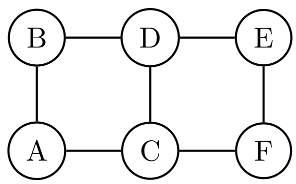
Trabalho Autónomo 3

Análise de Redes

Student's name: Diogo Alexandre Alonso de Freitas Freitas;
Student's number: 104841;
Student's mail: daafs@iscte-iul.pt

Q. 1 *A menos que seja especificado de outra forma, o comprimento de um caminho é o número de ligações nele contidas. Dados dois nós num grafo conexo arbitrário e não dirigido, deve existir algum caminho mais curto entre eles. Verdadeiro ou Falso: Podem existir vários caminhos mais curtos?*

- **Verdadeiro**, escolhendo dois nós aleatoriamente de um determinado grafo, tendo em conta que este é conexo arbitrário (é possível ir de qualquer node a qualquer outro, seguindo as linhas), podem existir vários caminhos curtos. Abaixo podemos visualizar um grafo que será utilizado como exemplo.



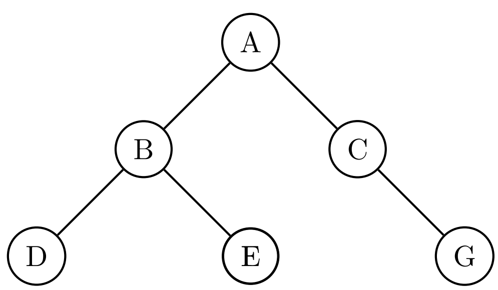
Imaginemos que queremos ir do ponto D ao ponto F, existem 2 caminhos curtos que podem ser percorridos:

I. $D \to E \to F$

II. $D \to C \to F$

Q. 2 *Verdadeiro ou Falso: Dados dois nós quaisquer de uma árvore (não dirigida), existe exatamente um caminho entre esses dois nós?*

- **Verdadeiro**, numa árvore (não dirigida), existe exatamente um caminho entre quaisquer dois nós, pois uma árvore é conexa e aciclica. Abaixo podemos visualizar uma árvore que corrobora o que foi dito anteriormente:

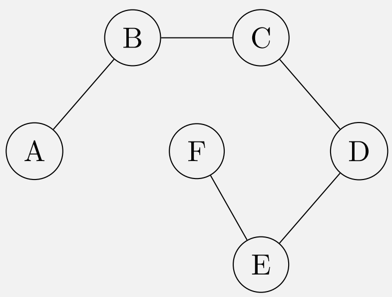


Q. 3 *Considere uma rede conexa e não dirigida com N nós. Qual é o número mínimo de ligações que a rede pode ter? Se não exigirmos que a rede esteja ligada, este número mínimo de ligações altera-se?*

O número mínimo de arestas que um grafo depende se esta necessita de possuir todos os nodos conectados ou não. Abaixo serão citados 2 casos diferentes, dependendo da interpretação do mesmo:

- **Grafo Vazio:** Um grafo que não possui arestas (ou seja, não conecta nenhum par de nós) é o caso mínimo. Portanto, para um grafo simples com $\boldsymbol{n}$ nós, o número mínimo de arestas é 0.
- **Grafo Conectado:** Se for necessário que todos os nós tenham alguma ligação, ou seja, haja um caminho entre qualquer par de nós, o número mínimo de arestas é $\boldsymbol{n}$ - 1. Isso corresponde a um grafo em forma de árvore, onde cada nó é conectado de forma que não haja ciclos. Abaixo, podemos visualizar um exemplo deste grafo com 6 nodos:

Exemplo 1: Número minimo de arestas de um grafo conectado com 6 nós



Cálculos auxiliares:

$$8 - 1 = 7$$

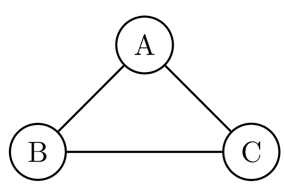
Pelo cálculo acima, podemos concluir que o número mínimo de arestas que um grafo conectado com 7 nodos pode ter é 6.

Q. 4 *Recorde-se que uma árvore de n nós contém n - 1 ligações. Verdadeiro ou Falso: Qualquer rede ligada e não dirigida de n nós e n - 1 ligações deve ser uma árvore?*

Na **Q. 2** podemos visualizar uma árvore com n nodos e $n - 1$ ligações e, por definicção, qualquer rede é uma árvore se for conexa e acíclica (que é o caso destas)

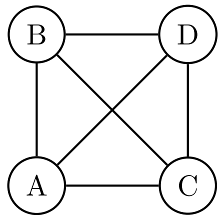
Q. 5 *Verdadeiro ou Falso: Qualquer rede não dirigida de N nós com pelo menos N ligações deve conter um ciclo?*

Verdadeiro, uma rede não dirigida com N nós e pelo menos N ligações sempre conterá um ciclo, abaixo podemos visualizar o exemplo de uma rede com 3 nodos e 3 ligações (visivel que possui ciclos):



Q. 6 *Verdadeiro ou Falso: Qualquer rede não dirigida de N nós com pelo menos N ligações deve conter um ciclo?*

Verdadeiro, imaginemos uma rede como um conjunto de cidades (nós) conectadas por estradas (ligações). Se tivermos mais estradas do que cidades, inevitavelmente haverá cruzamentos, ou seja, caminhos que se encontram, formando um ciclo. Abaixo podemos visualizar uma rede com o máximo de ligações possível:

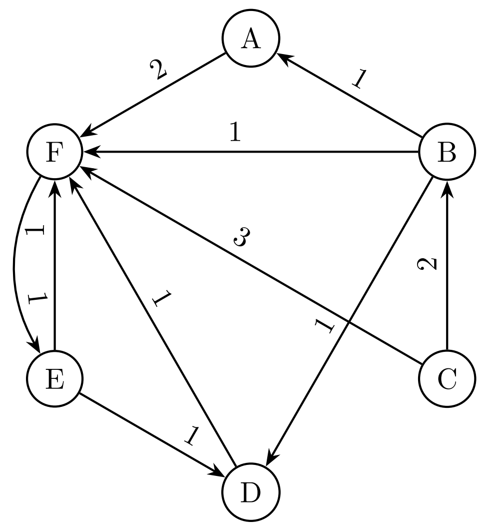


É notório que ela é cíclica e, se retirar-mos as ligações centrais que se cruzam, ela continua a ser cíclica.

Q. 7 *Considere a rede definida pela matriz de adjacência acima. Existem ciclos nesta rede? Está fortemente conectado? Fracamente ligado?*

Para uma visualização mais simples do problema, abaixo podemos visualizar, não só a matriz de adjacência, como também a respetiva rede:

$$
\begin{array}{c c}
 & \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\
\begin{array}{c} A \\ B \\ C \\ D \\ E \\ F \end{array} &
\begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
2 & 1 & 3 & 1 & 1 & 0
\end{pmatrix}
\end{array}
$$



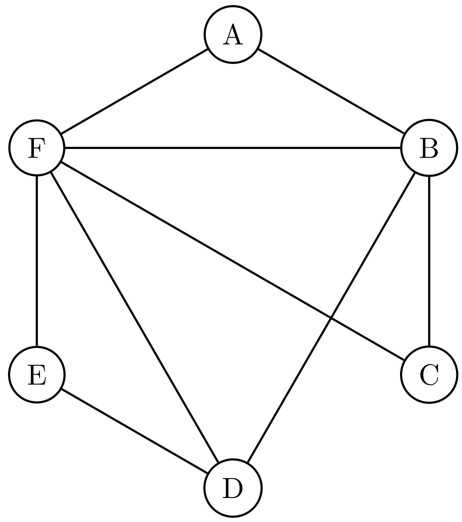
I. A rede definida acima é cíclica: $D \to F \to E \to D$

II. Uma rede orientada é fortemente conexa se existe, pelo menos, um caminho entre qualquer par de nodos (i, j), de i para j e vice-versa. Como não existe uma ligação do nodo C para B, esta rede não é fortemente conexa.

III. Uma rede orientada é fracamente conexa se, quando se ignora a orientação das ligações, existe, pelo menos, um caminho entre qualquer par de nodos (i, j). No grafo é notório que não existe ligação entre C e D.

Q. 8 *Considere a versão não ponderada e não dirigida da rede definida pela matriz de adjacência acima. Esta rede é uma árvore?*

Abaixo podemos visualizar a rede, apresentada na **Q. 7**, mas não ponderada e não dirigida:



Para uma rede ser uma árvore, ela necessita de seguir os seguintes passos:

I. Ser uma rede **não orientada**;

II. Ser uma rede **conexa**;

III. Ser uma rede sem **cíclos**;

IV. N nodos e $N-1$ ligações;

O grafo acima é **não orientada e conexa**, possui **cíclos** e possui **6 nodos** e **9 ligações**. Logo, esta rede **não** é uma árvore.

Q. 9 *Considere a versão não ponderada e não dirigida da rede definida pela matriz de adjacência acima. Qual é o diâmetro desta rede?*

A respetiva rede pode ser visualizada na questão: **Q. 8**.
O diâmetro de uma rede é definido como:

$$l_{max} = max_{i,j} l_{i,j}$$

Logo, o diâmetro desta rede é 5 (D $\rightarrow$ E $\rightarrow$ F $\rightarrow$ A $\rightarrow$ B $\rightarrow$ C).

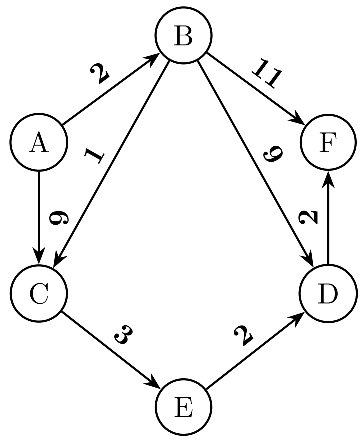
Q. 10 *Se converter uma rede dirigida fracamente ligada numa rede não dirigida, a rede resultante será ligada? Explique porquê ou por que não.*

Sim, a antiga rede dirigida **fracamente ligada** tornar-se-á uma rede ligada caso seja convertida em uma rede não dirigida. Isto ocorre, pois, por ela ser fracamente liganda, temos à priori a informação de que "existe pelo menos uma ligação entre qualquer par de nodo (i, j), caso se ignore a respetiva direção da ligação".

Q. 11 *Considere uma rede não dirigida arbitrária, e não completa. Agora adicione um único link. Como é que o número de nós na componente gigante desta rede mudou como resultado desta adição?*

- ○ **A. Diminuiu estritamente.**
- ○ **B. Diminuiu ou manteve-se o mesmo.**
- ● **C. Aumentou ou manteve-se o mesmo.**
- ○ **D. Aumentou estritamente**

Q. 12 *Considere a seguinte rede dirigida. Qual das seguintes alternativas descreve com maior precisão a conectividade desta rede?*

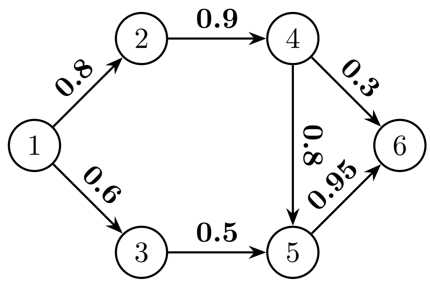


- ○ **A. Fortemente conectado**
- ● **B. Fracamente conectado**
- ○ **C. Desconectado**
- ○ **D. Nenhuma das anteriores**

Q. 13 *Quantos nós existem no maior componente fortemente ligado da rede na figura acima?*

Na rede do exercicio **Q. 12**, não existe nenhuma componente fortemente ligada maior que 1.

Q. 14 *Considere a rede seguinte. Qual das seguintes alternativas descreve com maior precisão a conectividade desta rede?*



- ○ **A. Fortemente conectado**
- ● **B. Fracamente conectado**
- ○ **C. Desconectado**
- ○ **D. Nenhuma das anteriores**

Q. 15 *Os pesos das ligações podem representar qualquer coisa sobre a relação entre os nós: força da relação, distância geográfica, tensão que flui através de um cabo de ligação, etc. os pesos estão relacionados com as distâncias. O comprimento de um caminho entre dois nós é então a soma das distâncias dos links nesse caminho. O caso mais simples ocorre quando os pesos dos links representam a distância espacial. Considere a rede do exercicio **Q. 14** e suponha que os pesos das ligações representam distâncias. Utilizando esta métrica de distância espacial, qual é o caminho mais curto entre os nós 1 e 6?*

Abaixo, podemos visualizar todos os caminhos possíveis de 1 até 6, estando selecionada o caminho mais curto entre estes:

I) $1 \xrightarrow{0.6} 3 \xrightarrow{0.5} 5 \xrightarrow{0.95} 6 \Rightarrow 2.05 \times$

II) $1 \xrightarrow{0.8} 2 \xrightarrow{0.9} 4 \xrightarrow{0.3} 6 \Rightarrow 2.00 \checkmark$

III) $1 \xrightarrow{0.8} 2 \xrightarrow{0.9} 4 \xrightarrow{0.8} 5 \xrightarrow{0.95} 6 \Rightarrow 3.45 \times$

Pelos cálculo acima, é notório que o caminho mais curto corresponde ao ponto **II)**

Q. 16 *Uma forma comum de definir a distância entre dois nós é o inverso (ou recíproco) do peso da ligação. Considere a rede do exercício **Q. 14** e suponha que a distância entre dois nós adjacentes é definida como o inverso do peso da ligação. Utilizando esta métrica de distância, qual é o caminho mais curto entre os nós 1 e 6?*

Por invertermos os pesos das ligações, o novo caminho mais curto passa a ser o antigo caminho mais longo do exercicio **Q. 15**, sendo esse o caminho **III)**:

$$1 \xrightarrow{\frac{1}{0.8}} 2 \xrightarrow{\frac{1}{0.9}} 4 \xrightarrow{\frac{1}{0.8}} 5 \xrightarrow{\frac{1}{0.95}} 6 \Rightarrow \frac{3.19}{0.684} \approx 4.66374$$

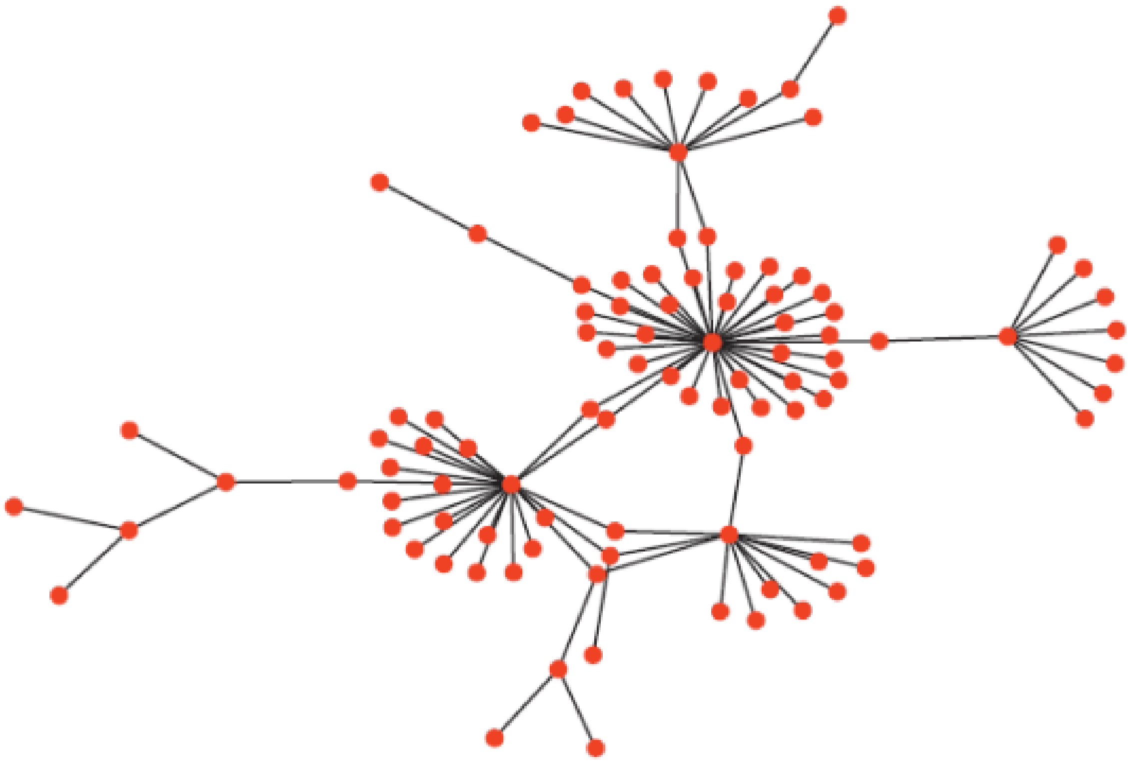


Figura 1: Rede Questão 17

Q. 17 *Considere a rede da figura acima. Qual das seguintes alternativas é a melhor estimativa do diâmetro desta rede?*

- ○ **A. 2**
- ○ **B. 4**
- ● **C. 10**
- ○ **D. 20**

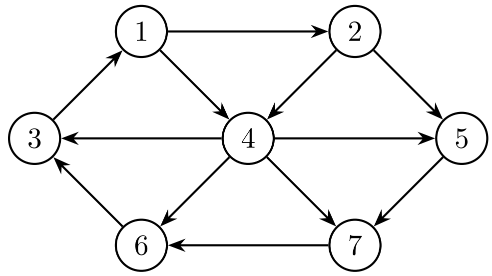
Q. 18 *Considere a rede da figura acima. Qual das seguintes alternativas é a melhor estimativa para o coeficiente de agrupamento médio deste gráfico?*

- ● **A. 0.05**
- ○ **B. 0.5**
- ○ **C. 0.75**
- ○ **D. 0.95**

Q. 19 *Seria provável que uma rede social tivesse o diâmetro e o coeficiente de agrupamento do gráfico da figura acima?*

A combinação de um diâmetro de 10 e um coeficiente de agrupamento de 0,05 em uma rede social é altamente improvável. Redes sociais reais tendem a exibir um equilíbrio entre alta conectividade (diâmetro baixo) e alta agrupamento (coeficiente de agrupamento alto). Essa combinação contraintuitiva poderia surgir em cenários muito específicos e artificiais, mas não é característica de redes sociais típicas.

Q. 20 *Considere a seguinte rede. Qual das seguintes alternativas descreve com maior precisão a conectividade desta rede?*



- **A. Fortemente conectado**
- **B. Fracamente conectado**
- **C. Desconectado**
- **D. Nenhuma das anteriores**

Q. 21 *Qual é o diâmetro da rede da questão **Q. 20**?*

O diâmetro da rede da **Q. 20** é 6: $1 \to 2 \to 4 \to 5 \to 7 \to 6 \to 3$

Q. 22 *Considere uma versão não dirigida da rede da figura **Q. 20**. Qual é o diâmetro desta rede?* O diâmetro da rede, não dirigida, será 3: $1 \to 4 \to 7$

Q. 23 *Considere um qualquer grafo dirigido arbitrário D juntamente com a sua versão não dirigida G. Verdadeiro ou Falso: Se o comprimento e o diâmetro médio do caminho mais curto do grafo dirigido existirem, podem ser menores do que os da versão não dirigida?*

O diâmetro do grafo **não dirigido** é **igual ou menor** do que o do **grafo dirigido**, pois a versão **não dirigida** permite mais caminhos entre os nós (podendo existir algum caminho mais "eficiente"entre dois dados nós). Portanto, o diâmetro da versão **não dirigida** nunca será superior ao da versão **dirigida**.

Q. 24 *Suponha que tem um grafo com 100 nós e 200 ligações. Qual é o grau médio de nós nesta rede?*

Para calcular o grau médio de nós nesta rede iremos usar a seguinte fórmula:

$$\text{Grau médio} = \frac{2 \times \text{Número de arestas}}{\text{Número de nodos}}$$

Logo:

$$\text{Grau médio} = \frac{2 \times 200}{100} = 4$$

Q. 25 *Considere uma rede constituída por 250 estudantes num dormitório. As ligações nesta rede representam relações entre companheiros de quarto: dois nós estão ligados se forem atualmente colegas de quarto. Neste dormitório, os quartos são maioritariamente para ocupação dupla, com alguns triplos e quádruplos.*

(a) **Este gráfico está ligado?**
Com base na descrição inicial, o grafo parece ser desconectado, já que os quartos (ou "cliques") não estão diretamente ligados entre si. Para o grafo ser ligado, seria necessário que pelo menos alguns estudantes servissem como "pontes"entre diferentes quartos, conectando as várias componentes. Se os estudantes permanecem sempre com os mesmos companheiros de quarto, a rede não será ligada.

(b) **Qual é a moda (valor mais frequente) da distribuição de graus dos nós?**
A moda da distribuição de graus dos nós nesta rede é 1, visto que a maioria dos estudantes tem 1 companheiro de quarto.

(c) **Quantos nós estão no maior clique?**
O maior clique nesta rede será composto pelos estudantes que partilham quartos de quádrupla ocupação. Logo, o maior clique contém 4 nós, que representam os estudantes num quarto quádruplo.

(d) **Esperaria que este gráfico tivesse algum hub?**
É É possível que haja estudantes com algumas ligações adicionais (como aqueles em quartos quádruplos), mas esses não seriam considerados hubs em um sentido forte, pois não atingiriam um número de ligações suficientemente alto para influenciar a conectividade da rede de forma significativa. Portanto, é mais realista concluir que **não há hubs** na rede de estudantes, dada a sua configuração típica.

Q. 26 *No NetworkX, como pode encontrar um nó com um maior grau de centralidade numa rede? E como obteria também o grau desse nó?*

Nó com maior grau de centralidade de uma rede

```
import networkx as nx

# Criar um grafo simples (exemplo)
G = nx.Graph()
G.add_edges_from([(1, 2), (1, 3), (2, 3), (3, 4), (4, 5)])

# Calcular a centralidade de grau
degree_centrality = nx.degree_centrality(G)

# Encontrar o nó com maior centralidade de grau
max_degree_node = max(degree_centrality, key=degree_centrality.get)

# Obter o grau do nó
degree_of_max_node = degree_centrality[max_degree_node]
```

Q. 27 *Suponha que tem um gráfico NetworkX G de funcionários. Os nomes dos nós são IDs de colaboradores e os nós possuem atributos para o nome completo, departamento, cargo e salário. Qual das seguintes opções lhe dará o salário do funcionário com o ID 5567?*

- ○ **A. G.node (5567) ('salário')**
- ○ **B. G[5567]**
- ○ **C. G.nó [5567] salário'**
- ○ **D. G(5567)**
- ● **E. Nenhuma das anteriores (Deveria ser: G.nodes[5567]['salario'])**

Q. 28 *Tem um gráfico NetworkX G e está prestes a desenhá-lo com o seguinte comando: nx.draw(G, node size=node size list). Qual das seguintes alternativas é a forma correta de forma a que os nós sejam dimensionados de acordo com o seu grau?*

- ○ **A. nodesizelist = [G(n) for n in G.nodes]**
- ○ **B. nodesizelist = G.degree()**
- ○ **C. nodesizelist = [G.degree() para n em G.nodes]**
- ● **D. nodesizelist = [G.degree(n) para n em G.nodes]**
- ○ **E. nodesizelist = [d para d em G.degree()]**

Q. 29 *Uma rede de colaboração académica é um tipo de rede social. Nesta rede, um nó com grau dois significa que:*

- ○ **A. Um académico foi coautor de um artigo com outro académico**
- ● **B. Um académico é coautor de publicações com outros dois académicos**
- ○ **C. Um estudioso é autor de duas publicações**
- ○ **D. Uma publicação foi da coautoria de dois estudiosos**

Q. 30 *Numa rede social, qual das seguintes afirmações se esperaria ser verdadeira sobre o grau dos seus nós?*

- ○ **A. A maioria dos nós liga-se num único hub grande**
- ● **B. Uma variedade de graus pode ser encontrada**
- ○ **C. Todos os nós têm mais ou menos o mesmo grau**
- ○ **D. Todos os nós têm um grau muito**

Q. 31 *Que propriedade necessita uma rede de ter para que a centralidade de proximidade fique bem definida?*

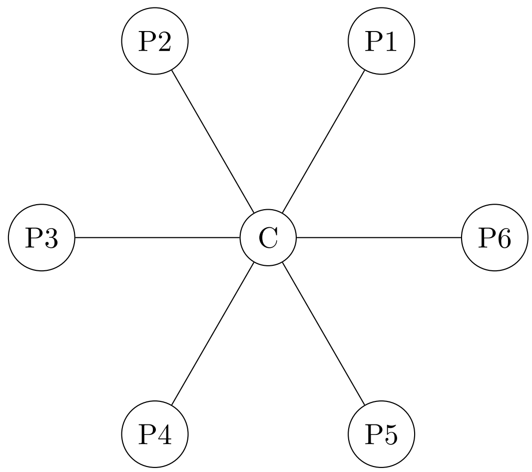
Para que a centralidade de proximidade esteja bem definida, a rede precisa de ser conexa. A centralidade de proximidade mede a inversa da soma das distâncias mínimas (caminhos mais curtos) de um nó para todos os outros nós na rede. Se a rede não for conexa, existem pares de nós entre os quais não há caminho, tornando as distâncias infinitas. Numa rede desconexa, a centralidade de proximidade ficaria indefinida ou teria de ser ajustada para lidar com as distâncias infinitas, o que não reflete corretamente a proximidade entre os nós.

Q. 32 *Apresente exemplos de redes tais que:*

(a) **O nó com maior grau não é aquele que tem maior proximidade**

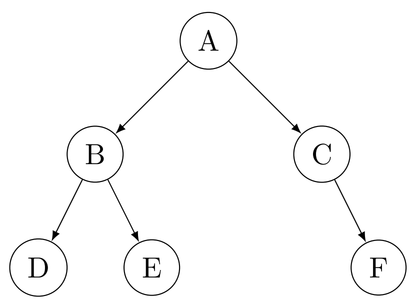
Um bom exemplo seria a rede **Estrela**:

i. O nó central (hub) tem o maior grau, pois está conectado a todos os outros.
ii. No entanto, os nós satélites estão mais próximos entre si do que do hub. Assim, os satélites podem ter uma centralidade de proximidade maior que o hub.



(b) **O nó com maior intermediação não é aquele que tem maior proximidade**

Um bom exemplo seria a rede em **Árvore**:



i. Nó A tem a maior intermediação, pois está localizado no centro da árvore e é atravessado pelos caminhos mais curtos entre os nós em diferentes ramos. Por exemplo, qualquer caminho de D para F, ou de E para F, passa obrigatoriamente por A.
ii. A centralidade de proximidade de um nó é definida pela soma das distâncias mínimas até todos os outros nós. Neste caso, nós B e C estão, em média, mais próximos de todos os outros nós do que A

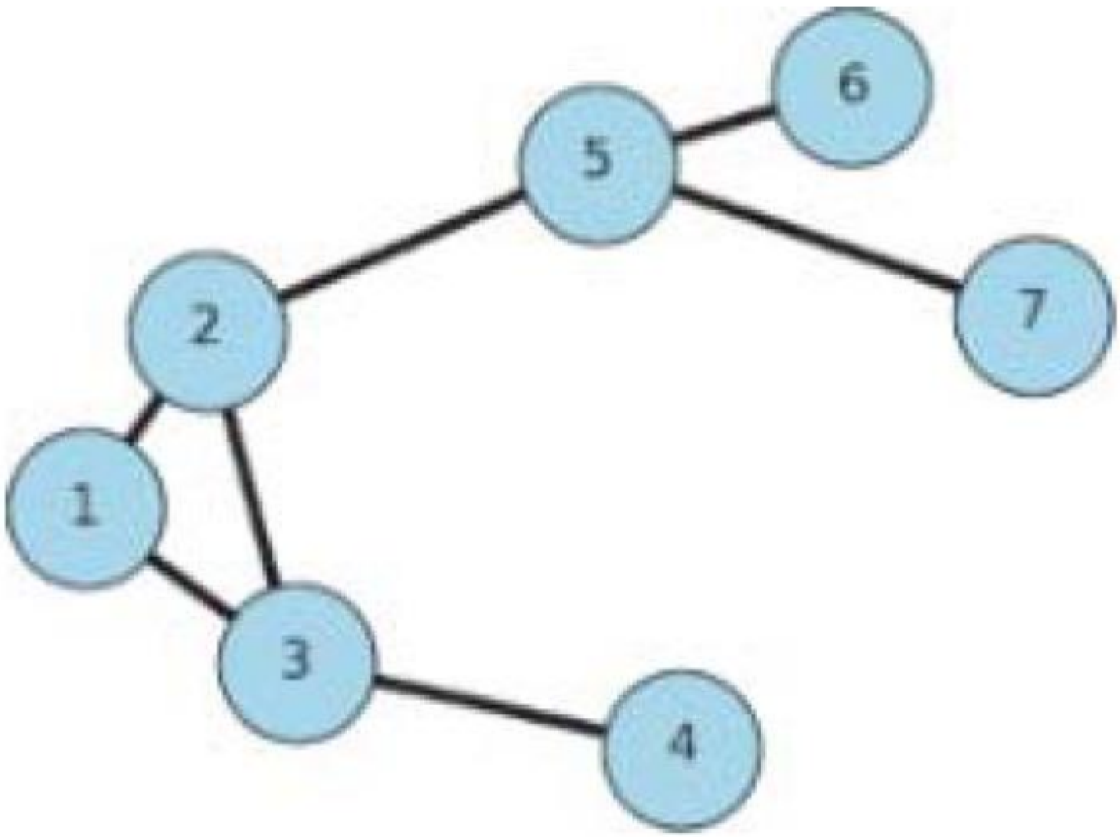


Figura 2: Rede Questão 33

Q. 33 **Considere a rede da figura acima para responder às próximas questões. Para cada questão, em caso de empate, responda com todos os nós superiores empatados.**

Criação da respetiva rede no networkingX

```
import networkx as nx
import matplotlib.pyplot as plt

# Criar um grafo vazio
G = nx.Graph()

# Adicionar arestas (conexões) entre os nós, conforme a imagem
G.add_edges_from([(1, 2), (1, 3), (2, 3), (2, 5), (3, 4), (5, 6), (5, 7)])
```

Acima é o código base, que consiste na criação da rede do exercicio, e será utilizado para a realização dos exercícios a seguir:

(a) **Qual o nó que tem o maior grau de centralidade?**

```
# Calcular o grau de centralidade de todos os nós
degree_centrality = nx.degree_centrality(G)

# Mostrar o grau de centralidade de cada nó
for node, centrality in degree_centrality.items():
    print(f"Nó {node}: Grau de centralidade = {centrality:.2f}")

# Identificar o nó com o maior grau de centralidade
max_node = max(degree_centrality, key=degree_centrality.get)
print(f"\nO nó com o maior grau de centralidade é o nó {max_node} com
↪  grau {degree_centrality[max_node]:.2f}")


----------------------------------------------------------------------

                            *Python Console*

>> Nó 1: Grau de centralidade = 0.33
>> Nó 2: Grau de centralidade = 0.50
>> Nó 3: Grau de centralidade = 0.50
>> Nó 5: Grau de centralidade = 0.50
>> Nó 4: Grau de centralidade = 0.17
>> Nó 6: Grau de centralidade = 0.17
>> Nó 7: Grau de centralidade = 0.17

>> O nó com o maior grau de centralidade é o nó 2 com grau 0.50
```

Acima, podemos visualizar que, pelo código, o nodo com maior grau de centralidade é o nó 2. Abaixo podemos visualizar o cálculo do grau de centralidade do nó 2.

$$C_D(2) = \frac{3}{7-1} = \frac{1}{2}$$

(b) **Qual o nó com maior centralidade de intermediação?**

```
# Calcular a centralidade de intermediação de todos os nós
betweenness_centrality = nx.betweenness_centrality(G)

# Mostrar a centralidade de intermediação de cada nó
for node, centrality in betweenness_centrality.items():
    print(f"Nó {node}: Centralidade de intermediação =
     ↪  {centrality:.2f}")

# Identificar o nó com a maior centralidade de intermediação
max_betweenness_node = max(betweenness_centrality,
 ↪  key=betweenness_centrality.get)
print(f"\nO nó com a maior centralidade de intermediação é o nó
 ↪  {max_betweenness_node} com valor
 ↪  {betweenness_centrality[max_betweenness_node]:.2f}")


----------------------------------------------------------------------

                            *Python Console*

>> Nó 1: Centralidade de intermediação = 0.00
>> Nó 2: Centralidade de intermediação = 0.60
>> Nó 3: Centralidade de intermediação = 0.33
>> Nó 5: Centralidade de intermediação = 0.60
>> Nó 4: Centralidade de intermediação = 0.00
>> Nó 6: Centralidade de intermediação = 0.00
>> Nó 7: Centralidade de intermediação = 0.00

>> O nó com a maior centralidade de intermediação é o nó 2 com valor
 ↪  0.60
```

Acima, podemos visualizar que, pelo código, o nodo com maior centralidade de intermediação é o nó 2. Abaixo podemos visualizar a fórmula para obter o grau de centralidade.

$$C_B(v) = \sum_{s,t \in V, s \neq t \neq v} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

(c) **Qual o nó com maior centralidade de proximidade?**

```
# Calcular a centralidade de proximidade de todos os nós
closeness_centrality = nx.closeness_centrality(G)

# Mostrar a centralidade de proximidade de cada nó
for node, centrality in closeness_centrality.items():
    print(f"Nó {node}: Centralidade de proximidade =
     ↪  {centrality:.2f}")

# Identificar o nó com a maior centralidade de proximidade
max_closeness_node = max(closeness_centrality,
 ↪  key=closeness_centrality.get)
print(f"\nO nó com a maior centralidade de proximidade é o nó
 ↪  {max_closeness_node} com valor
 ↪  {closeness_centrality[max_closeness_node]:.2f}")


----------------------------------------------------------------------

                           *Python Console*

>> Nó 1: Centralidade de proximidade = 0.50
>> Nó 2: Centralidade de proximidade = 0.67
>> Nó 3: Centralidade de proximidade = 0.55
>> Nó 5: Centralidade de proximidade = 0.60
>> Nó 4: Centralidade de proximidade = 0.38
>> Nó 6: Centralidade de proximidade = 0.40
>> Nó 7: Centralidade de proximidade = 0.40

>> O nó com a maior centralidade de proximidade é o nó 2 com valor
 ↪  0.67
```

$$C_p(v) = \frac{1}{\sum_{u \in V} d(v,u)}$$