

# Maximum Parsimony

- **Input:** Set  $S$  of  $n$  aligned sequences of length  $k$
- **Output:**
  - A phylogenetic tree  $T$  leaf-labeled by sequences in  $S$
  - additional sequences of length  $k$  labeling the internal nodes of  $T$

such that 
$$\sum_{(i,j) \in E(T)} H(i,j)$$

is minimized, where  $H(i,j)$  denotes the Hamming distance between sequences at nodes  $i$  and  $j$