

**Problems for 640:338, Spring 2008;      Week 13**

(Note: do the two global alignment problems. Only attempt the local alignment and overlap problems once that the material has been covered in class.)

**1.** Given two protein sequences,  $x = \text{HWAEGGKL}$  and  $y = \text{PAWLGEK}$ , use BLOSUM50 and a linear gap penalty with  $d = 6$  to solve

- a) Draw and score a path for the global alignment
 

$H - \text{WAEGGKL}$	
$\text{PAW} - \text{LGEK} -$	·
  
- b) Draw and score a path for the local alignment
 

$A - \text{EGGK}$	
$\text{AWL} - -G$	·
  
- c) Draw and score a path for the overlap alignment

$\text{HWAEGGKL}$   
 $\text{PAWL} - \text{GE} - -K$

**2.** a) (From Ewens and Grant, p.215) Using BLOSUM50, and a linear gap penalty with  $d=5$ , find all highest scoring global alignments between  $\text{EATGHAG}$  and  $\text{EEAWHEAE}$ .

b) Find the highest scoring repeat match alignment using a linear gap penalty with  $d = 6$  and a motif alignment closing penalty  $-T$  with  $T = 8$ .

**3.** a) (From Ewens and Grant, p.215) Find the best local alignments of the DNA sequences  $\text{AAGTATCG}$  and  $\text{AAGTTAGTT}$  with the scoring scheme  $S(x, x) = 1$ ,  $S(x, y) = -1$  if  $x \neq y$ , and  $d = 2$ ; (use a linear gap penalty).

b) Using the same data, find the the highest scoring overlap alignment.