

# Modeling and Reasoning with Bayesian Networks: Inference with Local Structure 13.1-13.3

Relatório semana 8 - MAC0215 (Atividade Curricular em Pesquisa)  
Aluno: Renato Lui Geh (Bacharelado em Ciência da Computação)  
Orientador: Denis Deratani Mauá

## 1 ATIVIDADES REALIZADAS NA SEMANA

Durante a semana foi feito um resumo dos seguintes tópicos do livro *Modeling and Reasoning with Bayesian Networks*[1]:

### 13 - Inference with Local Structure

#### 13.1 - Introduction

#### 13.2 - The impact of local structure on inference complexity

##### 13.2.1 - Context-specific independence

##### 13.2.2 - Determinism

##### 13.2.3 - Evidence

##### 13.2.4 - Exposing local structure

#### 13.3 - CNF encodings with local structure

##### 13.3.1 - Encoding network structure

## 2 DEFINIÇÃO DAS ATIVIDADES

Foram estudadas as vantagens de se fazer inferência por estrutura local (*local structure*) explorando algumas propriedades dos parâmetros da rede, ao invés de algoritmos baseados em estruturas (*structure-based*) como vistos nos relatórios anteriores[3] [4].

Neste relatório vamos definir algumas propriedades de se fazer inferência por estrutura local e explicar o porquê de, em alguns casos, local structure ser melhor que structure-based. Vamos separar o relatório nos seguintes tópicos:

1. Estrutura local e baseado em estrutura
2. Impacto de estrutura local na complexidade da inferência
3. Independência contexto-específica
4. Determinismo
5. Evidência

## 2.1 ESTRUTURA LOCAL E BASEADO EM ESTRUTURA

Nos relatórios anteriores vimos que podemos tornar a Rede Bayesiana mais simples fazendo transformações na rede de tal forma que tornamos inferência mais tratável. No entanto, tais algoritmos são exponenciais em tempo e espaço à *treewidth*[4] da rede. Além disso, esses algoritmos tentam melhorar a performance da rede pela estrutura da rede sem tomar em conta os valores dos parâmetros. Elas são, portanto, chamadas de algoritmos baseados em estrutura (*structure-based*).

No entanto, podemos melhorar a eficiência dos algoritmos de inferência se manipularmos a rede levando em conta os valores específicos dos parâmetros da rede. As propriedades dos parâmetros da rede que permitem tal manipulação são conhecidas como paramétrica (*parametric*) ou estrutura local (*local structure*). Esse tipo de estrutura normalmente se manifestam em redes que envolvem restrições lógicas (*logical constraints*), independência contexto-específica ou modelos locais de interação.

## 2.2 IMPACTO DE ESTRUTURA LOCAL NA COMPLEXIDADE DA INFERÊNCIA

Podemos simplificar uma Rede Bayesiana parametricamente ao compilarmos a rede e então analisarmos os parâmetros do circuito resultante. Por exemplo, se um circuito tiver parâmetros zero, podemos reduzir bastante o tamanho do circuito, assim como se um circuito possuir parâmetros distintos com valores iguais.

Além disso, considere que tenhamos uma rede tal que os nós pais de  $Y$ , denominados de  $X_1, \dots, X_n$ , formam uma relação de ou-lógico com  $Y$ . Adicionalmente, suponha que toda variável possua um valor em  $\{0, 1\}$  e que as variáveis  $X_i$  têm distribuição normal. Ou seja:

1.  $y = 1$  e  $x_i = 1$  para algum  $x_i$ , ou
2.  $y = 0$  e  $x_i = 0$  para todo  $x_i$ .

Então podemos ver que  $\theta_{y|x_1, \dots, x_n} = 1$  se e somente se obedecermos (1) ou (2).

Além disso, a *treewidth* desta rede é  $n$ , o que torna inferência em *structure-based* intratável se  $n$  for grande o suficiente. No entanto, podemos fatorar a *network polynomial* [2] de tal forma que tenhamos uma *polynomial* representável por um circuito aritmético [2] de tamanho  $O(n)$ , mesmo que a rede por trás tenha uma *treewidth*  $n$ . Portanto, podemos fazer inferência em tempo e espaço de forma linear em *treewidth*  $n$  devido aos valores específicos dos parâmetros da rede.

## 2.3 INDEPENDÊNCIA CONTEXTO-ESPECÍFICA

A rede anteriormente discutida é um exemplo de tipo de local-structure chamada de independência contexto-específica (*context-specific independence*). Particularmente, se um nó  $X_i$  tem valor 1, a probabilidade de  $Y$  se tornar independente dos outros nós pais  $X_j$  para todo  $j \neq i$  é:

$$Pr(Y|x_i = 1) = Pr(Y|x_1, \dots, x_{i-1}, x_i = 1, x_{i+1}, \dots, x_n) \quad (1)$$

Para todo valor  $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ . Esse tipo de independência é chamada de context-specific pois tal independência não se mantém para certos valores de  $X_i$ , ou seja, se o valor de  $X_i$  for 0, a independência não se manteria. Context-specific independence é portanto uma função dos parâmetros da rede, e portanto podem ser perdidas se mudarmos o valor dos parâmetros.

## 2.4 DETERMINISMO

Considere a rede da Figura 1, composta de  $n \times n$  nós binários. A treewidth dessa rede cresce linearmente em  $n$ , mas se assumirmos que cada nó assume uma função Booleana nos nós pais, então podemos descrever um circuito aritmético de tamanho  $O(n^2)$  para essa rede, já que escolhendo um valor para o nó  $I$  garante a implicação do valor de todo outro nó na rede. Portanto o número de termos não-desvanecentes (*nonvanishing terms*) na polynomial network precisa ser igual ao número de valores para o nó  $I$  e cada termo tem tamanho  $O(n^2)$ . A polynomial network terá então dois termos de tamanho  $O(n^2)$  cada.

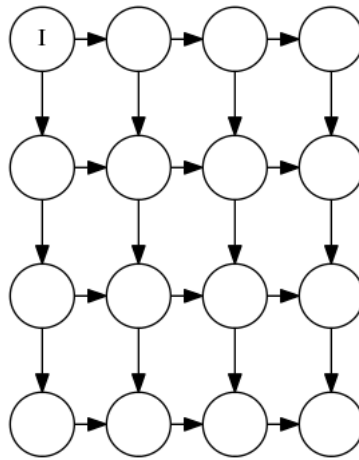


Figura 1: Uma Rede Bayesiana com estrutura em grade com  $n^2$  nós. A treewidth dessa classe de rede cresce linearmente em  $n$ .

Vamos definir de forma mais geral o exemplo acima. Considere Redes Bayesianas onde cada nó não-raiz é funcionalmente determinada pelos seus pais, ou seja, para cada nó  $X$  com pais  $U \neq \emptyset$ , temos que  $Pr(x|\mathbf{u}) = 1$  para cada instância pai  $\mathbf{u}$  e algum valor  $x$ . Isso quer dizer que  $Pr(x'|\mathbf{u}) = 0$  para todo outro valor  $x' \neq x$  e, portanto o valor  $\mathbf{u}$  dos pais  $\mathbf{U}$  implica o valor  $x$  do nó filho  $X$ . Para essas redes, chamadas de redes funcionais (*functional networks*), o número de nonvanishing terms na network polynomial é exponencial somente no número de raízes da rede, independentemente da treewidth.

Determinismo é um tipo mais geral de local structure onde parâmetros da rede tem valor zero. Functional networks exibem um tipo forte de determinismo, porém tal forma de local structure ocorre mesmo quando um nó não é funcionalmente determinado pelos seus nós pais. Seja uma variável  $X$ , por exemplo, ela pode não ser funcionalmente determinada pelos seus pais  $\mathbf{U}$ , porém um de seus valores  $x$  pode ser impossível dada uma instância de seus pais  $\mathbf{u}$ ,  $\theta_{x|\mathbf{u}} = 0$ . Tal impossibilidade é chamada de restrição lógica (*logical constraint*) e pode reduzir bastante a complexidade da inferência mesmo que não correspondam a dependência funcional.

## 2.5 EVIDÊNCIA

Local structure é especialmente eficaz dadas certas evidências. Considere a rede na Figura 2 (a) como exemplo e assuma que cada nó  $S_i$  é o ou-lógico de seus pais. Supondo que a evidência dada é  $S_1 = false$ ,  $S_2 = true$  e  $S_3 = true$ . Como  $S_1$  é falso e dados pais  $D_1$ ,  $D_2$  e  $D_3$ , temos que todos esses pais devem ser também *false*. Usando o corte anteriormente visto [4], podemos aparar todas as arestas que partem destes nós, resultando na Figura 2 (b). Note que apesar de termos visto corte de uma rede como uma técnica structure-based, devido à estrutura local dada pela evidência e a relação de parentesco entre os nós, pudemos aplicar o corte à rede.



Figura 2: Na esquerda (a) a Rede Bayesiana onde cada nó folha é um ou-lógico de seus pais. Na direita (b) vemos o resultado do corte[4] (arestas pontilhadas) da rede dados  $S_1 = false$ ,  $S_2 = true$  e  $S_3 = true$ .

## REFERÊNCIAS

- [1] Adnan Darwiche. *Modeling and Reasoning with Bayesian Networks*. 1st Edition. Cambridge University Press, 2009.
- [2] Renato Lui Geh. *Modeling and Reasoning with Bayesian Networks: Compiling Bayesian Networks*. 1. 2015. URL: <http://www.ime.usp.br/~renatolg/mac0215/doc/reports/week1/relatorio.pdf>.
- [3] Renato Lui Geh. *Modeling and Reasoning with Bayesian Networks: Inference by Variable Elimination 6.1-6.5*. 2. 2015. URL: <http://www.ime.usp.br/~renatolg/mac0215/doc/reports/week2/relatorio.pdf>.
- [4] Renato Lui Geh. *Modeling and Reasoning with Bayesian Networks: Inference by Variable Elimination 6.6-6.9*. 3. 2015. URL: <http://www.ime.usp.br/~renatolg/mac0215/doc/reports/week5/relatorio.pdf>.