Sankoff's algorithm – recurrence relation

- Traverse the tree in post-order and update $s(v,t)$ as follows

–assume node v has children u and w

–$s(v,t) = \min_i \{s(u,i) + score(i,t)\} + \min_j \{s(w,j) + score(j,t)\}$

- the minimum parsimony score is given by the smallest score $s(root,t)$ over all symbols t

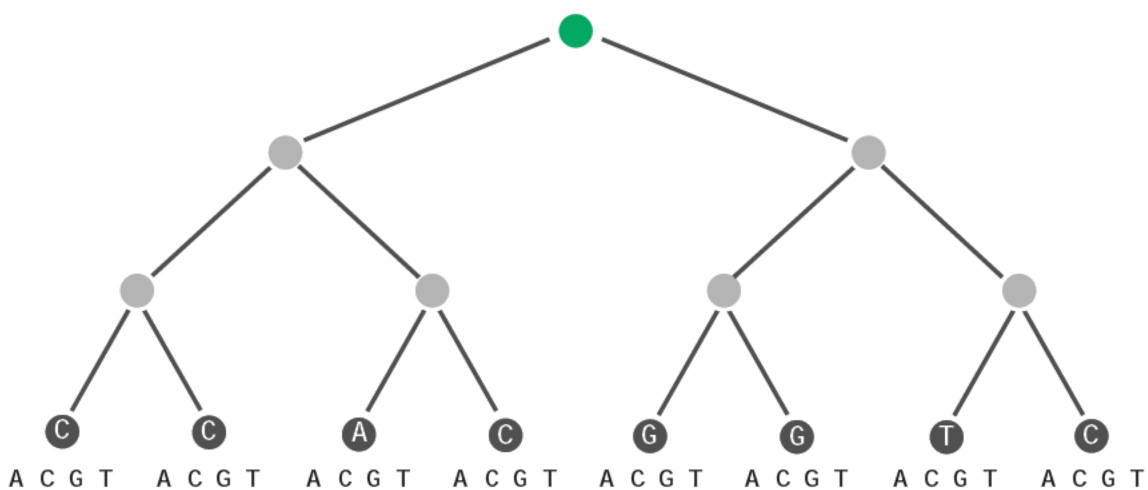
- backtrack to fill the values for the other nodes

- Note – this solves the weighted version. For unweighted set score $(i,i) = 0$, $score(i,j) = 1$ for any i,j

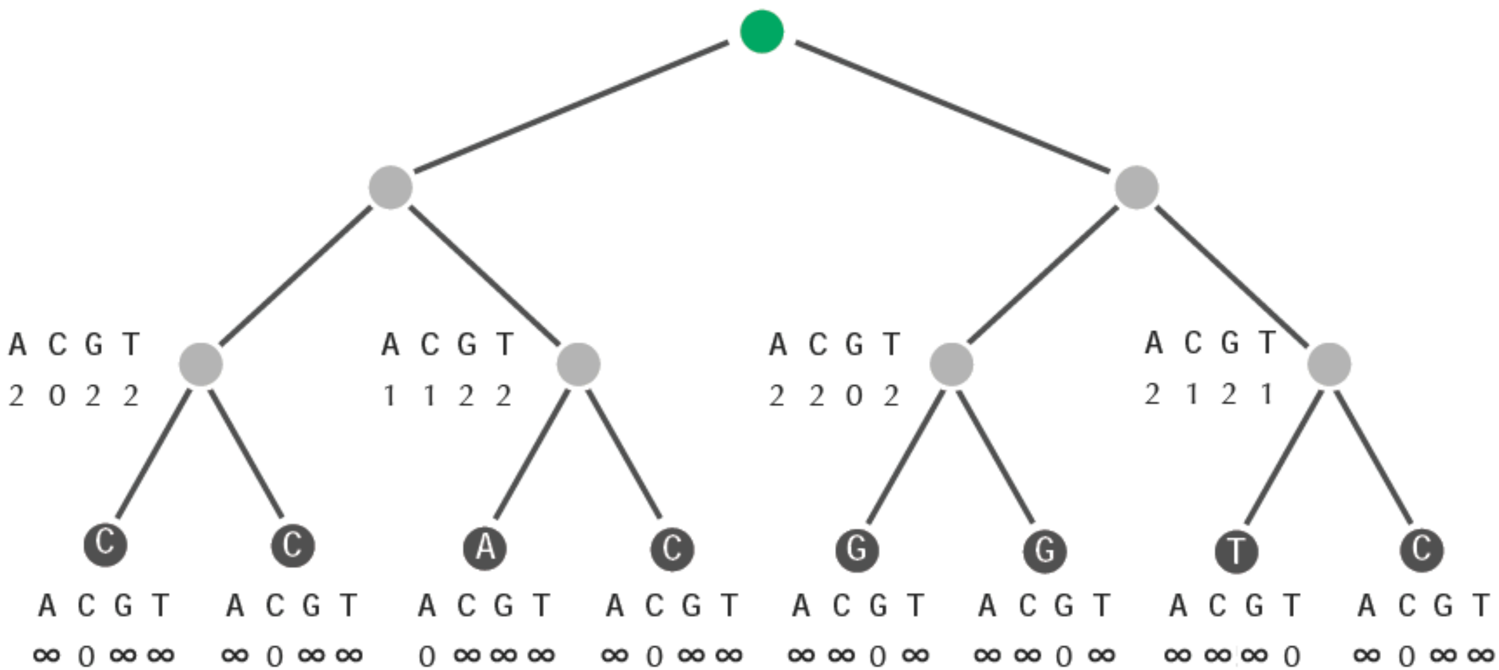
Sankoff's algorithm – recurrence relation

- At each node v in the tree store s(v,t) – best parsimony score for subtree rooted at v if character stored at v is t

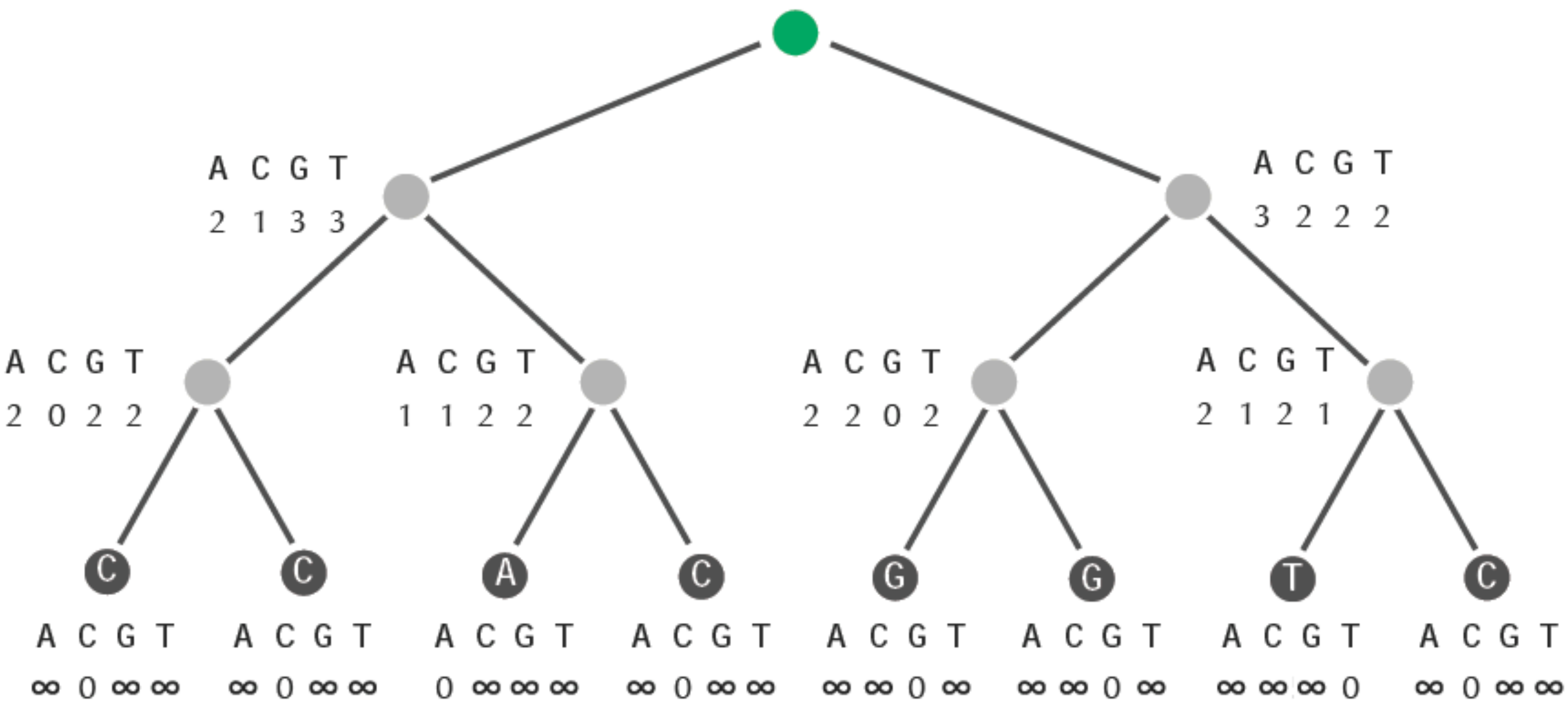
What is the base case?



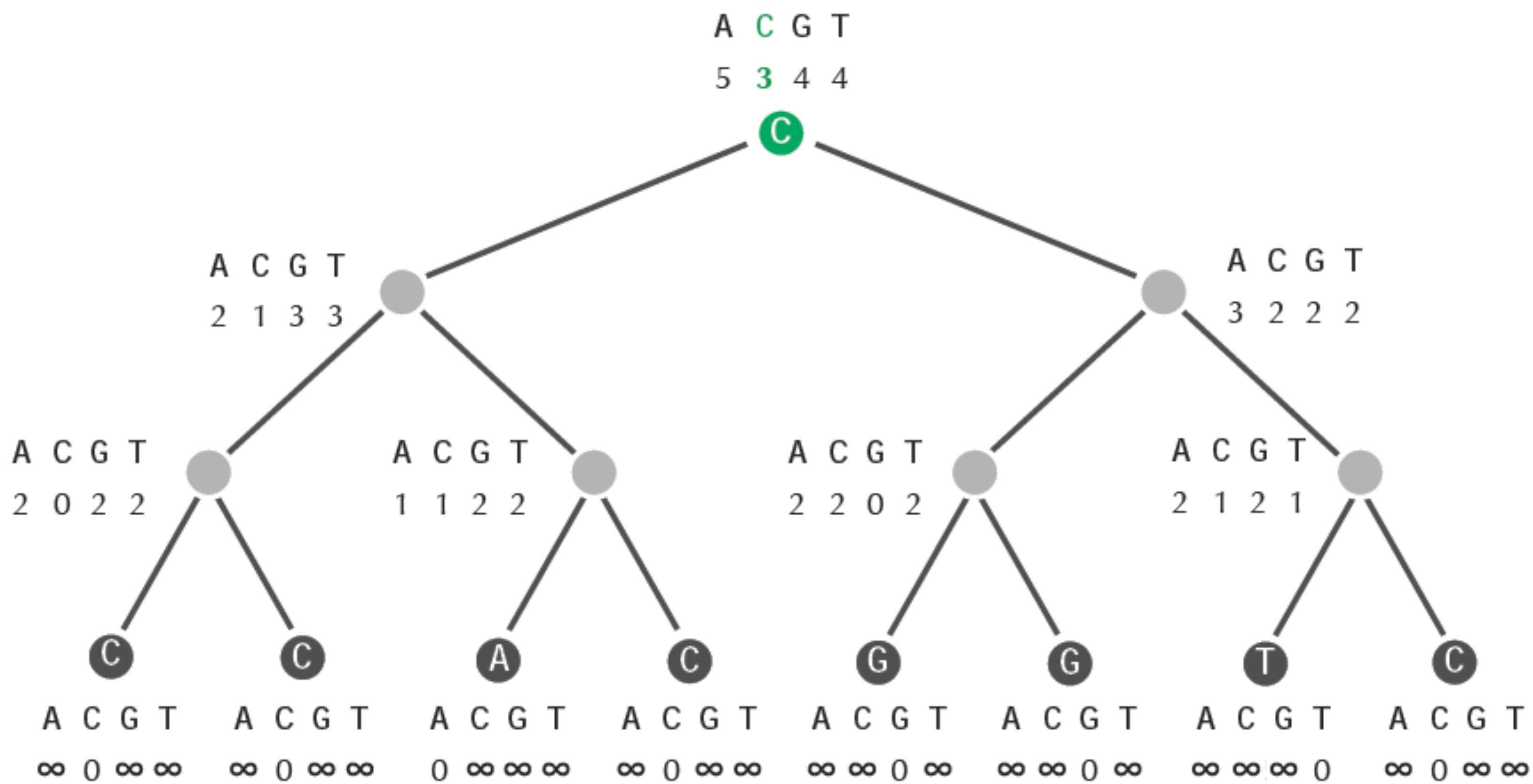
Sankoff's algorithm – example (continued)



Sankoff's algorithm – example (continued)



Sankoff's algorithm – example (continued)



Optimal labeling can be computed in linear time O(nk) where n is number of leaves and k is number of character states