

Imię i nazwisko:

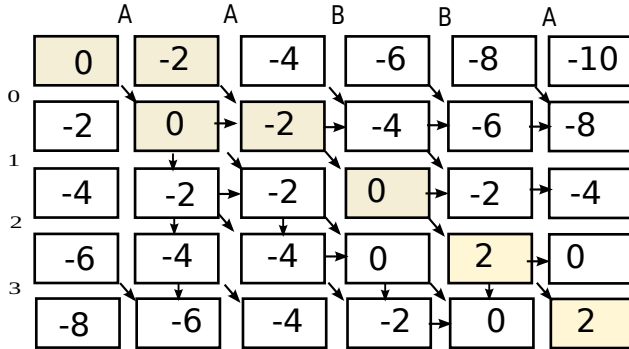
Zadanie 1 (9 pkt)

Uzupełnij macierz dla algorytmu, który bada podobieństwo globalne (algorytm Needlemana-Wunscha). Podaj najlepsze rozwiązania dla sekwencji **AABBA** oraz profilu:

A	0.5	0	0	0.5
B	0.5	1	1	0.5

	A	B
A	2	-2
B	-2	2

Macierz podobieństwa pokazano obok. Stosujemy liniową karę za przerwę, $d = -2$.



Rozwiązania:

A A B B A
- 0 1 2 3

A A B B A
0 - 1 2 3

Ilość rozwiązań: 2

Zadanie 2 (7 pkt)

Opracowano test, który zwraca prawdopodobieństwo wystąpienia choroby, jego wyniki są przedstawione w tabeli. Podaj macierz pomyłek, zakładając, że traktujemy wynik testu powyżej 0.55 jako pozytywny (osoba chora), następnie podaj tę macierz dla progu 0.15 oraz 0.25. Narysuj krzywą ROC, a na niej 5 punktów (dla progów: 0.0, 0.15, 0.25, 0.55, 1.0).

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

Próg 0.15:

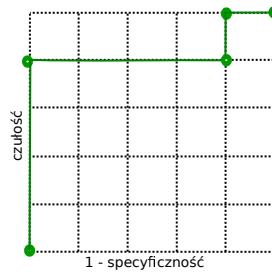
TP = 5	FP = 4
FN = 0	TN = 1

Próg 0.25:

TP = 4	FP = 4
FN = 1	TN = 1

Próg 0.55:

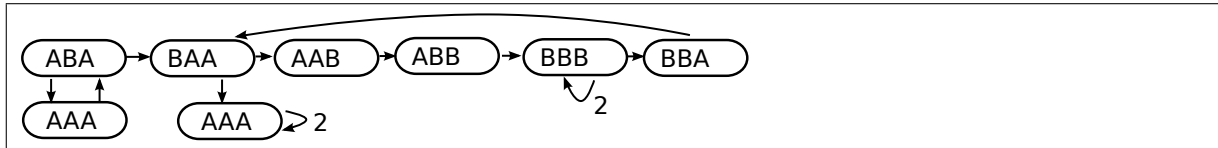
TP = 4	FP = 0
FN = 1	TN = 5



Zadanie 3 (8 pkt)

Dla zestawu odczytów: BAAAAA, ABBBBBA, ABAABB, ABABA, BBAA, 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.

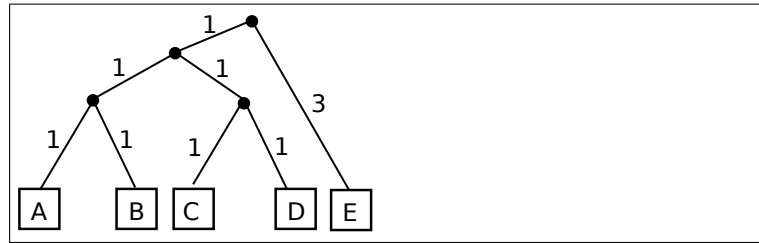


Sekwencja: **ABABAABBBBBBAAAAA**

Zadanie 4 (7 pkt)

Podaj drzewo filogenetyczne dla taksonów A, B, C, D, E dla pokazanej macierzy odległości. Wykorzystaj metodę grupowania parami ze średnią arytmetyczną (UPGMA).

	B	C	D	E
A	2	5	3	5
B		3	5	7
C			2	7
D				5



Zadanie 5 (9 pkt)

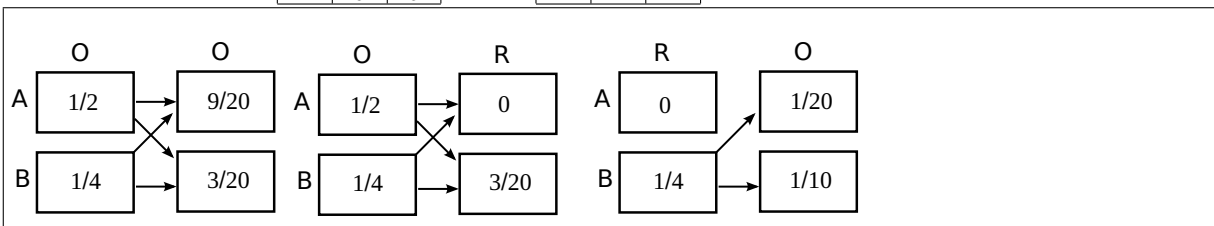
Posługujemy się monetami A i B, obserwując sekwencje rzutów (orły i reszki). Zakładając, że to doświadczenie jest opisywane ukrytym modelem Markowa przedstawionym niżej, podaj prawdopodobieństwo wyrzucenia sekwencji OO. Następnie podaj prawdopodobieństwo wyrzucenia sekwencji OR oraz prawdopodobieństwo wyrzucenia sekwencji RO.

$Q = \{A, B\}$
 $V = \{O, R\}$
 $P_A = \frac{1}{2}, P_B = \frac{1}{2}$

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

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

$P_{OO} = 0.6$
 $P_{OR} = 0.15$
 $P_{RO} = 0.15$



Notatki lub uwagi do prowadzącego