Comunicação e Redes – 2020.QS

Lista 1 – Universidade Federal do ABC

Aritanan Gruber

aritanan.gruber@ufabc.edu.br

http://professor.ufabc.edu.br/∼aritanan.gruber

Exercícios marcados com ▶ devem ser entregues via Tidia até 16/11/2020 às 00h00m.

Notação. Exercícios marcados com: ♦ são operacionais, e envolvem a execução de cálculos simples ou o fornecimento de (contra) exemplos; ♥ são sobre teoria de redes, e geralmente requerem provas matemáticas; ♣ também são sobre redes, mas requerem a construção de um algoritmo; ♠ envolvem matemática de maneira mais geral, não limitada a grafos e variações; ★ são exercícios ligeiramente mais difíceis; finalmente, ▶ são exercícios que devem ser entregues via Tidia.

1. ♠ *Balls and bins.* Quantas possibilidades existem para se distribuir k reais entre n pessoas de forma que cada uma receba ao menos um real? Suponha agora que não é necessário às pessoas receberem ao menos um real. Qual o número de diferentes formas de ditribuição neste caso?

2. ▶ ♦ *Packing.* Alice deseja enviar sete soluções químicas a Bob: A, B, C, D, E, F e G. No entanto, qualquer par de soluções dentre A, B, C e G não podem ser acondicionadas numa mesma caixa, pois podem reagir entre sí e causar explosões. Além disso, A e F reagem com D ou E. Alice deseja enviar todas as soluções no menor número possível de caixas – pois ela pagará por caixa enviada. Modele o problema que Alice deseja resolver com um grafo e determine o menor número de caixas que ela deve utilizar. Justifique sua resposta.

3. ♥ *Regularity.* Um grafo é *k-regular* se todos os vértices possuem grau igual a k. Seja $\{R, S\}$ uma bipartição de um grafo k-regular G. Mostre que se $k > 0$ então $|R| = |S|$.

4. ♥ *Double counting.* Prove as seguintes generalizações da identidade de Euler (*handshaking lemma*):

   (a) $\displaystyle\sum_{x\in X} d(x) = \sum_{e\in E} |\, e \cap X|$, para qualquer $X \subseteq V$.

   (b) $\displaystyle\sum_{v\in V} d(v)^2 = \sum_{(x,y)\in E} (d(x) + d(y)) = \sum_{e\in E}\sum_{f\in E} |\, e \cap f|$.

5. ▶ ♥ *Unit distances.* Seja S um conjunto de n pontos no plano cuja distância entre quaisquer dois pontos em S é pelo menos 1. Mostre que existem no máximo $3n$ pares de pontos de S situados a distância exatamente 1.

6. ♥ *Degree and total number of vertices.* Sejam $G = (X \uplus Y, E)$ um grafo bipartido sem vértices isolados e $f : Y \to [0, \infty)$ uma função. Prove que se para toda aresta $(x, y) \in E$, a desigualdade $d_G(y) \le d_G(x) \cdot f(y)$ é válida, então $|X| \le \sum_{y\in Y} f(y)$.

7. ▶ ♥ *Connectivity thresholds.*

   (a) Seja G um grafo com n vértices e m arestas tal que $m > \binom{n-1}{2}$. Mostre que G é conexo. Mostre que para $n > 1$ a desigualdade é justa, exibindo um grafo não conexo com $m = \binom{n-1}{2}$.

   (b) Seja G um grafo com n vértices tal que $\delta > \frac{1}{2}(n-2)$. Mostre que G é conexo. Mostre que para n par a desigualdade é justa, exibindo um grafo não conexo com $\delta = \frac{1}{2}(n-2)$.

8. ▶ ♣ *Path-reliability.* Seja $D = (V, A)$ um digrafo em que os vértices representam servidores e os arcos representam canais unidirecionais de comunicação. Seja $c : A \to [0, 1]$ uma função que associa um valor de **confiabilidade** a cada arco de D. Isto é, $c(a)$ é a probabilidade do arco $a \in A$ continuar operando ao longo do tempo.

   A confiabilidade de um caminho $P = \langle x_0, x_1, \ldots, x_k\rangle$ em D é dada por

   $$c(P) := \prod_{i=1}^{k} c(x_{i-1}, x_i).$$

   Escreva um algoritmo que recebe o par (D, c) e determina os caminhos de maior confiabilidade entre todos os pares de vértices de D. Sugestão: $\log \prod_j x_j = \sum_j \log x_j$, e $\max f(x) = -\min -f(x)$.

9. ♠ *Linearity of expectations, moments, and variances.* Seja $\Omega \subseteq \mathbb{Z}$ e sejam $X, Y : \Omega \to \mathbb{R}$ variáveis aleatórias reais. Para um inteiro positivo k, o *k-ésimo momento* de uma variável aleatória X é dado por

$$\mathbb{E}[X^k] := \sum_{\omega \in \Omega} \omega^k \cdot \mathsf{Pr}[X = \omega].$$

A *esperaça* (média) de X é definida como $\mathbb{E}[X]$, o primeiro momento de X. A *variância* de X é definida como $\mathsf{Var}[X] := \mathbb{E}[(X - \mathbb{E}[X])^2]$. Prove que:

(a) $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$ para constantes a e b. Em outras palavras, o operador esperança é linear.

(b) Se X e Y são *independentes*, então $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$ e $\mathsf{Var}[X + Y] = \mathsf{Var}[X] + \mathsf{Var}[Y]$.

(c) $\mathsf{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}^2[X]$.

10. ▶ ♥ *Laplacians.* Sejam $G = (V, E)$ um grafo e $\vec{G} = (V, \vec{E})$ uma orientação qualquer das arestas de G. As matrizes de *adjacência* $\mathbf{A} = \mathbf{A}(G) \in \{0,1\}^{V \times V}$ e de *valência* $\mathbf{D} = \mathbf{D}(G) \in \mathbb{N}^{V \times V}$ de G e de *incidência* $\mathbf{B} = \mathbf{B}(\vec{G}) \in \{-1, 0, 1\}^{V \times E}$ de $\vec{G}$ são dadas por

$$\mathbf{A}_{u,v} = \begin{cases} 1 & \text{se } (u,v) \in E, \\ 0 & \text{caso contrário,} \end{cases} \quad \text{e} \quad \mathbf{D}_{u,v} = \begin{cases} d(u) & \text{se } u = v, \\ 0 & \text{caso contrário.} \end{cases} \quad \text{e} \quad \mathbf{B}_{u,\vec{e}} = \begin{cases} -1 & \text{se } \vec{e} = (u,v), \\ 1 & \text{se } \vec{e} = (v,u), \\ 0 & \text{caso contrário,} \end{cases}$$

Claramente, $\mathbf{A}$ é simétrica e $\mathbf{D}$ é diagonal. A matriz *Laplaciana* $\mathbf{L} = \mathbf{L}(G)$ de G é definida como $\mathbf{L} = \mathbf{D} - \mathbf{A}$. Prove que $\mathbf{B}\mathbf{B}^\top = \mathbf{L}$.

11. ♠ ♥ *Eigenvalues of graphs.* Os **autovalores de uma matriz** quadrada $\mathbf{A}$ são as raízes de seu **polinômio característico** $\det(\mathbf{A} - \lambda\mathbf{I})$, em que $\mathbf{I}$ é a matriz identidade com mesmas dimensões de $\mathbf{A}$ e λ é a variável do polinômio. Isto equivale a dizer que um escalar λ_1 é um autovalor de $\mathbf{A}$ se $\mathbf{A}\mathbf{x} = \lambda_1\mathbf{x}$ – em sendo o caso, $\mathbf{x}$ é um **autovetor** associado a λ_1. Os **autovalores de um grafo** são os autovalores de sua matriz de adjacências. Seja G um grafo k-regular com n vértices. Mostre que:

(a) $\mathbf{B}\mathbf{B}^\top = \mathbf{A} + k\mathbf{I}$, em que $\mathbf{B} = \mathbf{B}(G)$ é a matriz de incidências de G.

(b) k é um autovalor de G, com correspondente autovetor $\mathbf{1}$, o vetor cujas n entradas são iguais a 1.

(c) Todo autovalor de G é um número real. Em outras palavras, grafos não admitem autovalores complexos.

(d) Todo autovalor racional de G é na realidade um inteiro.

12. ▶ ♦ *Centralities.* Considere o grafo definido pelas lista de arestas abaixo:

u	Alice	Carl	Alice	Alice	Alice	Bob	Gail	Harry	Jen	Harry	Irene	Irene	Ernst	David	Carl
v	Bob	Alice	David	Ernst	Frank	Gail	Harry	Jen	Gail	Irene	Gail	Jen	Frank	Carl	Frank

(a) Esboce um desenho do grafo.

(b) Quantos vértices e quantas arestas existem? Qual a densidade do grafo?

(c) Determine o grau de cada vértice e identifique os vértices centrais.

(d) Calcule a centralidade por autovetor de cada vértice e identifique os vértices centrais.

(e) Calcule a centralidade por proximidade de cada vértice e identifique os vértices centrais.

(f) Calcule a centralidade por proximidade harmônica de cada vértice e identifique os vértices centrais.

(g) Calcule a centralidade por intermediação (*betweenness*) de cada vértice e identifique os vértices centrais.

(h) Calcule o coeficiente de clusterização da cada vértice e o coeficiente médio de clusterização do grafo.

(i) Existe uma conspiração contra o vértice central por intermediação determinado no item (g). Se você fosse adicionar uma única aresta ao grafo de forma a reduzir a centralidade por intermediação de tal vértice, qual seria?

(j) Esboce o gráfico dos graus versus coeficientes de intermediação dos vértices. Estas variáveis são correlacionadas? Calcule o coeficiente de correlação e seu. Quais são os maiores *outliers* nesta relação? Utilize a figura feita no item (a) na interpretação dos *outliers*.

13. ► ♠ *Probabilistic double counting.* Alice está tentando enviar a Bob um único *bit* — dígito 0 ou 1. Após enviado, o bit percorre um caminho passando por n retransmissores até chegar a Bob. Devido a limitações técnicas, cada retransmissor pode introduzir um ruído na transmissão, invertendo o valor do bit com probabilidade p.

   (a) Mostre que a probabilidade de Bob receber o bit enviado por Alice corretamente é

   $$\sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} p^{2k}(1-p)^{n-2k}.$$

   (b) Considere agora uma forma alternativa de calcular tal probabilidade. Dizemos que o retransmissor tem um *viés* q se a probabilidade com que inverte o valor do bit é $(1-q)/2$. O viés q é, portanto, um valor real no intervalo $[-1,1]$. Prove que o envio de um bit via dois retransmissores com viés q_1 e q_2 é equivalente a enviar o bit através de um único retransmissor com viés q_1q_2.

   (c) Prove que a probabilidade de Bob receber o bit enviado por Alice corretamente após passar por n retransmissores como descrito no enunciado é igual a

   $$\frac{1+(1-2p)^n}{2}.$$

14. ► ♦ Para esta questão, vamos utilizar o software *Gephi* (disponível em https://gephi.org/). Gere grafos aleatórios em $\mathscr{G}([100], p)$ para $p \in \{0.001, 0.01, 0.03, 0.06, 0.1\}$ – utilize o layout *Fruchterman Reingold* para melhor vizualização.

   (a) Em questões estruturais, quais as diferenças entre os grafos gerados?

   (b) Determine as estatísticas (grau médio, diâmetro, densidade, pagerank, centralidade por autovetor, coeficiente de clusterização, etc.) disponíveis no menu à direita para cada grafo. Os valores encontrados são consistentes com o observado em (a)?

   (c) Repita seu experimento gerando novamente os grafos. Há flutuações nos resultados em (b)? Em caso afirmativo, o que você deve fazer para obter resultados em que a possibilidade de flutuações tenha relevância pequena?

15. ♠ *Coupon-collector problem.* Alice possui uma empresa de alimentos e decidiu pegar carona na popularidade de um evento esportivo para incrementar as vendas. A "promoção" consiste em cartões como fotos e estatísticas de atletas escondidos dentro das caixas de um certo cereal. Cada caixa possui somente um cartão e existem n cartões diferentes. Por algum motivo, você decidiu colecionar os cartões. Prove que sem colaborar com outras pessoas, você precisará comprar pelo menos $n \ln n + \Theta(n)$ caixas de cereal até obter pelo menos uma unidade de cada cartão.

   A notação $\Theta(f(n))$ representa o conjunto de todas as funções $g(n)$ tais que $\ell f(n) \leq g(n) \leq u f(n)$ para todo $n \geq n_0$, com ℓ, u, n_0 constantes reais positivas.

   Sugestão: Utilize $\ln n + \Theta(1)$ como aproximação para o *número Harmônico* $H(n) := \sum_{k=1}^{n} \frac{1}{k}$.

16. ♠ *Harmonic number's approximation.* Prove que o *número Harmônico* $H(n) := \sum_{k=1}^{n} \frac{1}{k}$ satisfaz $H(n) = \ln + \Theta(1)$.

   Sugestão: $\displaystyle\int_1^n \frac{1}{x}dx \leq \sum_{k=1}^{n} \frac{1}{k}$ e $\displaystyle\sum_{k=2}^{n} \frac{1}{k} \leq \int_1^n \frac{1}{x}dx.$

17. ♥ *Connected components.* Sejam $n > k \geq 2$ inteiros e seja X_k a variável aleatória que denota o número de componentes conexos com exatos k vértices em $G \sim \mathscr{G}(n,p)$, para uma medida de probabilidade $p = p(n)$. Mostre que

   $$\mathbb{E}[X_k] \leq \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)}.$$

18. ► ♥ *Monotone properties.* Suponha que $\mathscr{P}$ é uma propriedade monótona crescente e que $0 \leq p_1 < p_2 \leq 1$. Sejam $G_i \sim \mathscr{G}(n, p_i)$, com $i \in \{1,2\}$. Prove que

   $$\Pr[G_1 \in \mathscr{P}] \leq \Pr[G_2 \in \mathscr{P}].$$