

Universidade de São Paulo  
Instituto de Matemática e Estatística  
Bachalerado em Ciência da Computação

Tiago Madeira

**Geração uniforme de  $k$ -trees para  
aprendizado de redes bayesianas**

Supervisor: Prof. Dr. Denis Deratani Mauá

São Paulo  
Novembro de 2016



# Resumo

O resumo ainda não foi escrito.

**Palavras-chave:** sem, resumo, por, enquanto.



# Abstract

The abstract has not been written yet.

**Keywords:** no, abstract, yet.



# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Fundamentos</b>	<b>3</b>
2.1	Grafos . . . . .	3
2.2	$k$ -clique e $k$ -tree . . . . .	4
2.3	Probabilidade . . . . .	4
2.4	Redes bayesianas . . . . .	5
<b>3</b>	<b>Geração aleatória de <math>k</math>-trees</b>	<b>7</b>
3.1	Introdução à codificação de $k$ -trees . . . . .	7
3.2	A solução de Caminiti et al . . . . .	7
3.3	Experimentos e resultados . . . . .	8
<b>4</b>	<b>Aprendizado de redes bayesianas</b>	<b>9</b>
<b>5</b>	<b>Conclusão</b>	<b>11</b>





# Capítulo 1

## Introdução

A ser escrita.



# Capítulo 2

## Fundamentos

Neste capítulo, apresentamos algumas definições que o leitor deve conhecer para compreender o trabalho.

Partimos do pressuposto de que o leitor conhece notações básicas de conjuntos.

### 2.1 Grafos

**Definição 1 (grafo).** [1] Um grafo é um par ordenado  $G = (V, E)$ . Os elementos de  $V$  são chamados de vértices de  $G$ . Os elementos de  $E$  são chamados de arestas de  $G$  e consistem em pares (não-ordenados) de vértices. Dados  $u, v \in V$ , se  $(u, v) \in E$  dizemos que  $u$  e  $v$  são adjacentes em  $G$ .

**Definição 2 (grafo completo).** [1] Um grafo  $G = (V, E)$  é dito completo se  $(u, v) \in E$  para todo  $u, v \in V, u \neq v$ .

**Definição 3 (subgrafo induzido).** [1] Dado um grafo  $G = (V, E)$  e um subconjunto  $V'$  de  $V$ , o subgrafo de  $G$  induzido por  $V'$ ,  $G' = (V', E')$ , é o grafo formado pelos vértices  $V' \subseteq V$  e arestas que só contém elementos de  $V'$ , ou seja,  $E' = \{(u, v) \in E \mid u, v \in V'\}$ .

**Definição 4 (caminho).** [1] Dado um grafo  $G = (V, E)$ , um caminho é uma sequência de arestas que conectam uma sequência de vértices adjacentes distintos.

**Definição 5 (distância).** [1] Dado um grafo  $G = (V, E)$  e dois vértices  $(u, v) \in V$ , a distância entre  $u$  e  $v$  é o número de arestas num menor caminho que os conecte.

**Definição 6 (árvore).** [1] Dado um grafo  $G = (V, E)$ , dizemos que ele é uma árvore se cada dois vértices  $u, v \in V$  são conectados por exatamente um caminho.

## 2.2 $k$ -clique e $k$ -tree

**Definição 7 ( $k$ -clique).** [1] Seja  $G = (V, E)$  um grafo. Um  $k$ -clique é um subconjunto dos vértices,  $C \subseteq V$ , tal que  $(u, v) \in E \ \forall \ u, v \in C, u \neq v$  (ou seja, tal que o subgrafo induzido por  $C$  é completo).

**Definição 8 ( $k$ -tree).** [3] Uma  $k$ -tree é definida da seguinte forma recursiva:

1. Um grafo induzido por um  $k$ -clique é uma  $k$ -tree.
2. Se  $T'_k = (V, E)$  é uma  $k$ -tree,  $K \subseteq V$  é um  $k$ -clique e  $v \notin V$ , então  $T_k = (V \cup \{v\}, E \cup \{(v, x) \mid x \in K\})$  é uma  $k$ -tree.

**Definição 9 ( $k$ -tree enraizada).** [2] Uma  $k$ -tree enraizada é uma  $k$ -tree com um  $k$ -clique destacado  $R = \{r_1, r_2, \dots, r_k\}$  que é chamado de *raiz* da  $k$ -tree enraizada.

## 2.3 Probabilidade

[4]

## 2.4 Redes bayesianas

[4]



# Capítulo 3

## Geração aleatória de $k$ -trees

O problema de gerar  $k$ -trees está intimamente relacionado ao problema de codificá-las. De fato, se há um código bijetivo que associa  $k$ -trees à *bytes*, basta gerar *bytes* aleatórios para gerar  $k$ -trees aleatórias.

Neste capítulo, apresentamos o problema de codificar  $k$ -trees, discutimos a solução linear e bijetiva para codificar/decodificar  $k$ -trees proposta por Caminiti et al[2], explicamos como ela foi implementada neste trabalho para gerar  $k$ -trees aleatórias e mostramos os resultados obtidos.

### 3.1 Introdução à codificação de $k$ -trees

### 3.2 A solução de Caminiti et al

**Definição 10 ( $k$ -tree de Rényi).** [5] Uma  $k$ -tree de Rényi  $R_k$  é uma  $k$ -tree enraizada com  $n$  vértices rotulados em  $[1, n]$  e raiz  $R = \{n - k + 1, n - k + 2, \dots, n\}$ .

**Definição 11 (esqueleto de uma  $k$ -tree enraizada).** [2] O esqueleto de uma  $k$ -tree enraizada  $T_k$  com raiz  $R$ , denotado por  $S(T_k, R)$ , é definido da

seguinte forma recursiva:

1. Se  $T_k$  é apenas o  $k$ -clique  $R$ , seu esqueleto é uma árvore com um único vértice  $R$ .
2. Dada uma  $k$ -tree enraizada  $T_k$  com raiz  $R$ , obtida por  $T'_k$  enraizada em  $R$  através da adição de um novo vértice  $v$  conectado a um  $k$ -clique  $K$  (ver definição 8), seu esqueleto  $S(T_k, R)$  é obtido adicionando a  $S(T'_k, R)$  um novo vértice  $X = \{v\} \cup K$  e uma nova aresta  $(X, Y)$ , onde  $Y$  é o vértice de  $S(T'_k, R)$  que contém  $K$  com uma distância mínima da raiz. Chamamos  $Y$  de pai de  $X$ .

**Definição 12 (árvore característica).** [2] A árvore característica  $T(T_k, R)$  de uma  $k$ -tree enraizada  $T_k$  com raiz  $R$  é obtida rotulando os vértices e arestas de  $S(T_k, R)$  da seguinte forma:

1. O vértice  $R$  é rotulado 0 e cada vértice  $\{v\} \cup K$  é rotulado  $v$ ;
2. Cada aresta do vértice  $\{v\} \cup K$  ao seu pai  $\{v'\} \cup K'$  é rotulada com o índice do vértice em  $K'$  (visualizando-o como um conjunto ordenado) que não aparece em  $K$ . Quando o pai é  $R$  a aresta é rotulada  $\varepsilon$ .

Note que a existência de um único vértice em  $K' \setminus K$  é garantida pela definição 11. De fato,  $v'$  precisa aparecer em  $K$ , caso contrário  $K' = K$  e o pai de  $\{v'\} \cup K'$  contém  $K$ . Isso contradiz o fato de que cada vértice em  $S(T_k, R)$  é ligado à distância mínima da raiz.

### 3.3 Experimentos e resultados



# Capítulo 4

## Aprendizado de redes bayesianas

A ser escrito.



# Capítulo 5

## Conclusão

Ainda não foi escrita.



# Referências Bibliográficas

- [1] J. A. Bondy and U. S. R. Murty. *Graph Theory*. Springer, 2008.
- [2] Saverio Caminiti, Emanuele G. Fusco, and Rossella Petreschi. Bijective linear time coding and decoding for  $k$ -trees. *Theory of Computing Systems*, 46:284–300, 2010.
- [3] Frank Harary and Edgar M. Palmer. On acyclic simplicial complexes. *Mathematika*, 15:115–122, 1968.
- [4] Daphne Koller and Nir Friedman. *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press, 2009.
- [5] C. Rényi and A. Rényi. The prüfer code for  $k$ -trees. *Combinatorial Theory and its Applications*, pages 945–971, 1970.