

BIOINFORMATICA
Compito scritto dell'11 Febbraio 2008

1) Si mostri la matrice di allineamento delle sequenze $s_1=AGCTAA$, $s_2=AGTAT$, utilizzando come score per la edit distance -2 per gli indel, 1 per i match, e -1 per i mismatch.

- a) Qual'è lo score dell'allineamento ottimo?
- b) Si mostri un allineamento ottimo.

2) Si concepisca un algoritmo efficiente che risolva il seguente problema:

INPUT: un vettore A di n colori nell'insieme {giallo, verde, rosso}.

OUTPUT: il colore più frequente.

- a) Si descriva a parole, il più possibile rigorosamente e sinteticamente, l'algoritmo concepito.
- b) Se ne scriva lo pseudocodice.
- c) Si valuti la complessità dell'algoritmo proposto, motivando la risposta.

3) Si mostri il suffix tree generalizzato delle sequenze $s_1=GATTA$ e $s_2=TGATA$.