

Proszę wpisywać odpowiedzi w ramkach

Zadanie 1 (8 pkt)

Badamy podobieństwo globalne algorytmem Needlemana-Wunscha dla sekwencji BAABB i ABBA. Stosujemy liniową karę za przerwę $d = -N$, gdzie N jest liczbą liter Twojego nazwiska podzieloną na 2, dzielenie całkowite, bez reszty. Macierz podobieństwa pokazano obok.

	A	B
A	$2 * d$	$-d$
B	$-d$	$2 * d$

Uzupełnij macierz podobieństwa, macierz pomocniczą. Podaj najlepsze dopasowanie, jeżeli jest kilka, podaj jedno z nich.

		B	A	A	B	B
A						
B						
B						
A						

Rozwiązanie:

Zadanie 2 (8 pkt)

Seqwencje CCACA, ABCCC i ABCAB użyto do asemblacji pod-grafem de Bruijna (bez wag w krawędziach) 4 rzędu (w wierzchołkach są sekwencje o długości 3 symboli).

Narysuj graf.

Odpowiedz na pytanie dotyczące wyniku asemblacji, wpisując a, b albo c:

- a) wynik nie istnieje,
- b) istnieje dokładnie jedna sekwencja wyjściowa,
- c) istnieje więcej niż jedna sekwencja wyjściowa.

Jeżeli istnieje sekwencja wyjściowa (jedna lub kilka) jest nią:

Zadanie 3 (8 pkt)

Posługujemy się monetami A, B i C, obserwując sekwencje rzutów *ORRR* (orły i reszki). Zakładając, że to doświadczenie jest opisywane ukrytym modelem Markowa przedstawionym odpowiedz na pytania.

$$Q = \{A, B, C\}$$
$$V = \{O, R\}$$
$$P_A = 1, P_B = 0, P_C = 0$$

	A	B	C
A	$\frac{3}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
B	$\frac{1}{5}$	$\frac{3}{5}$	$\frac{1}{5}$
C	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{3}{5}$

	O	R
A	$\frac{1}{2}$	$\frac{1}{2}$
B	1	0
C	0	1

Uzupełnij macierz pomocniczą w algorytmie Viterbiego

	O	R	R	R
A	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
B	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
C	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Podaj najbardziej prawdopodobną sekwencję użytych monet:

<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
----------------------	----------------------	----------------------	----------------------

Zadanie 4 (8 pkt)

Poniżej podano fragment pliku SAM (kolumny POS CIGAR SEQ).
1 6M GAAGTT
2 6M AAGTTG
4 6M GTTGCC
5 6M TTGCC

oraz odpowiadający fragment sekwencji kodującej

>chrZ:1-15
GAAGTCGCCCTCCC

Opis standardu CIGAR oraz tablica kodonów jest na kolejnej stronie.

1. Określ sekwencję aminokwasów, które są kodowane przez tą sekwencję zakładając, że na pozycji 1 zaczyna się nowy kodon

2. Czy dane z pliku SAM pozwalają wnioskować o wystąpieniu wariantu(ów)?

Jeżeli tak: a) określ typ wariantu(ów) b) jaki jest jego (ich) efekt na białko c) czy jest (są) to wariant(y) hetero czy homo-zygotyczny(e)?

3. Zaproponuj modyfikację pliku SAM tak aby znalazły się dowody na wystąpienie jedno-nukleotydowej delecji heterozygotycznej.

Imię i nazwisko: MBI, 17.06.2021,

zadanie 5, 8 pkt.

Dany jest roztwór wodny oligonukleotydów DNA, na którym implementowana jest logika zdań w postaci bazy wiedzy. W bazie tej dane są zdania logiczne w koniunkcyjnej postaci normalnej:

$A = B_1 \text{ and } B_2 \text{ and } \dots \text{ and } B_n$

gdzie B_i - klauzula, która jest alternatywą dowolnej liczby literałów:

$B_i = b_1 \text{ or } b_2 \text{ or } \dots \text{ or } b_m$

przy czym literał to pojedyncza zmienna zdaniowa lub negacja.

Także dany jest enzym BseXI, który rozpoznaje ciąg nukleotydów: 5' - **GCAGC**- 3' .

Tnie on następnie zadaną molekułę DNA w odległości 8 i 12 nukleotydów:

5' - (...) **GCAGC**NNNNNNNNNNNNNNN(...) - 3'

3' - (...) **CGTCG**NNNNNNNNNNNNNNN(...) - 5'

Po rozcięciu za pomocą BseXI dostajemy molekuły:

5' - (...) **GCAGC**NNNNNNNNN-3' 5' - NNNN(...) - 3'

3' - (...) **CGTCG**NNNNNNNNNNNNN-5' 3' - (...) - 5'

gdzie N dowolny z czterech nukleotydów: A, T, C, G,

(...) dowolny ciąg nukleotydów.

Niech baza wiedzy składa się z dwóch faktów: **p**, (**p v q**) kodowanych następująco:

gdzie podkreślenie oznacza negację.

Niech klauzulę faktu **p** reprezentuje molekuła:

5' - (50)CCGGCCG**GCAGC**GCTG - 3'

3' - (50)GGCCGGC**CGTCG**CGAC(**p**) - 5'

natomiast molekuła reprezentująca klauzulę (**p v q**) ma postać:

5' - (**p**) CAGCGCTG

3' - GTCGCGAC (**q**)

Polecenia do wykonania

Dokonać wnioskowania stosując regułę odrywania na powyżej zdefiniowanej bazie wiedzy:

- Dokonać ligacji faktu **p** z klauzulą (**p v q**).
- Dokonać rozcięcia otrzymanego wyniku enzymem BseXI.
- Przedstawić molekułę reprezentującą wniosek częściowy.

Odpowiedź (można kontynuować z drugiej strony):