



Genómica Computacional – Tarea 1

Cova Pacheco Felipe de Jesús

2019-2

Profesor: Sergio Hernández López
Ayudante: Juan Antonio Arias del Ángel
Ayudante Lab.: Rafael López Martínez

1.

a) $GF[AT]CC$

b)

2. G C T T G G

3. G C A C G G

Son todas las cadenas que empiecen con G, seguidas de una A o una C, seguidas de muchas T o AC, y que terminan con GG

c)

Se encontró TATATA en la posición 3:9

Se encontró TA en la posición 11:13

Se encontró TA en la posición 14:16

Se encontró TATA en la posición 17:21

La expresión permite que haya TA en la expresión al menos una vez.

d) En 5:8

e) Son cadenas que empiecen con G y terminen con G, sin importar los caracteres que se encuentren entre ellas, y pueden ser 0 o más caracteres.

Ejemplos:

GAAAAG, GG, GA12345G, GAHUEHAHFRYUAGASAG

2.**a)**1. 001011110101011110000000101011

El alfabeto es 0,1

Frecuencia de 0: 0.5

Frecuencia de 1: 0.5

$$H(X) = -[(0,5 \log_2(0,5)) + (0,5 \log_2(0,5))]$$

$$H(X) = -[(-0,5) + (-0,5)]$$

$$H(X) = -[-1]$$

$$H(X) = 1 \text{ bit}$$

2. 100101100111101110111011100111

El alfabeto es 0,1

Frecuencia de 0: 0.333

Frecuencia de 1: 0.667

$$H(X) = -[(0,333 \log_2(0,333)) + (0,667 \log_2(0,667))]$$

$$H(X) = -[(-0,528) + (-0,39)]$$

$$H(X) = -[-0,9183]$$

$$H(X) = 0,91 \text{ bits}$$

3. 000000000000000000000000000000

El alfabeto es 0

Frecuencia de 0: 1

$$H(X) = -[(1 \log_2(1))]$$

$$H(X) = -[(0)]$$

$$H(X) = -[0]$$

$$H(X) = 0 \text{ bit}$$

b) El alfabeto es A,C,G,T, ' '

Frecuencia de A: 0.059

Frecuencia de C: 0.353

Frecuencia de G: 0.176

Frecuencia de T: 0.118

Frecuencia de ' ': 0.294

$$H(X) = -[(0,059 \log_2(0,059)) + (0,353 \log_2(0,353)) + (0,176 \log_2(0,176)) + (0,118 \log_2(0,118)) + (0,294 \log_2(0,294))]$$

$$H(X) = -[(-0,519) + (-0,24) + (-0,53) + (-0,442) + (-0,363)]$$

$$H(X) = -[-2,09486]$$

$$H(X) = 2,09\text{bits}$$

c) Un proceso Bernoulli solo puede tener 2 valores, 0 y 1. Sea X un proceso Bernoulli. $P(X = 1) = p$

$$P(X = 0) = 1 - p$$

La entropía de Shanon para X se define como

$$H(X) = -[(p \log_2(p)) - (1 - p \log_2(1 - p))]$$

Sin embargo, ya sea que $X = 0$ o $X = 1$, la entropía será 0, porque el hecho de asegurar que algo vaya o no vaya a suceder elimina toda la incertidumbre de un evento. Es por esto por lo que cuando $0 \leq p \leq 1$, la entropía es mayor a cero, ya que hay posibilidades de que algo pase o no.