos credais. Assim, na prática é necessário selecionar aquela que é mais adequada para a aplicação desejada [5]. Este trabalho considera a tarefa de processar a integração de evidências em redes credais que assumem o conceito de independência forte [8] usando a regra de Jeffrey. Duas variáveis X e Y são fortemente independentes quando X e Y são independentes em cada vértice de XY . Isto é, dado XY , cada vértice satisfaz $X|Y = X$ e $Y|X = Y$. Uma rede credal expressa relações de independência forte se para cada uma de suas variáveis i tem-se que i é fortemente independente de seus nós não pais e de seus não descendentes quando o estado de seus pais é conhecido. O maior conjunto credal multivariado que atende a esta restrição da rede é dito uma extensão forte.

Uma inferência em uma rede credal calcula limites para a probabilidade de algum evento de interesse na extensão associada. Se qa é o evento de interesse e \mathbf{E} representa a evidência, o objetivo de uma inferência é determinar o intervalo $[qa|\mathbf{E}, qa|\mathbf{E}]$ em que $qa|\mathbf{E}$ e $qa|\mathbf{E}$ são chamados probabilidades inferiores e superiores, respectivamente. Campos e Cozman [11] apresentaram um algoritmo para inferência exata e aproximada em redes credais. Inicialmente, o algoritmo constrói um problema de programação multilinear [13] cujas restrições são obtidas por um procedimento que: (a) explora o algoritmo de eliminação de variáveis para gerar expressões simbólicas relativas às restrições multilineares associadas a uma inferência; (b) enumera expressões que são oriundas dos axiomas da teoria da probabilidade; e (c) enumera as restrições definidas pelos conjuntos credais especificados na rede.

O Algoritmo 2, apresentado a seguir, descreve o procedimento que gera o programa multilinear. Neste algoritmo, $w_{a_j^i}$ representa a_j^i , sendo a_j^i um evento atômico definido sobre as variáveis de A_i ; $w_{a_u^j|b_v^j}$ representa $A_j = a_u^j|B_j = b_v^j$; S é uma coleção de restrições; inicialmente vazia; e X^e é o conjunto de variáveis observadas e Z como $\{q\} \cup X^e \cup X^r$. Novamente assume-se que as variáveis irrelevantes foram removidas da rede.

- **Algoritmo 2:** gera o programa multilinear relativo a uma inferência em redes credais
 - entrada: \mathcal{C}, qt ;
 - 1. ordene as variáveis de $X \setminus Z$;
 - 2. para cada variável i gere um fator B_i ;
 - 3. para cada B_i segundo a ordem de eliminação faça;
 - (a) insira em B_i todas as funções probabilísticas em que i aparece e as renomeie como $f_1(A_1|B_1) \dots f_d(A_d|B_d)$; observe que $i \in A_j \cup B_j$ e $A_j \cap B_j \equiv \emptyset$ para $j = 1..d$; estas funções são aquelas associadas aos conjuntos credais e nós da rede ou aquelas oriundas do processo de eliminação de variáveis.
 - (b) para cada $f_j(A_j|B_j)$:
 - para cada b_v^j , adicione a restrição $\sum_u w_{a_u^j|b_v^j} = 1$ em S ;
 - (c) sejam $A_s \cup B_s$ as variáveis do separador de B_i e $w_{a_u^s|b_v^s} = a_u^s|b_v^s$; para cada a_u^s e b_v^s acrescente $\sum_{i=1}^{m_i} \prod_{j=1}^d w_{a_u^j|b_v^j} = w_{a_u^s|b_v^s}$ as restrições em S ;
 - 4. descreva os conjuntos credais de \mathcal{C} como inequações lineares e acrescente-as à S [2];
 - 5. repita o processo acima para B_q ;
 - 6. selecione a variável $w_{q,t}$ em B_q como a função objetivo a ser minimizada/maximizada.

Na última etapa o máximo e mínimo da probabilidade de interesse são obtidos pela solução do problema de programação multilinear. Para tanto emprega-se um resolvidor do tipo ramificar-e-limitar [23].

4 Inferência em redes credais usando a regra de Jeffrey

Esta seção apresenta um procedimento para o cômputo de intervalos de probabilidades em redes credais dadas evidências que devem ser incorporadas ao modelo segundo a regra de Jeffrey. O procedimento proposto é descrito no Algoritmo 3. Ele inicia com uma rede credal \mathcal{C} e um conjunto de evidências o_1, \dots, o_s que estão na forma de distribuições marginais definidas sobre algumas das variáveis de X . É importante destacar que este procedimento assume as seguintes condições: (a) cada evidência é independente da outra e (b) as evidências devem ser integradas de forma simultânea. A segunda condição evita a necessidade de se considerar dificuldades advindas da não comutatividade da regra de condicionamento de Jeffrey.

O objetivo é calcular o intervalo $[qa, qa]$ sujeito as restrições impostas por este tipo de evidência probabilística. Sem perda de generalidade, o Algoritmo 3 é apresentado para o caso em que se deseja computar qa . Como pode ser visto, ele é uma extensão do Algoritmo 2 na qual a idéia básica é explorar as relações multilineares da rede credal e acrescentar a elas aquelas que são provenientes da evidência.

- **Algoritmo 3** : Inferência multilinear com a regra de Jeffrey

- entrada: a rede credal \mathcal{C} , o evento qa e as evidências o_1, \dots, o_s ;
 - saída: qa
1. empregue o Algoritmo 2 para gerar o programa multilinear referente ao cômputo de qa ; proceda como se cada variável o_i estivesse associada a uma evidência virtual; isto é, acrescente as expressões simbólicas como se φ_{o_i} estivesse presente;
 2. para cada o_i
 - (a) para cada evento $o_i j$ adicione a restrição $\varphi_{o_i j} \cdot o_i j = o_i j$ ao conjunto S ;
 - (b) gere as restrições de normalização $\sum_j o_i j = 1$ e insira-as na coleção S ;
 - (c) para cada categoria $o_i j$ de o_i use o Algoritmo 3.1, descrito a seguir, para gerar as restrições associadas a $o_i j$;
 3. resolva o problema multilinear minimizando a função objetivo $w_{q,t}$.

A seguir são analisadas as principais diferenças entre este algoritmo e o Algoritmo 2. Para começar o primeiro passo do Algoritmo 3 faz uma chamada ao Algoritmo 2 para que este gere um programa multilinear. Esta chamada é acrescida com relatórios de evidência virtual, um para cada o_s . Deve-se observar que neste ponto cada elemento de φ_{o_i} é uma variável ilimitada, entretanto o passo 2a resolve isto ao conectá-la a o_s com o emprego da transformação de Chan e Darwiche [4]. A utilização da transformada de Chan e Darwiche

insere o termo o_{ij} cujas restrições de normalização são tratadas no passo 2b.

A probabilidade o_{ij} também é limitada por restrições que podem ser obtidas a partir da rede credal. O Algoritmo 3.1 descreve uma subrotina que gera as inequações e equações de o_{ij} . Este algoritmo é uma versão modificada do Algoritmo 2 e, diferentemente daquele, quando executado sobre o_{ij} não produz um problema de otimização para esta probabilidade. Em vez disto retorna somente as restrições que a rede credal impõe sobre o_{ij} .

- **Algoritmo 3.1:** gera as restrições multilineares associadas às evidências.

– entrada: \mathcal{C}, o_{ij} ;

1. execute o Algoritmo 2 para o_{ij} , contudo não gere um problema de otimização, para isto basta saltar o último passo do Algoritmo 3;
2. selecione a variável $w_{o_{i,j}}$ em B_{o_i} e adicione a expressão $w_{o_{i,j}} = o_{ij}$ em S .

5 Exemplos de aplicação

Esta seção ilustra os métodos descritos na seção anterior com dois exemplos. O primeiro exemplo ilustra o processo de forma simbólica. Nele supõe-se que, dada a rede da Figura 1, um agente deseja usar a regra de Jeffrey para computar 50 condicionada pela evidência 2.

O primeiro passo do Algoritmo 3 estipula que deve ser executada uma chamada do Algoritmo 2 para gerar a função objetivo e as restrições do programa multilinear associado a 50. Nesta chamada deve ser “simulada” a presença de um relatório de evidência virtual $\varphi(2)$. O resultado desta chamada é problema multilinear:

$$50 = \min 50 : s.a. S, \quad (3)$$

Alguns elementos do conjunto de restrições S são:

$$50 = \sum_{i=0;1} 50|4i \cdot 4i;$$

$$4j, 3k = \sum_{i=0;1} 4j|2i, 3k \cdot 2i, 3k \cdot \varphi_{2i}.$$

$$2j, 3k = \sum_{i=0;1} 2j|1i3j|1i \cdot 1i;$$

$$4j = \sum_{i=0;1} 4l, 3i;$$

No passo 2.a o Algoritmo 3 associa o termo φ_{2i} com a evidência 2 usando a transformada de Chan e Darwiche. Isto produz a restrição em $\varphi_{2j} \cdot 2j = 2j$ em X , para $j = 0, 1$. Este passo também introduz um novo termo no problema, a variável $2j$, cujas restrições de normalização são tratadas no passo 2.b. Esta variável do problema multilinear é limitada por uma série de equações que são extraídas da rede credal pelo Algoritmo 3.1, entre elas a expressão:

$$2j = \sum_{i=0;1}^{1;1} 2j|1i3k|1i \cdot 1i.$$

O último passo do Algoritmo 3 chama a rotina de ramificar-e-limitar empregada por Campos and Cozman [11] para resolver as inferências em redes credais.

Para manter a simplicidade do exemplo as restrições de normalização associadas às distribuições de probabilidades e as restrições oriundas da especificação dos conjuntos credais da rede não foram enumeradas neste exemplo. Assim, se neste exemplo \underline{l} e \bar{l} são os limites inferior e superior informados para 10 então as equações $10 + 11 = 1$, $10 \geq \underline{l}$ e $10 \leq \bar{l}$ devem ser adicionadas ao programa multilinear. De forma similar a restrição de normalização $\sum_{j=0}^1 \sum_{k=0}^1 2j, 3k = 1$, sobre os termos referentes a 2, 3 nas restrições que definem o programa multilinear, também deve ser inserida em S . Deve ser notado que o algoritmo original de Campos e Cozman especifica a inserção destas expressões no programa multilinear. Também deve ser observado que alguns símbolos podem aparecer mais de uma vez durante a geração das restrições multilineares - neste exemplo, o símbolo 20 aparece duas vezes. Portanto, é necessário manter um registro das variáveis que compõe o programa multilinear de tal forma que cada variável w_* seja associada a um único símbolo de medida de probabilidades $*$ e vice-versa.

O segundo exemplo ilustra o impacto deste tipo de processamento em termos numéricos. O objetivo é computar 10 na rede credal da Figura 2. Nesta rede a variável 1 simboliza a proposição *as roupas que estão no interior da caixa são coloridas* e a variável 2 representa a proposição *as roupas serão vendidas rapidamente*. O arco que conecta estas duas variáveis indica que o fato das roupas serem coloridas influencia o resultado das vendas. Ambas as variáveis são proposicionais e seus valores *0 and *1 denotam as constantes *verdadeiro* e *falso*, respectivamente. Os conjuntos credais⁴ da rede são listados abaixo e especificam a crença do agente no que se refere à distribuição marginal de 1 e às funções de probabilidades condicionais de 2 dados os valores de 1, conforme informado por um especialista hipotético.

$$\begin{aligned}
 1 &= CC(\{(0.7, 0.3); (0.6, 0.4)\}) \\
 2|10 &= CC(\{(0.9, 0.1); (0.5, 0.5)\}) \\
 2|11 &= CC(\{(0.3, 0.7); (0.6, 0.4)\})
 \end{aligned}$$

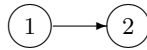


Figure 2. Rede credal do segundo exemplo.

A evidência a ser integrada pela regra de Jeffrey é fornecida por um especialista em vendas que inspecionou as roupas antes do embarque e declarou que $2 = (0.8, 0.2)$. O objetivo é calcular $10 = \min 10$ dada a evidência. O primeiro passo do Algoritmo 3 gera as restrições lineares oriundas dos conjuntos credais da rede e as expressões multilineares provenientes do procedimento de eliminação de variáveis, entre elas:

$$10 = \sum_{j=0}^1 2^j | 10 \cdot 10 \cdot \varphi_{2j}.$$

Aqui 10 indica o valor de 10 atualizado pela evidência.

Os passos 2a e 2b adicionam as restrições

$$\varphi_{20} \cdot 20 = 20,$$

$$\varphi_{21} \cdot 21 = 21$$

e

$$20 + 21 = 1$$

⁴ $CC(\cdot)$ denota o operador de casco convexo.

ao conjunto R .

O passo 2c adiciona expressões multilineares para 20 e 21. Depois deste processamento tem-se:

$$\begin{aligned}
 10 &= \sum_{j=0}^1 2j | 10 \cdot 10 \cdot \varphi_{2j}; & 20 &= 0.8; \\
 & & 20 &= \varphi_{20} \cdot 20; & 21 &= \varphi_{21} \cdot 21; \\
 20 &= \sum_{j=0}^1 20 | 1j \cdot 1j; & 21 &= 0.2; \\
 21 &= \sum_{j=0}^1 21 | 1j \cdot 1j; & 20 + 21 &= 1;
 \end{aligned}$$

O último passo do Algoritmo 3 executa um procedimento de otimização para obter $10 = \min 10$. Dadas as evidências fornecidas neste segundo exemplo a resposta encontrada é $10 = 0,666$. Este valor difere daquele que é obtido com o conhecimento *a priori*, $10 = 0,6$.

6 Considerações finais

Este trabalho apresentou um método para integrar evidências probabilísticas em redes credais segundo a regra de Jeffrey. Com o uso desta regra um agente de software que empregue uma rede credal para tomar decisões em modelos probabilísticos imprecisos pode atualizar os conjuntos credais da rede de maneira que a extensão da mesma concorde com as distribuições de probabilidade definidas sobre algumas variáveis da rede. Estas informações poderiam ser fornecidas por outro agente cujas funcionalidades lhe permitam a aquisição de informações mais precisas sobre algumas variáveis do sistema. Após a obtenção das informações externas, o primeiro agente pode usar os algoritmos descritos neste trabalho para calcular a probabilidade posterior das demais variáveis do seu modelo. Este tipo de inferência não é implementada pelos algoritmos atuais.

A abordagem proposta explora o fato de que este tipo de informação pode ser relacionada com relatórios de evidência virtual pelo emprego da transformação de Chan e Darwiche. Este procedimento gera uma série de restrições multilineares que são adicionadas àquelas que compõem a extensão forte da rede. Na sequência os intervalos de probabilidade dos eventos de interesse podem ser obtidos com a solução de um problema de otimização multilinear.

A principal contribuição deste trabalho é explorar o emprego do condicionamento de Jeffrey em redes credais, uma estratégia de condicionamento que pode ser usada na execução

de raciocínio baseado em evidências em sistemas multiagentes. Os algoritmos apresentados utilizam métodos para inferência em redes credais que codificam extensões fortes. Os procedimentos desenvolvidos exploram os algoritmos de inferência exata mais eficientes disponíveis. Apesar disto, a complexidade computacional da inferência - o problema é NP^{PP} -completo [10] - dificulta o desenvolvimento de aplicações com muitas variáveis ou cujos grafos sejam densos. Assim, um tópico de pesquisa a ser abordado no futuro é o desenvolvimento de algoritmos mais rápidos para inferência de intervalos aproximados.

7 References

- [1] A. Antonucci. *Imprecise Probabilistic Graphical Models: Equivalent Representations, Inference Algorithms and Application*. PhD thesis, Faculty of Informatics of the University of Lugano, Lugano, Italy, Abril 2008.
- [2] D. Bremner, K. Fukuda, and A. Marzetta. Primal-dual methods for vertex and facet enumeration. In J.-D. Boissonnat, editor, *13th ACM Symposium on Computational Geometry*, pages 49–56, Nice, France, June 1997.
- [3] A. Cano and S. Moral. A review of propagation algorithms for imprecise probabilities. In G. Cooman, F. Cozman, S. Moral, and P. Walley, editors, *1st International Symposium on Imprecise Probabilities and Their Applications*, pages 51–60, Ghent, Belgium, 29 June - 2 July 1999 1999.
- [4] H. Chan and A. Darwiche. On the revision of probabilistic beliefs using uncertain evidence. *Artificial Intelligence*, 163(1):67–90, 2005.
- [5] I. Couso, S. Moral, and P. Walley. A survey of concepts of independence for imprecise probabilities. *Risk, Decision and Policy*, (5):165–185, 2000.
- [6] F. Cozman. Generalizing variable elimination. In *Workshop on Probabilistic Reasoning in Artificial Intelligence*, pages 1–12, Atibaia, Brazil, 2000.
- [7] F. Cozman. Graphical models for imprecise probabilities. *International Journal of Approximate Reasoning*, 39(2-3):167–184, 2005.
- [8] F. G. Cozman. Robustness analysis of Bayesian networks with local convex sets of distributions. In D. Geiger, P. Shenoy, and H. Horvitz, editors, *13th Annual Conference on Uncertainty in Artificial Intelligence Conference*, pages 393–405, San Francisco, July 13-15 1997.
- [9] F. G. Cozman. Credal networks. *Artificial Intelligence*, 120(2):199–233, 2000.

- [10] C. de Campos and F. Cozman. The inferential complexity of bayesian and credal networks. In *19th International Joint Conference on Artificial Intelligence*, pages 1313–1318, 2005.
- [11] C. P. de Campos and F. G. Cozman. Inference in credal networks using multilinear programming. In *Second Starting AI Researcher Symposium (STAIRS)*, pages 50–61, Valencia, Spain, 2004.
- [12] P. Diaconis and S. L. Zabell. Updating subjective probability. *J. American Stat. Assoc.*, 77(380):822–830, 1982.
- [13] R. F. Drenick. Multilinear programming: Duality theories. *Journal of Optimization Theory and Applications*, 72(3):459–486, 1992.
- [14] H. Edelsbrunner. *Algorithms in Computational Geometry*. Springer-Verlag, Berlin, 1987.
- [15] I. Levi. *The Enterprise of Knowledge*. MIT Press, Cambridge, 1980.
- [16] S. Moral. Algorithms for imprecise probabilities. Technical report, Universidade de Granada, DCCIA, Universidade de Granada, 1999.
- [17] R. E. Neapolitan. *Probabilistic Reasoning in Expert Systems*. Prentice Hall, Englewood Cliffs, 1990.
- [18] K. Parsaye and M. Chignell. *Expert systems for experts*. John Wiley and Sons, New York, 1988.
- [19] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Francisco, 1988.
- [20] J. C. F. Rocha. *Algoritmos para Inferência em Redes Credais*. Doutorado, Universidade de São Paulo, São Paulo, Brasil, 2004.
- [21] S. Russell and P. Norvig. *Artificial Intelligence: A modern approach*. Prentice Hall, Upper Saddle River, 1995.
- [22] G. Shafer. Jeffrey’s rule of conditioning. *Philosophy of Science*, 48(3):337–362, 1981.
- [23] H. Sherali and C. Tuncbilek. A global optimization algorithm for polinomial programming problems using a reformulation-linearization technique. *Journal of Global Optimization*, 2(1):101–112, 1992.
- [24] N. L. Zhang and D. Poole. A simple approach to bayesian network computations. In *Workshop on Real-Time Decision Support and Diagnosis System*, pages 171–178, Banff, Canada, 2002.