

Imię i nazwisko:

Zadanie 1 (8 pkt)

Uzupełnij macierz dla algorytmu, który bada podobieństwo lokalne globalne (algorytm Needlemana-Wunscha). Podaj najlepsze rozwiązania dla sekwencji **ACABA** oraz **BABA**. Macierz podobieństwa pokazano obok. Stosujemy liniową karę za przerwę, $\gamma(n) = n * d, d = -3$.

	A	B	C
A	2	-2	-3
B	-2	2	-2
C	-3	-2	2

	A	C	A	B	A	
B	0	-3	-6	-9	-12	-15
A	-3	-2	-5	-8	-7	-10
B	-6	0	-3	-3	-6	-5
A	-9	-3	-2	-5	-1	-4
B	-12	-6	-5	0	-3	1

Rozwiązania:

ACABA
B-ABA

Ilość rozwiązań 1

Zadanie 2 (8 pkt)

Opracowano test, który zwraca prawdopodobieństwo wystąpienia choroby, jego wyniki są przedstawione w tabeli. Podaj macierz pomyłek, zakładając, że traktujemy wynik testu powyżej 0.35 jako pozytywny (osoba chora), następnie podaj tę macierz dla progów 0.45, 0.55 oraz 0.65. Narysuj krzywą ROC, a na niej 6 punktów (dla progów: 0.0, 0.35, 0.45, 0.55, 0.65, 1.0).

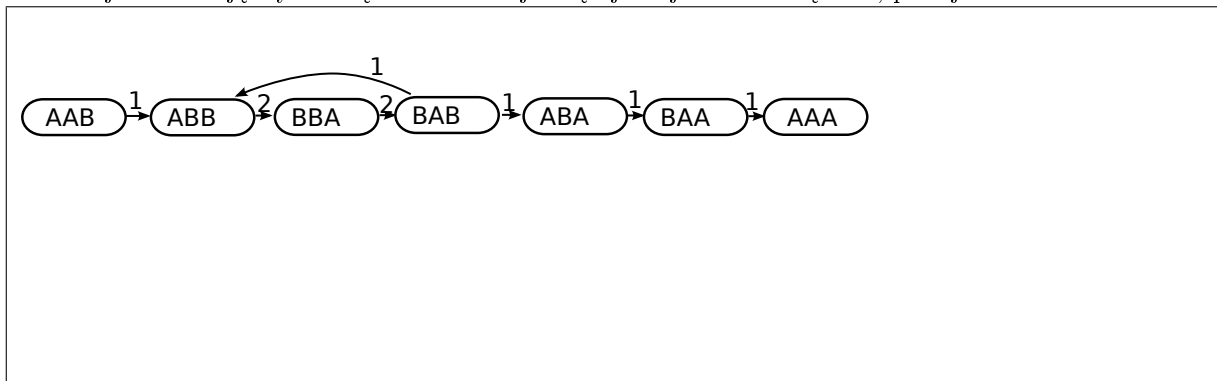
osoba	stan	wynik testu
A	zdrowa	0.1
B	zdrowa	0.2
C	zdrowa	0.3
D	chora	0.4
E	chora	0.5
F	zdrowa	0.5
G	zdrowa	0.6
H	chora	0.7
I	chora	0.8
J	chora	0.9

Próg 0.35:	TP = 5	FP = 2
	FN = 0	TN = 3
Próg 0.45:	TP = 4	FP = 2
	FN = 1	TN = 3
Próg 0.55:	TP = 3	FP = 1
	FN = 2	TN = 4
Próg 0.65:	TP = 3	FP = 0
	FN = 2	TN = 5

Zadanie 3 (8 pkt)

Dla zestawu odczytów: AABB, ABBAB, BABAAA, BABAB, zbuduj multi-graf de Bruijna 4 rzędu (w wierzchołkach są sekwencje o długości 3 symboli, krawędzie przechowują informację, ile razy krawędź ma być dodana do ścieżki wynikowej).

Podaje sekwencję wynikową. Jeżeli istnieje więcej niż jedno rozwiązanie, podaj dowolne z nich.



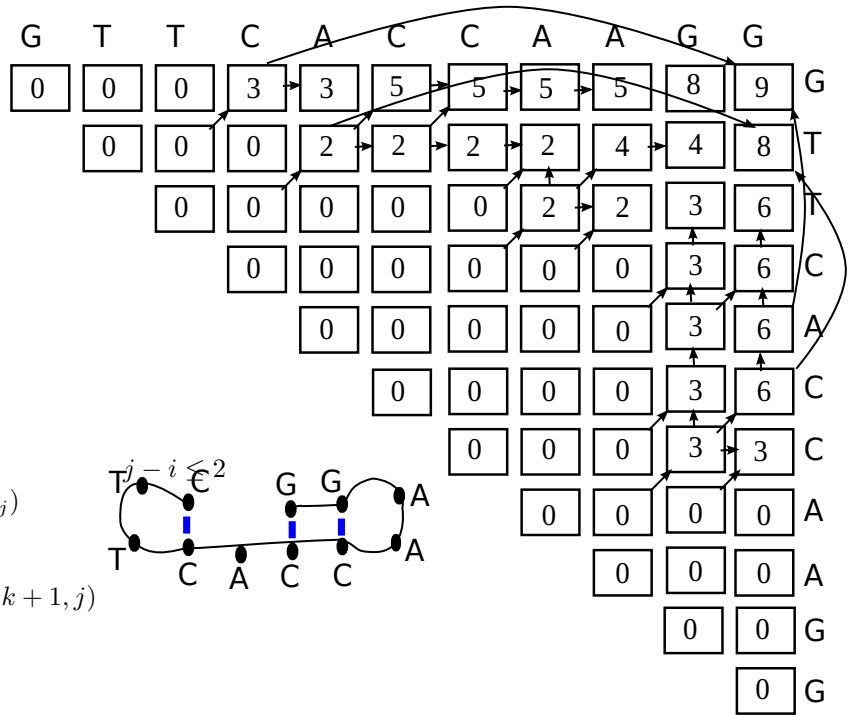
Sekwencja: AABBABBABAAA

Zadanie 4 (8 pkt)

Podaj strukturę drugorzędową cząsteczki GTTCACCAAGG. Wykorzystaj algorytm Nussinov (patrz wzór poniżej) i uzupełnij macierz przedstawioną obok. Macierz energii jest następująca:

	A	C	G	T
A	0	0	0	2
C	0	0	3	0
G	0	3	0	0
T	2	0	0	0

$$F(i, j) = \max \begin{cases} 0 \\ F(i+1, j-1) + e(x_i, x_j) \\ F(i+1, j) \\ F(i, j-1) \\ \max_{k:i \leq k < j} F(i, k) + F(k+1, j) \end{cases}$$



(0,3), (5,10), (6,9)

Zadanie 5 (8 pkt)

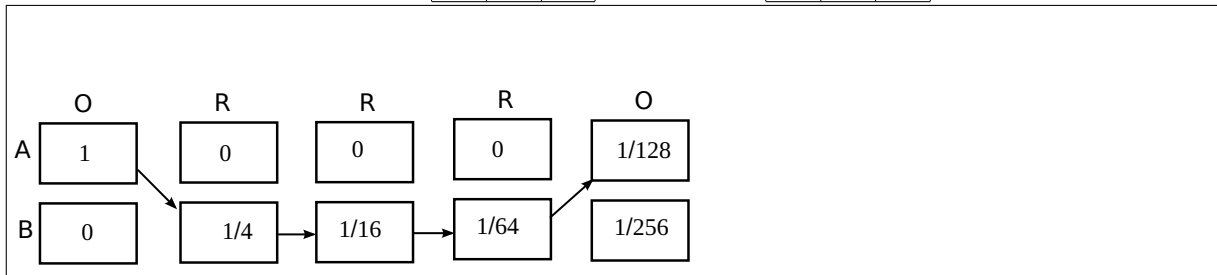
Posługujemy się monetami A i B, obserwując sekwencje rzutów (orły i reszki). Zakładając, że przedstawione doświadczenie jest opisywane ukrytym modelem Markowa przedstawionym obok, podaj wszystkie najbardziej prawdopodobne sekwencje stanów (sekwencję użytych monet), gdy obserwujemy ORRRO.

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

	A	B
A	$\frac{1}{2}$	$\frac{1}{2}$
B	$\frac{1}{2}$	$\frac{1}{2}$

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

ABBBA



Notatki lub uwagi do prowadzącego