

# Matemáticas Computacionales TC2020

## Tarea EX 02

02.11.18

### Gramáticas Libres de Contexto (30% +9%)

El *Biologic Space Lab* está en problemas nuevamente. Debido a una brecha en un ducto de ventilación, ciertos parásitos lograron entrar a uno de los laboratorios de acceso restringido, en el cual se almacenaban muestras de ADN de especies extraterrestres inteligentes para fines taxonómicos. Algunos de los bancos de datos están corruptos debido a la infección, pero otros aún están a salvo. HQ te ha asignado la tarea de diseñar un sistema de control para identificar aquellas muestras de ADN que aún pueden rescatarse, y te ha enviado la siguiente información:

- Los parásitos tienen la capacidad de alterar el genoma de las especies, copiando y replicando ciertos genes mediante mutación.
- El parásito replica los genes pertenecientes al número de extremidades inferiores y de globos oculares de manera arbitraria.
- Gracias al análisis de los datos, se determinó que los parásitos sólo están afectando a ciertas especies que consideran ‘aptas y balanceadas’, que son aquellas que tienen el mismo número de ojos que de piernas.
- Sin embargo, se ha confirmado que el número de genes replicados pertenecientes a los ojos nunca es igual que el de las extremidades inferiores en el mismo individuo—en un individuo mutado, el número final de ojos nunca es igual al número final de piernas.

Además, dado a la gran bio-diversidad del BSL, HQ te sugiere lo siguiente:

Tenemos cerca de 900,000 muestras en el databank, de las cuáles sólo 487,350 podrían estar infectadas. Nos interesa saber cuáles de esas muestras posiblemente infectadas están sanas para separarlas cuanto antes. Como son tantas especies distintas, creemos que es mejor enfocarse en encontrar diferencias en el número de genes en lugar de buscar una secuencia válida.

Diseña un sistema de reglas (Gramática) para generar subcadenas de genes de ojos y piernas de individuos sanos.

Para probar tu sistema, HQ te envió el ADN de un individuo de la especie *Etecoon*, que es un primate con antena y que por naturaleza debería tener dos ojos y dos piernas. Utilizando las reglas de tu gramática, el sistema intentará generar una palabra idéntica al genoma del *Etecoon*. Si lo logra, el *Etecoon* está sano.

Piensa de manera abstracta. Imagina que el *input* de tu sistema son las subcadenas de ADN específicas de ojos y piernas. Sugerencias:

- Genera el alfabeto de terminales. ¿Qué símbolo usarás para representar ojos y cuál para representar piernas? (2%)
- Describe con un enunciado el lenguaje que debería aceptar tu gramática (5%)
- Genera las reglas de tu GLC (10%)
- Describe formalmente tu GLC usando  $G = (V, \Sigma, R, S)$  (5%)
- Usando tu gramática, escribe el árbol de derivación de un *Dachora* sano, una especie de ave que tiene dos piernas y dos ojos (8%)
- Dibuja un *Etecoon* y un *Dachora* (+4%)
- Haz un Autómata de Pila que acepte el mismo lenguaje que tu GLC (+5%)