BMI/CS 576
HW1 Solution

Problem 2. Global alignment (10 points)

		A	G	A	T	T
	0	-3	-6	-9	-12	-15
A	-3	**1**	-2	-5	-8	-11
G	-6	-2	**2**	**-1**	-4	-7
T	-9	-5	-1	1	**0**	-3
T	-12	-8	-4	-2	2	**1**

x: A G A T T
y: A G - T T

Problem 3. Local alignment (10 points)

		G	A	A	G	A	G
	0	**0**	0	0	0	0	0
A	0	0	**1**	1	0	1	0
A	0	0	1	**2**	0	1	0
G	0	1	0	0	**3**	0	2
C	0	0	0	0	0	2	0

x: AAG
y: AAG

Problem 4. Concepts in probability (9 points)

(a) P(S=chips) = 3/9, P(S=cereal) = 6/9

(b) P(S=chips, D=coffee) = 2/9, P(S=chips, D=water)= 1/9,
P(S=cereal, D=coffee) = 4/9, P(S=cereal, D=water) = 2/9

(c) P(S=chips) = P(S=chips, D=coffee) + P(S=chips, D=water) = 3/9
P(S=cereal) = P(S=cereal, D=coffee) + P(S=cereal, D=water) = 6/9

(d) S is independent of D since for all s and d, P(S=s,D=d) = P(S=s)*P(D=d).

(e) S is not independent of D given T=morning, since
P(S=s,D=d | T=morning) is not P(S=s | T=morning)*P(D=d | T=morning).
P(S=cereal | T=morning) = 4/5
P(D=coffee | T=morning) = 4/5
P(S=cereal | T=morning) * P(D=coffee | T=morning) = 4/5 *4/5 = 15/25
P(S=cereal, D=coffee | T=morning) = 3/5

(f) S is independent of D given T=noon, since for all s and d, P(S=s,D=d | T=noon) = P(S=s | T=noon)*P(D=d | T=noon).

Problem 5. Applying concepts of probability to sequence alignment (6 points)

(a)
Related model (p_{xy}):

	A	T	G	C
A	10/145	20/145	10/145	5/145
T		30/145	20/145	10/145
G			15/145	10/145
C				15/145

Unrelated model (q_a):

q_A	55/290
q_T	110/290
q_G	70/290
q_C	55/290

(b)

$$S(x,y) = \sum_{i=1}^{n} log\left(\frac{p_{x_i y_i}}{q_{x_i} q_{y_i}}\right)$$

$$S(ATA, GTC) = log\left(\frac{p_{AG}}{q_A q_G}\right) + log\left(\frac{p_{TT}}{q_T q_T}\right) + log\left(\frac{p_{AC}}{q_A q_C}\right) = 0.3174 > 0$$

Since the log odd is positive, the alignment was more likely to be generated by the related model.