Metody Bioinformatyki, część informatyczna, 9 lutego 2015

Imię i nazwisko: .... | **O** | czas: 50 min

Zadanie 1 (7 pkt)

Uzupełnij macierz dla algorytmu, który bada podobieństwo lokalne (algorytm Smitha-Watermana). Podaj najlepsze rozwiązania dla sekwencji **BBABB** oraz **ABBA**. Macierz podobieństwa pokazano obok. Stosujemy liniową karę za przerwę, $\gamma(n) = n * d, d = -2$.

	A	**B**
A	1	-1
B	-1	1

		B	B	A	B	B
	0					
A						
B						
B						
A						

Rozwiązania:

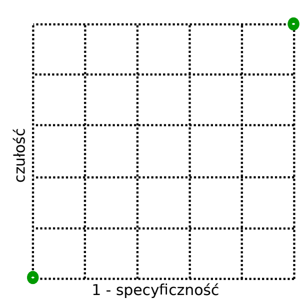
Ilość rozwiązań

Zadanie 2 (8 pkt)

Opracowano test X, który zwraca prawdopodobieństwo wystąpienia choroby. Wyniki testu są przedstawione w tabeli. Podaj macierz pomyłek, zakładając, że traktujemy wynik testu powyżej 0.35 jako pozytywny (osoba chora), następnie dla progu 0.5 oraz progu 0.65. Narysuj dodatkowe 3 punkty na krzywej ROC (dla progu 0.35, 0.5 i 0.65). Narysuj krzywą ROC.

osoba	stan	wynik testu
A	zdrowa	0.1
B	zdrowa	0.1
C	zdrowa	0.2
D	zdrowa	0.3
E	zdrowa	0.6
F	chora	0.6
G	chora	0.7
H	chora	0.8
I	chora	0.9
J	chora	0.9

Próg 0.35:	TP =	FP =
	FN =	TN =
Próg 0.5:	TP =	FP =
	FN =	TN =
Próg 0.65:	TP =	FP =
	FN =	TN =



Notatki lub uwagi do prowadzącego

Zadanie 3 (3 pkt)

Osoba X ma genotyp $A_1A_2B_1C_1C_2$. Obok podano genotypy 6 innych osób. Proszę wskazać osoby, które mogą być rodzicem X lub dzieckiem X.

a) $A_1A_3B_2C_1C_2$

b) $A_1A_3B_1B_2C_2C_3$

c) $A_3B_1C_1C_3$

d) $A_3B_1C_1C_2$

e) $A_1A_3B_1B_2C_1C_3$

f) $A_1A_3B_2C_1C_3$

Zadanie 4 (4 pkt)

Dla populacji obserwujemy dwa loci A i B, każde ma dwa warianty: A i a, oraz B i b. Podaj najbardziej podobny genotyp oraz prawdopodobieństwo jego wystąpienia, jeżeli prawdopodobieństwa haplotypów są następujące: $P(ab) = 0.5$, $P(Ab) = 0.4$ $P(aB) = 0$, $P(AB) = 0.1$. Zakładamy równowagę Hardy'ego-Weinberga dla obserwowanej populacji.

Zadanie 5 (7pkt)

Posługujemy się uczciwą oraz nieuczciwą monetą, obserwując sekwencje rzutów (orły i reszki). Zakładając, że przedstawione doświadczenie jest opisywane ukrytym modelem Markowa przedstawionym obok, podaj najbardziej prawdopodobną sekwencję stanów (sekwencję użytych monet), jeżeli wynikiem doświadczenia jest sekwencja $ROOR$.

- $Q = \{U, N\}$
- $V = \{O, R\}$
- $P_U = 0, P_N = 1$

	U	**N**
U	$\frac{2}{3}$	$\frac{1}{3}$
N	$\frac{1}{3}$	$\frac{2}{3}$

	O	**R**
U	$\frac{1}{2}$	$\frac{1}{2}$
N	$\frac{1}{4}$	$\frac{3}{4}$

Zadanie 6 (6pkt)

Dla zestawu bezbłędnych odczytów: `AABBAA`, `BBAAA`, `AAAABA`, `AAABAAA`, `AAABB`, `BAAAB` zbuduj graf de Bruijna 5 rzędu (w wierzchołkach są sekwencje o długości 4 symboli), a następnie podaj wynikową sekwencję.

Sekwencja:

Czy istnieje tylko jedna poprawna sekwencja wynikowa (tak/nie)?