

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA

LISTA DE EXERCÍCIOS 1

e1-i1

Pedro Cantarutti - pacantarutti@inf.ufrgs.br

Porto Alegre, 19 de Setembro de 2017

E1-1a

Algoritmo:

- Decidi ler o arquivo em um única string, eliminando caracteres como '\n', '\n' e espaços em branco . Assim, apenas verifiquei, utilizando *find*, comum a objetos do tipo str, se subsequência: *ACCTGGTGTTCCCA*, estava presente.

Resposta:

- Região de localização (posição do primeiro caractere): 12052650
- *Linha:*
TGTCTCTCACTATATACCTGGTGTTCCCAACCTTAGGTTCATAACTAGAATTG
CATAGACTTTTTCACTA

E1-1b

Algoritmo:

- Optei por usar *list comprehension*, basicamente é uma maneira mais enxuta de aplicar uma regra a elementos de uma dada lista. Em seguida, percorri a nova lista, e para cada elemento verifiquei se o mesmo era um palíndromo fazendo uso do *slicing* para obter o mesmo conjunto de caracteres invertido (*reversed*).

Resposta:

- 986

E1-1c

Algoritmo:

- Assim com na questão e1-1b, optei por usar *list comprehension*. Por fim, percorrer essa nova lista e adicionar cada elemento de tamanho 37 em um *set*, em Python *sets* são “listas” sem repetições de elementos.

Resposta:

- 1050079

E1-1d

Algoritmo:

- Optei por utilizar um *dict*, dicionário, para contabilizar a quantidade de cada nucleotídeo. Assim, apenas percorri a lista de caracteres do arquivo lido e fui incrementando a *value*, que inicialmente é zero, da chave correspondente do dict.

Resposta:

- `{'T': 11856330, 'A': 11820665, 'G': 8226382, 'C': 8185246}`
- A: 29.49 %
- T: 29.58 %
- C: 20.42 %
- G: 20.52 %