

Phylogenetic Trees: NJ

You are given the dissimilarity map δ below for the samples $\mathcal{X} = \{A, B, C, D, E\}$. Fill in the steps below to create the Neighbor-Joining tree for \mathcal{X} . First, in the space to the right of δ , draw the initialization of the NJ tree and let (lowercase) c be the center node. Write out N_c and $|N_c|$. To start, we let $d = \delta$.

δ	A	B	C	D	E
A	0	1	3	6	6
B		0	2	5	5
C			0	5	5
D				0	2
E					0

1. (a) The tables below show S and Q for the first iteration. Verify that S_C and $Q(D, E)$ are correct.

i	A	B	C	D	E
S_i	16	13	15	18	18

Q	B	C	D	E
A	-26	-22	-16	-16
B		-22	-16	-16
C			-18	-18
D				-30

- (b) Join D and E at internal vertex v and draw the new tree. Then calculate:

- $d(D, v) =$
- $d(E, v) =$

- (c) Fill in the updated d matrix:

d	A	B	C	v
A	0	1	3	
B		0	2	
C			0	
v				0

2. (a) Second iteration: verify S_v , $Q(A, B)$, and $Q(C, v)$.

i	A	B	C	v
S_i	9	7	9	13

Q	B	C	v
A	-14	-12	-12
B		-12	-12
C			-14

(b) We will arbitrarily choose A and B to join at vertex w . Draw the new tree and calculate:

- $d(A, w) =$
- $d(B, w) =$

(c) Fill in the updated d matrix:

d	w	C	v
w	0		
C		0	4
v			0

3. (a) Third iteration: fill in the tables below.

i	w	C	v
S_i			

Q	C	v
w		
C		

(b) We will arbitrarily choose w and C to combine, but we don't change the topology since they are already joined at center vertex c . We do need to fill in the remaining edge weights:

- $d(w, c) =$
- $d(C, c) =$

(c) To fill in the last edge weight we update the distance matrix d :

d	c	v
c	0	
v	0	4