

Q1) Calculate the scoring matrix S for following multiple alignment.

```

I A A V
I I A I
V A V I
I A V I

```

$S_{ij} = 2\log_2(P_{ij}/Q_iQ_j)$, where P_{ij} is the probability of substituting aa i with aa j. Q_i and Q_j are amino acid frequencies. $\log_2(x) = \ln(x)/\ln(2)$

$N_{II} =$	$P_{II} =$	$Q_I Q_I =$	$Q_I =$
$N_{IA} =$	$P_{IA} =$	$Q_I Q_A =$	$Q_A =$
$N_{IV} =$	$P_{IV} =$	$Q_I Q_V =$	$Q_V =$
$N_{AA} =$	$P_{AA} =$	$Q_A Q_A =$	
$N_{AI} =$	$P_{AI} =$	$Q_A Q_I =$	
$N_{AV} =$	$P_{AV} =$	$Q_A Q_V =$	
$N_{VV} =$	$P_{VV} =$	$Q_V Q_V =$	
$N_{VI} =$	$P_{VI} =$	$Q_V Q_I =$	
$N_{VA} =$	$P_{VA} =$	$Q_V Q_A =$	

	I	A	V
I			
A			
V			

Q2) Align the two sequences seq1 and seq2 using scores from the Blosum50 matrix and a gap penalty of -2.

Blosum50 Matrix

A	5																			
R	-2	7																		
N	-1	-1	7																	
D	-2	-2	2	8																
C	-1	-4	-2	-4	13															
Q	-1	1	0	0	-3	7														
E	-1	0	0	2	-3	2	6													
G	0	-3	0	-1	-3	-2	-3	8												
H	-2	0	1	-1	-3	1	0	-2	10											
I	-1	-4	-3	-4	-2	-3	-4	-4	-4	5										
L	-2	-3	-4	-4	-2	-2	-3	-4	-3	2	5									
K	-1	3	0	-1	-3	2	1	-2	0	-3	-3	6								
M	-1	-2	-2	-4	-2	0	-2	-3	-1	2	3	-2	7							
F	-3	-3	-4	-5	-2	-4	-3	-4	-1	0	1	-4	0	8						
P	-1	-3	-2	-1	-4	-1	-1	-2	-2	-3	-4	-1	-3	-4	10					
S	1	-1	1	0	-1	0	-1	0	-1	-3	-3	0	-2	-3	-1	5				
T	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	2	5			
W	-3	-3	-4	-5	-5	-1	-3	-3	-3	-3	-2	-3	-1	1	-4	-4	-3	15		
Y	-2	-1	-2	-3	-3	-1	-2	-3	2	-1	-1	-2	0	4	-3	-2	-2	2	8	
V	0	-3	-3	-4	-1	-3	-3	-4	-4	4	1	-3	1	-1	-3	-2	0	-3	-1	5
	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V

V

W

A

A

S

S

V

A

L