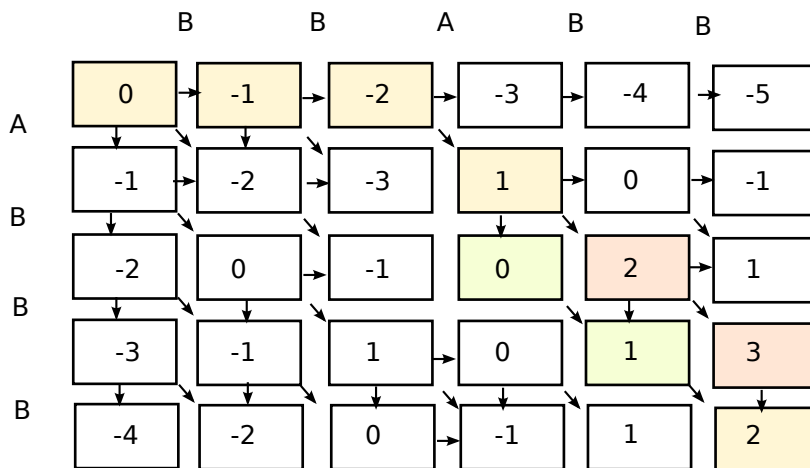


**Zadanie 1 (8 pkt)**

Uzupełnij macierz dla algorytmu, który bada podobieństwo globalne (algorytm Needlemana-Wunscha). Podaj najlepsze rozwiązania dla sekwencji **BBABB** oraz **ABBB**. Macierz podobieństwa pokazano obok. Stosujemy liniową karę za przerwę,  $\gamma(n) = n * d, d = -1$ .

	<b>A</b>	<b>B</b>
<b>A</b>	3	-2
<b>B</b>	-2	1



Rozwiązania:

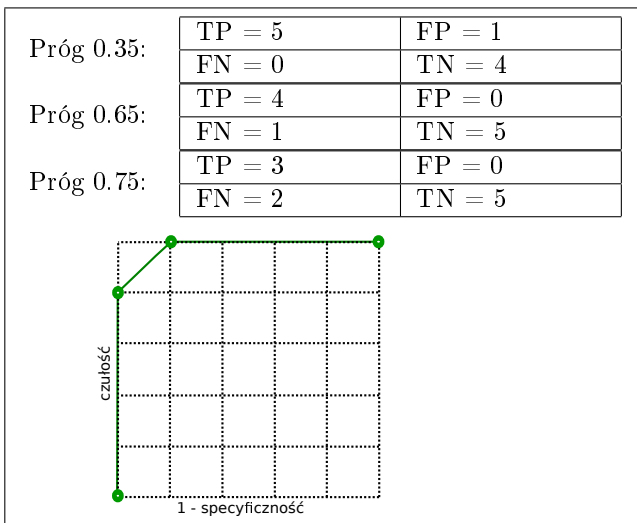
BBABB-	BBA-BB	BBAB-B
--ABBB	--ABBB	--ABBB

Ilość rozwiązań **3**

**Zadanie 2 (9 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 podaj tę macierz dla progu 0.65 oraz progu 0.75. Narysuj 3 punkty na krzywej ROC (0.35, 0.65, 0.75). Narysuj krzywą ROC.

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



Notatki lub uwagi do prowadzącego

**Zadanie 3 (4 pkt)**

Locus  $A$  posiada warianty  $A_1, A_2, A_3$ , locus  $B$  posiada warianty  $B_1$  i  $B_2$ , zaś locus  $C$  posiada warianty  $C_1, C_2$ . Zbadano mieszaninę DNA od 2 osób, uzyskano profil  $A_1A_2A_3B_1C_1C_2$ . Który genotyp nie może wchodzić w skład mieszaniny, jeżeli jedna z osób (ofiara) ma genotyp  $A_1B_1C_1C_2$ .

- a)  $A_2A_3B_2C_1C_2$
- b)  $A_2A_3B_1C_1$
- c)  $A_2B_1B_2C_2$
- d)  $A_2A_3B_1B_2C_1C_2$
- e)  $A_2A_3B_1C_2$

a) c) d)
----------------

**Zadanie 4 (3 pkt)**

Dla populacji obserwujemy dwa loci  $A$  i  $B$ , każde ma dwa warianty:  $A$  i  $a$ , oraz  $B$  i  $b$ . Podaj najbardziej prawdopodobny genotyp (lub genotypy). Prawdopodobieństwa haplotypów są następujące:  $P(ab) = 0.4, P(Ab) = 0.4, P(aB) = 0.2, P(AB) = 0$ . Zakładamy równowagę Hardy’ego-Weinberga dla obserwowanej populacji.

<b><math>P(AaB) = 0.32</math></b>
-----------------------------------

**Zadanie 5 (9 pkt)**

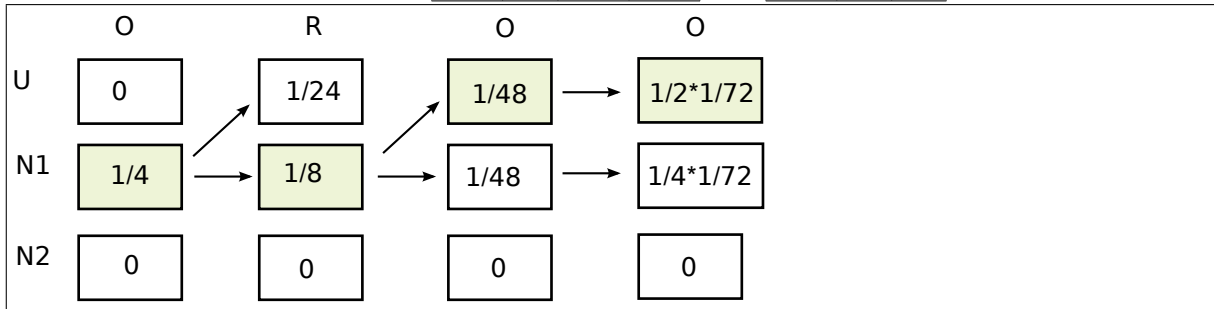
Posługujemy się dwoma monetami, uczciwą oraz nieuczciwą, obserwując sekwencje rzutów (orły i reszki). Zakładając, że przedstawione doświadczenie jest opisywane ukrytym modelem Markowa przedstawionym obok, podaj najbardziej prawdopodobną sekwencję stanów (sekwencję użytych monet), jeżeli wynikiem doświadczenia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

- $Q = \{U, N1, N2\}$
- $V = \{O, R\}$
- $P_U = 0, P_{N1} = 1, P_{N2} = 0$

	U	N1	N2
U	$\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{6}$
N1	$\frac{1}{3}$	$\frac{2}{3}$	0
N2	$\frac{1}{3}$	0	$\frac{2}{3}$

	O	R
U	$\frac{1}{2}$	$\frac{1}{2}$
N1	$\frac{1}{4}$	$\frac{3}{4}$
N2	0	1

<b>N1 N1 U U</b>
------------------

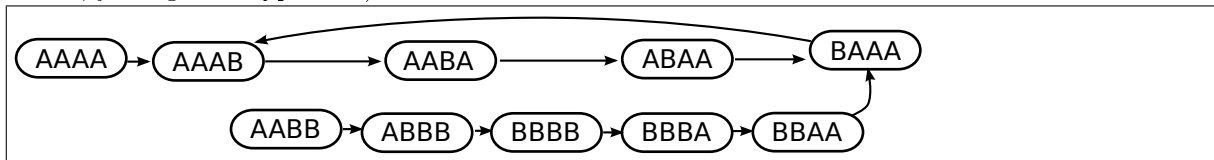


**Zadanie 6 (7pkt)**

Dla zestawu odczytów: AAAABA, ABBBBAAA, AAABA, AABAAA, BAAAB, AABBB zbuduj graf de Bruijna 5 rzędu (w wierzchołkach są sekwencje o długości 4 symboli).

Czy istnieje ścieżka Eulera? (tak/nie)? **nie**

Podaj sekwencję wynikową, która jest najdłuższą ścieżką w grafie (w szczególności jest to ścieżka Eulera, jeżeli graf taką posiada).



Sekwencja: **AABBBBBAAABAAA**