Homework 2

October 15, 2023

Jose Carlos Munoz

1

$$H(a,b) = 4$$
$$H(a,c) = 3$$

String a and c are the most similar becuase they have the lowest hamming distance.

2

α has a total of 10 elements. β has a total of 11 element. λ has a total of 12 element.

$$J(\alpha,\beta) = \frac{\alpha \cap \beta}{\alpha \cup \beta} = \frac{7}{14} = .5 \qquad J(\alpha,\lambda) = \frac{\alpha \cap \lambda}{\alpha \cup \lambda} = \frac{8}{14} = .571$$

α and λ are the most similar because they have the highest Jaccard coefficient.

3

$$E(human, animal) = 4$$
$$E(human, plant) = 4$$

Both animal and plant have the same similarity to human because their edit distance is 4 for each.

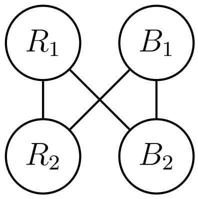
4

$\vec{X} * \vec{Y} = 1*0 + 2*2 + 2*0 = 4$. $\vec{X} * \vec{Z} = 1 * -2 + 2*2 + 2 * -1 = 0$

$$\begin{aligned}
\mid \vec{X} \mid &= \sqrt{(}1^2 + 2^2 + 2^2) & &= \sqrt{(}1+4+4) = \sqrt{(}9) & &= 3 \\
\mid \vec{Y} \mid &= \sqrt{(}0^2 + 2^2 + 0^2) & &= \sqrt{(}0+4+0) = \sqrt{(}4) & &= 2 \\
\mid Z \mid &= \sqrt{(}-2^2 + 2^2 + -1^2) & &= \sqrt{(}4+4+1) = \sqrt{(}9) & &= 3
\end{aligned}$$

$S_C(X,Y) = \frac{4}{3*2} = \frac{2}{3}$. $S_C(X,Z) = \frac{0}{3*3} = \frac{0}{9}$. Vector Y is the most similar to Vector Z. This is becuase the cosine value of $S_C(X,Y)$ is much greater than $S_C(X,Z)$

5

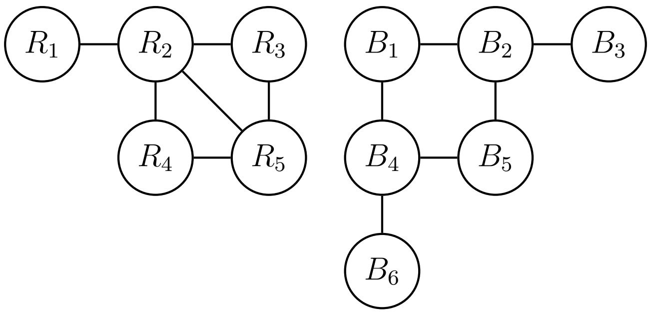


The figure above has a modularity of 0. This is because the Q_R and Q_B are the same. The calculation for both is

$$Q_{(R,B)} = \frac{1}{4} - \left(\frac{2+2}{2*4}\right)^2$$
$$Q_{(R,B)} = \frac{1}{4} - \left(\frac{1}{2}\right)^2$$
$$Q_{(R,B)} = \frac{1}{4} - \left(\frac{1}{4}\right)$$
$$Q_{(R,B)} = 0$$

Adding all of these Q, we get that the modularity of the network is 0

6



The figure above has a modularity of 0.5. The calculation for both is

$$
\begin{aligned}
Q_{(R)} &= \frac{6}{12} - \left(\frac{1+2+2+3+4}{2*12}\right)^2 & Q_{(B)} &= \frac{6}{12} - \left(\frac{1+3+2+2+3+1}{2*12}\right)^2 \\
Q_{(R)} &= \frac{6}{12} - \left(\frac{12}{2*12}\right)^2 & Q_{(B)} &= \frac{6}{12} - \left(\frac{12}{2*12}\right)^2 \\
Q_{(R)} &= \frac{1}{2} - \left(\frac{1}{2}\right)^2 & Q_{(B)} &= \frac{1}{2} - \left(\frac{1}{2}\right)^2 \\
Q_{(R)} &= \frac{1}{2} - \frac{1}{4} & Q_{(B)} &= \frac{1}{2} - \frac{1}{4} \\
Q_{(R)} &= \frac{1}{4} & Q_{(B)} &= \frac{1}{4}
\end{aligned}
$$

Adding all of these Q, we get that the modularity of the network is 0.5