

Tarea N°2

Teoría de la Computación

La facultad de medicina de la UNAB está realizando experimentos para la detección temprana de enfermedades que tienen buenas probabilidades de atenuación o de cura si se tratan antes de que se expresen los primeros síntomas. Para esto utiliza el ADN de las personas y de acuerdo a la secuencia del mismo se puede hacer un pronóstico.

La facultad de medicina ha solicitado la ayuda de los alumnos de Teoría de la Computación para que construyan un programa que los ayude a identificar las 3 secuencias que más son de interés. Las secuencias se centrarán en: adenina→A, timina→T, citosina→C y guanina→G

Las enfermedades a detectar son:

- 1- Alzheimer → Secuencia a reconocer ($A^n C^{2n} T^n G^{3n} / n \geq 1$)
- 2- Parkinson → Secuencia a reconocer ($T^{2n} C^n A^n G^{2n} / n \geq 1$)
- 3- Epilepsia → Secuencia a reconocer ($T^{2n} C^n G^n A^{3n} / n \geq 1$)

Para resolver este problema deberá construir un PDA que reconozca las secuencias indicadas anteriormente y luego implementarlo en un lenguaje de programación que usted seleccione (C, C++, Java o Python).

La entrada será un archivo de texto plano que contendrá una secuencia de ADN y la salida debe indicar la probabilidad de que el individuo en estudio tenga una de las tres enfermedades. El porcentaje se calculará de acuerdo a las cadenas que existan en la secuencia.

Por ejemplo una secuencia de 4 cadenas:

Entrada: *AACCCCTTGGGGGG ACCTGGG AAAAGGTTCCC TTCGAAA*

Salida: Individuo con 50% de probabilidad de Alzheimer y 25% de Epilepsia

Las dos primeras secuencias corresponden al Alzheimer y la última a Epilepsia. La tercera no corresponde a ninguna.