

Homework #7, CS 329E

Due: March 27, 2008.

1. Let  $u, v, w, x$  be a set of sequences with  $u = AATAG, v = AATGG, w = ACGAG, x = TCCTA$ . Is there any tree on which all the characters are compatible? Find the solution to maximum parsimony on this dataset.
2. Give an example of a set of four sequences for which all three trees are solutions to maximum parsimony (i.e., they all have the optimal score).
3. Give an example of a set of four sequences for which there are exactly two different trees, each with the optimal score.
4. Consider the recurrence relation  $t(n) = (2n-5)t(n-1)$ , with  $t(1) = t(2) = t(3) = 1$ .
  - Give  $t(i)$  for  $i = 3, 4, 5, 6, 7$ .
  - Find a closed form formula for  $t(n)$ .
  - Extra credit: how is this related to the number of unrooted trees on  $n$  leaves, with the leaves labelled  $1, 2, \dots, n$ ?