|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CORSO** | **Informatica – 12BHD** | | | | **DATA** | | **01/09/2014** | | | | |
| **NOME** |  | | | | **COGNOME** | |  | | | | |
| **MATRICOLA** | | **S** |  |  | |  | |  |  |  | **B/1** |
| ☐AAA-BARC ☐BARD-BOUH ☐BOUI-CART ☐CARU-CONS ☐CARU-CONS ☐CONT-DEMAR ☐DEMAS-FERRD ☐FERRE-GIAQ ☐GIAR-LAEZ ☐LAFA-MANC ☐MAND-MIQZ ☐MIRA-PAHZ  ☐PAIA-PODD ☐PODE-ROSSE ☐ROSSF-SIQZ ☐SIRA-TUCB ☐TUCC-ZZZ  ☐Poli@Home ☐5 Crediti ☐AAA-LIB/English ☐LIC-ZZZ/English ☐Altro:................... | | | | | | | | | | | |

|  |  |
| --- | --- |
| **DOMANDA 1** | *Risultato* |
| Dati i seguenti numeri interi N1, N2, N3, N4, espressi in modulo e segno a 8 bit, indicare quale è maggiore e giustificare con i passaggi  N1= 10110001 N3= 01000011  N2= 01101101 N4= 10100111 | Maggiore:  N3  N4  N2  N1 |
| Passaggi | |

|  |
| --- |
| **DOMANDA 2** |
| Dire se le due seguenti espressioni booleane sono equivalenti e spiegare: 1) ; 2) |
| Risposta |

|  |
| --- |
| **DOMANDA 3** |
| Spiegare qual è la differenza tra file sorgente e file eseguibile |
| Risposta |

|  |
| --- |
| **DOMANDA 4 (Programmazione)** |

Scrivere un programma in C che esegua un confronto tra una specifica sottosequenza (chiamata *seed*) e una serie di sequenze di DNA. Il programma deve leggere da un file formato ASCII contenente in ogni riga una sequenza di DNA (ovvero una stringa formata dai caratteri A, T, C, G) di lunghezza 50. **Il file di input, di lunghezza non nota a priori, viene passato come primo argomento sulla linea di comando**. Il seed da ricercare è specificato in una stringa passata **come secondo argomento da linea di comando** dopo il nome del file ed è di lunghezza 10.

Il programma deve cercare una sottosequenza corrispondente al seed specificato all’interno di ogni sequenza di DNA specificata nel file di input. A tale scopo si facciano le seguenti assunzioni:

- Sono ammessi al massimo N “mismatch” ovvero mancate corrispondenze tra le lettere della sequenza e il seed, dove **N è un numero passato come terzo argomento sulla linea di comando**;

- Non sono ammessi mismatch in corrispondenza della prima lettera del seed;

- In presenza di più sottosequenze soddisfacenti i requisiti di ricerca, il programma deve stampare la posizione di partenza di ciascuna sottosequenza con il corrispondente numero di mismatch.

Il programma deve **fornire in output su console la posizione di partenza della sottosequenza trovata, per ogni sequenza di DNA, il numero della riga contenente la sequenza esaminata e il numero di mismatch**. In caso il seed non venga trovato in nessuna sequenza data in input il programma deve stampare il messaggio: “Seed non presente in nessuna sequenza”.

Esempio di file di input:

NOTA:

Le lettere sottolineate rappresentano a titolo di esempio i mismatch tra la sottosequenza e il seed.

|  |
| --- |
| sequenze.txt |
| ATCGTTAATT**ATTGGGATTC**GCTTAGC**ATTGGGATTA**CTTAGCTTAGCTTA  ATCGTTAATTGCTTAGCTTAGCTTAGCTTAGCTTAGCTTAGCTTAGCTTA  ATCGTTAATTGCTTAGCTTA**ATTGGGATCC**ATCGGCTTAATCGGTCTTAA |

Esempio di esecuzione:

C:\> **findseed sequenze.txt ATTGGGATTA 2**

Seed trovato nella sequenza 1 in posizione 11 con 1 mismatch.

Seed trovato nella sequenza 1 in posizione 28 con 0 mismatch.

Seed trovato nella sequenza 3 in posizione 21 con 2 mismatch.

C:\> **findseed sequenze.txt AGGAAGATTA 2**

Seed non presente in nessuna sequenza.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CORSO** | **Informatica – 12BHD** | | | | **DATA** | | **01/09/2014** | | | | |
| **NOME** |  | | | | **COGNOME** | |  | | | | |
| **MATRICOLA** | | **S** |  |  | |  | |  |  |  | **B/2** |
| ☐AAA-BARC ☐BARD-BOUH ☐BOUI-CART ☐CARU-CONS ☐CARU-CONS ☐CONT-DEMAR ☐DEMAS-FERRD ☐FERRE-GIAQ ☐GIAR-LAEZ ☐LAFA-MANC ☐MAND-MIQZ ☐MIRA-PAHZ  ☐PAIA-PODD ☐PODE-ROSSE ☐ROSSF-SIQZ ☐SIRA-TUCB ☐TUCC-ZZZ  ☐Poli@Home ☐5 Crediti ☐AAA-LIB/English ☐LIC-ZZZ/English ☐Altro:................... | | | | | | | | | | | |

|  |  |
| --- | --- |
| **DOMANDA 1** | *Risultato* |
| Dati i seguenti numeri interi N1, N2, N3, N4, espressi in modulo e segno a 8 bit, indicare quale è maggiore e giustificare con i passaggi  N1= 00110001 N3= 11000011  N2= 11101101 N4= 00100111 | Maggiore:  N3  N4  N2  N1 |
| Passaggi | |

|  |
| --- |
| **DOMANDA 2** |
| Dire se le due seguenti espressioni booleane sono equivalenti e spiegare: 1) ; 2) |
| Risposta |

|  |
| --- |
| **DOMANDA 3** |
| Spiegare la differenza tra un “file di testo” e un “file in formato ASCII” |
| Risposta |

|  |
| --- |
| **DOMANDA 4 (Programmazione)** |

Scrivere un programma in C che esegua un confronto tra una specifica sottosequenza (chiamata *seed*) e una serie di sequenze di DNA. Il programma deve leggere da un file formato ASCII contenente in ogni riga una sequenza di DNA (ovvero una stringa) formata dalle lettere A, T, C, G di lunghezza 50. **Il file di input, di lunghezza non nota a priori, viene passato come primo argomento sulla linea di comando**. Il seed da ricercare è specificato da una stringa passata **come secondo argomento da linea di comando** dopo il nome del file ed è di lunghezza 10.

Il programma deve cercare una sottosequenza corrispondente al seed specificato all’interno di ogni sequenza di DNA specificata nel file di input. A tale scopo si facciano le seguenti assunzioni:

- Sono ammessi al massimo N “mismatch” ovvero mancate corrispondenze tra le lettere della sequenza e il seed, dove **N è un numero passato come terzo argomento sulla linea di comando**;

- Non sono ammessi mismatch in corrispondenza dell’ultima lettera del seed;

- In presenza di più sottosequenze (corrispondenti ai criteri di ricerca) trovate in una stessa riga, è sufficiente che il programma identifichi solo la prima sottosequenza trovata;

Il programma deve **fornire in output su console se è stato trovato il seed oppure no, e in caso positivo il numero della riga contenente la sequenza esaminata ed il numero di mismatch** per ogni sequenza di DNA. In caso il seed non venga trovato in nessuna sequenza data in input il programma deve stampare il messaggio: “Seed non presente in nessuna sequenza”.

Esempio di file di input:

NOTA:

Le lettere sottolineate rappresentano a titolo di esempio i mismatch tra la sottosequenza e il seed.

|  |
| --- |
| sequenze.txt |
| ATCGTTAATT**TTGGGATTA**GCTTAGCTTAGCTTAGCTTAGCTTAGCTTA  ATCGTTAATTGCTTAGCTTAGCTTAGCTTAGCTTAGCTTAGCTTAGCTTA  ATCGTTAATTGCTTAGCTTA**TATGGGATTA**ATCGGCTTAATCGGTCTTAA |

Esempio di esecuzione:

C:\> **findseed sequenze.txt ATTGGGATTA 2**

Seed trovato nella sequenza 1 con 1 mismatch.

Seed trovato nella sequenza 3 con 2 mismatch.

C:\> **findseed sequenze.txt AGGAAGATTA 2**

Seed non presente in nessuna sequenza.