

## 1 Esercizio Lode

**ATTENZIONE:** Questo esercizio consente di conseguire la lode, e viene valutato se e solo se tutti gli altri esercizi sono stati svolti correttamente.

Un DNA è una lunga catena di unità chiamate nucleotidi. Nel DNA ci sono 4 tipi di nucleotidi denotati con le lettere A, C, G, e T. Gli Esseri Umani (Homo Sapiens) hanno 3 milioni di coppie di nucleotidi Per esempio, una piccola porzione del DNA umano potrebbe essere:

```
ACTAGGGATCATGAAGATAATGTTGGTGTTCAGACAATT
```

In questo esempio, viene denominato k-mer di lunghezza 4 (e viene chiamato 4-mer) una sequenza di 4 nucleotidi consecutivi (i.e. lettere). Alcuni esempi di 4-mers derivati dall'esempio sono:

```
ACTA, CTAG, TAGG, AGGG, GGGA, etc.
```

Completare il programma `lode.cc` inserendo la dichiarazione e la definizione della funzione non ricorsiva `genera_k_mer` che prende come argomento un intero  $k$ , e calcola e stampa a video tutti i possibili k-mer, e ritorni il numero di k-mer calcolati. Matematicamente si tratta di generare tutte le possibili permutazioni di lunghezza  $k$  su insieme di caratteri A, C, G, e T, e corrisponde a generare  $4^k$  possibili k-mer.

**Note:**

- Scaricare il file `lode.cc`, modificarlo solo per inserire la dichiarazione e la definizione della funzione `generare_k_mers`, e caricare il file risultato delle vostre modifiche a soluzione di questo esercizio nello spazio apposito.
- Cercate di usare il numero minore di cicli (while, for, do, ...) innestati possibili. Ad esempio, esiste una possibile soluzione che richiede solo due cicli innestati per il calcolo di tutti i possibili k-mer.
- Un possibile suggerimento per risolvere il problema consiste nell'assumere ordine A, C, G, T per i caratteri; partire da una stringa di lunghezza  $k$  che contiene sole A, e modificarla fino a quando non si ottiene la stringa di lunghezza  $k$  che contiene solo T; , ad ogni cambiamento effettuato si rimpiazza A con C, C con G, G con T, e T con A. Ovviamente, questo è solo un suggerimento, esistono diverse soluzioni del problema.
- All'interno di questo programma **non** è ammesso l'utilizzo di variabili globali o di tipo `static` e di funzioni di libreria al di fuori di quelle definite in `cstring`, `cstdlib`, `iostream`.

Il programma per essere eseguito si aspetta di ricevere come argomento il numero positivo  $k$  di cui generare tutti i k-mer.

Questi sono due esempi di esecuzione (i puntini indicano testo rimosso per rendere leggibile l'output):

```
computer > ./a.out 1
Start
A
C
G
T
Numero di k-mer generati: 4
Finish!
```

```
computer > ./a.out 4
Start
AAAA
CAAA
GAAA
TAAA
ACAA
....
....
ATTT
CTTT
GTTT
TTTT
Numero di k-mer generati: 256
Finish!
```