

Zadanie 3 (3 pkt)

Zbuduj drzewo filogenetyczne metodą UPGMA (średnich połączeń) dla 5 taksonów. Macierz odległości podano obok.

	B	C	D	E
A	4	8	8	8
B		16	8	12
C			5	7
D				2

Zadanie 4 (3 pkt)

Posługujemy się monetami A i B, obserwując sekwencje rzutów *ORRR* (orły i reszki). Podaj najbardziej prawdopodobną sekwencję użytych monet, zakładając, że to doświadczenie może być opisane ukrytym modelem Markowa pokazanym niżej.

$$Q = \{A, B\},$$

$$V = \{O, R\},$$

$$P_A = 1, P_B = 0.$$

	A	B
A	$\frac{4}{5}$	$\frac{1}{5}$
B	$\frac{2}{5}$	$\frac{3}{5}$

	O	R
A	$\frac{1}{2}$	$\frac{1}{2}$
B	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"/>

Podaj najbardziej prawdopodobną sekwencję użytych monet:

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

Pytanie (1 pkt)

Jakie zagadnienie / zagadnienia warto byłoby dodać do MBI

Pytanie (1 pkt)

Ile godzin poświęciłeś na MBI (sumarycznie wykłady, ćwiczenia, praca własna):

Notatki lub uwagi do R. Nowaka:

Imię i nazwisko:

Zadanie 5 (3 pkt)

Poniżej podano fragment sekwencji referencyjnej oraz pliku SAM (kolumny POS i SEQ, odpowiadające pozycji oraz sekwencji) dla czterech kolejnych odczytów.

sekwencja referencyjna:
 >chr1:1-20
 atgctgacatgacccagtc

plik SAM:
 1 atgctgac
 4 cctgacatga
 12 gtccagtc
 7 gacatgtccc

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

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

Jeżeli tak:

- Określ typ wariantu(ów) (SNV / Indel ?),
- Jaki jest ich efekt na białko,
- Czy są to wariant(y) hetero- czy homozygotyczny(e)?

- Dodaj dwa odczyty (podaj ich lokalizacje i sekwencje), które wprowadzą heterozygotyczny wariant na pozycji 10.

Tablica kodonów:

		Second base					
		U	C	A	G		
First base	U	UUU } Phenylalanine F UUC } UUA } Leucine L UUG }	UCU } Serine S UCC } UCA } UCG }	UAU } Tyrosine Y UAC } UAA } Stop codon UAG } Stop codon	UGU } Cysteine C UGC } UGA } Stop codon UGG } Tryptophan W	U	C
	C	CUU } Leucine L CUC } CUA } CUG }	CCU } Proline P CCC } CCA } CCG }	CAU } Histidine H CAC } CAA } Glutamine Q CAG }	CGU } Arginine R CGC } CGA } CGG }	C	A
	A	AUU } Isoleucine I AUC } AUA } AUG } Methionine start codon M	ACU } Threonine T ACC } ACA } ACG }	AAU } Asparagine N AAC } AAA } Lysine K AAG }	AGU } Serine S AGC } AGA } Arginine R AGG }	A	G
	G	GUU } Valine V GUC } GUA } GUG }	GCU } Alanine A GCC } GCA } GCG }	GAU } Aspartic acid D GAC } GAA } Glutamic acid E GAG }	GGU } Glycine G GGC } GGA } GGG }	G	
						U	C
						A	G

Notatki lub uwagi do Tomasza Gambina:

MBI, 02.02.2024, czas: 60 min

Imię i nazwisko:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Zadanie 6 (3 pkt)

Korzystając z metody implementacji biokomputerów na DNA prezentowanej na wykładzie z MBI, których zasada działania opiera się na koncepcji nanotechnologii przepływowej Lab-on-the-chip, gdzie przetwarza się roztwory na zasadzie splatania, odpowiedź na następujące pytania:

- Jakie są podstawowe mechanizmy działania biokomputerów na zasadzie splatania?

- Wyjaśnić, do czego służy molekula terminalna.

- Opisać kodowanie literałów w tej metodzie.

Notatki lub uwagi do J. Mulawki: