

# Esame 24/06/2019

Andrea Passerini  
andrea.passerini@unitn.it

Informatica

## Programma python

Scrivere un programma che prenda in ingresso i nomi di due file contenenti ciascuno una sequenza in formato FASTA, ed un valore  $k$ , e stampi:

- le sottosequenze di lunghezza  $k$  per le quali la somma della frequenza con cui appaiono nella prima e nella seconda sequenza è massima
- le frequenze con cui tali sottosequenze appaiono nella prima e nella seconda sequenza

# Esempio esecuzione

```
> python max_common_substring.py
Nome file prima sequenza: fastas/1i1iP.f
Nome file seconda sequenza: fastas/1n7uA.f
Dimensione sottosequenze: 2
substr  1i1iP    1n7uA
AA      7        8
```

## Programma python: suggerimento

Si possono implementare 5 funzioni separate:

- 1 una che legga un file FASTA e restituisca una coppia nome-sequenza
- 2 una che data una sequenza ed un valore  $k$ , restituisca il dizionario con le frequenze delle sue sottosequenze di lunghezza  $k$
- 3 una che dati due dizionari di frequenze, estragga una lista di triplette (sottosequenza, freq\_1, freq\_2) per tutte le sottosequenze a per cui  $\text{freq\_1} + \text{freq\_2}$  è massima
- 4 una che dati i nomi delle sequenze e la lista di sottosequenze più comuni con le loro frequenze, stampi le informazioni come da esempio mostrato
- 5 una (o un main) che realizzi il programma richiesto usando le funzioni di cui sopra

## Programma python: versione semplificata

In caso di difficoltà provare a risolvere una versione leggermente semplificata in cui si assume che la sottosequenza a frequenza massima sia sempre una (la prima trovata, o l'ultima)

## Esercizi da linea di comando

Elencare i nomi dei file nella directory `fastas/` con le 5 sequenze piu' lunghe, in ordine crescente di lunghezza.

## Risultato atteso

```
1ibrB.f  
1dl2A.f  
1fp4B.f  
1n7uA.f  
1i1iP.f
```

## Esercizi da linea di comando

Elencare i nomi dei file all'interno della directory `fastas/` le cui sequenze cominciano con Alanina (A) o Glicina (G) e contengono una coppia di Acido Glutammico (EE).

## Risultato atteso

```
last_.f  
1d8wA.f  
1d12A.f  
...
```

## Modalita' di esecuzione e consegna

- 1 Avviare la macchina in modalita' `ESAME`
- 2 Autenticarsi con nome utente `sci-esame` e password fornita dal docente
- 3 Il testo del compito ed i file necessari si trovano in una cartella `Testo` sul Desktop
- 4 Realizzare il programma python come file `programma.py` e scrivere gli esercizi da linea di comando in un file di testo `linea_di_comando.txt`
- 5 Creare sul Desktop una cartella con *nome\_cognome* e metterci i due file realizzati.
- 6 Eseguire il logout ma NON spegnere la macchina