

The Neighbor-Joining Algorithm

Episode 6789
of the famous TELEKOLLEG Series
First aired on February 25, 2004

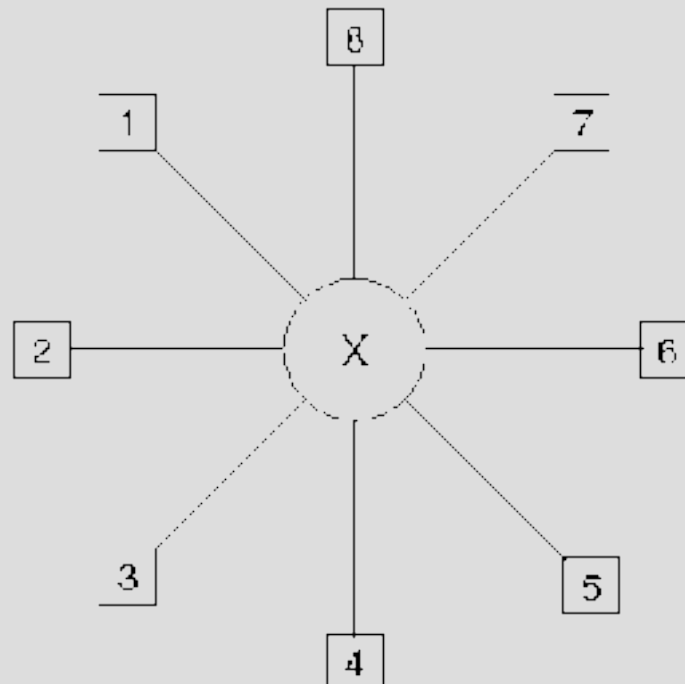
NJ: Introduction

- Objective
 - Tree reconstruction for N objects
 - Topology
- Conditions
 - Distance metric for *any* two of total objects
 - OTU-criterion
- Function
 - Branch length minima* (RMSD)
- Result
 - Unique tree
 - ME

NJ: Algorithm (SN)

- Iterations
 - Determine S_{ij}
 - Find i and j with $S_{ij} = \min (ij = \text{OUT})^*$
 - Introduce inner node replacing ij
- Termination condition
 - $N = 3$

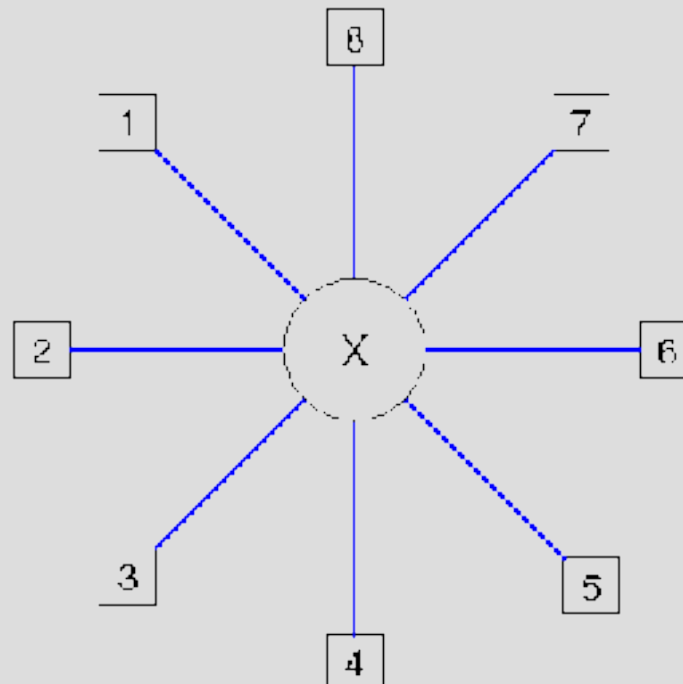
NJ: Example (SN)



Dij=

0								
7	0							
8	5	0						
11	8	5	0					
13	10	7	8	0				
16	13	10	11	5	0			
13	10	7	8	6	9	0		
17	14	11	12	10	13	8	0	

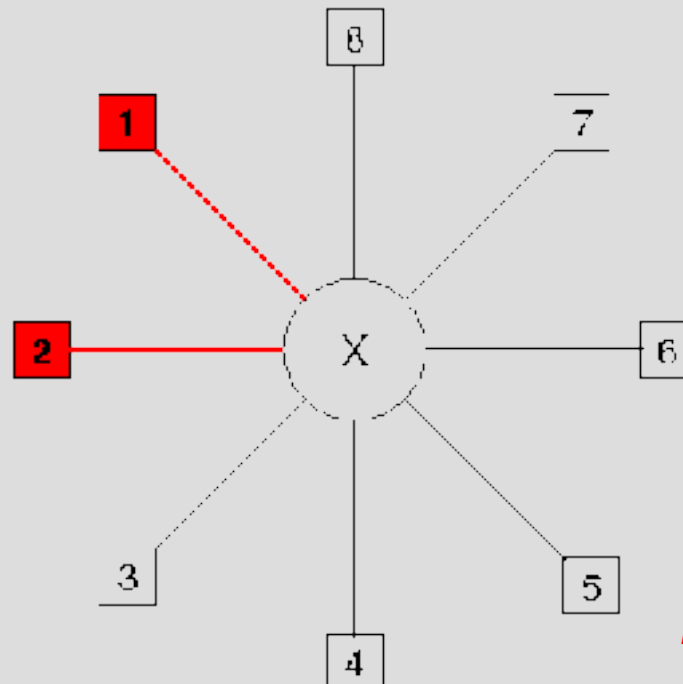
NJ: Example (SN)



S_{ij}=

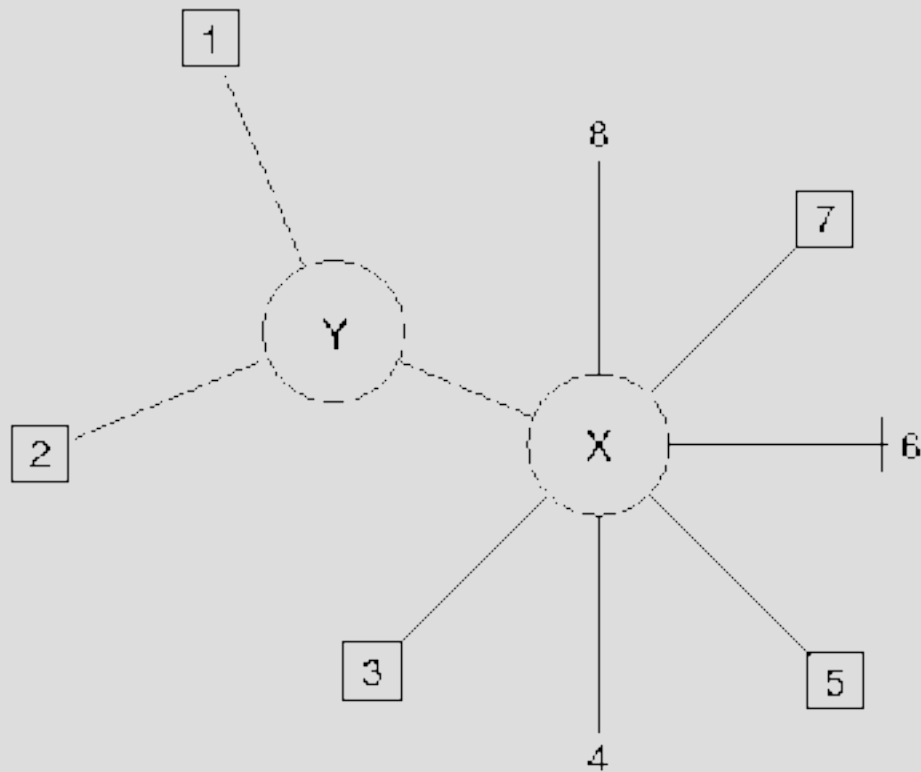
	0							
36.67	0							
38.33	38.33	0						
39.00	39.00	38.67	0					
40.33	40.33	40.00	39.67	0				
40.33	40.33	40.00	39.67	37.00	0			
40.17	40.17	39.83	39.50	38.83	38.83	0		
40.17	40.17	39.83	39.50	38.83	38.83	37.57	0	

NJ: Example (SN)

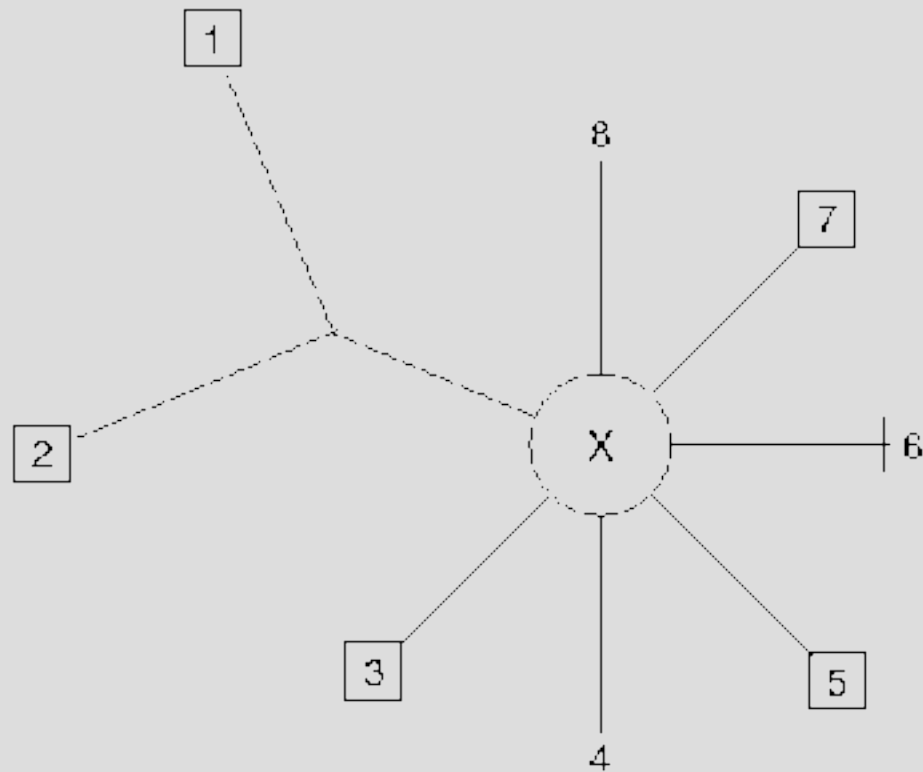


	<i>i=1</i>	2	3	4	5	6	7	8	
	0								<i>j=1</i>
	36.67	0							2
	38.33	38.33	0						3
	39.00	39.00	38.67	0					4
<i>S_{ij}</i> =	40.33	40.33	40.00	39.67	0				5
	40.33	40.33	40.00	39.67	37.00	0			6
	40.17	40.17	39.83	39.50	38.83	38.83	0		7
	40.17	40.17	39.83	39.50	38.83	38.83	37.57	0	8

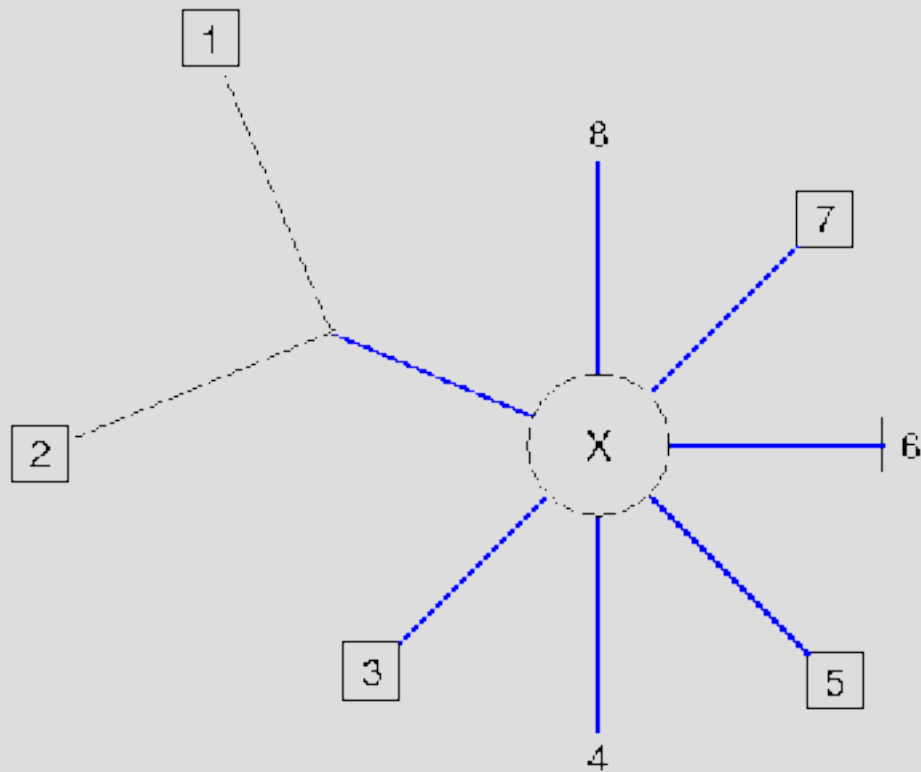
NJ: Example (SN)



NJ: Example (SN)

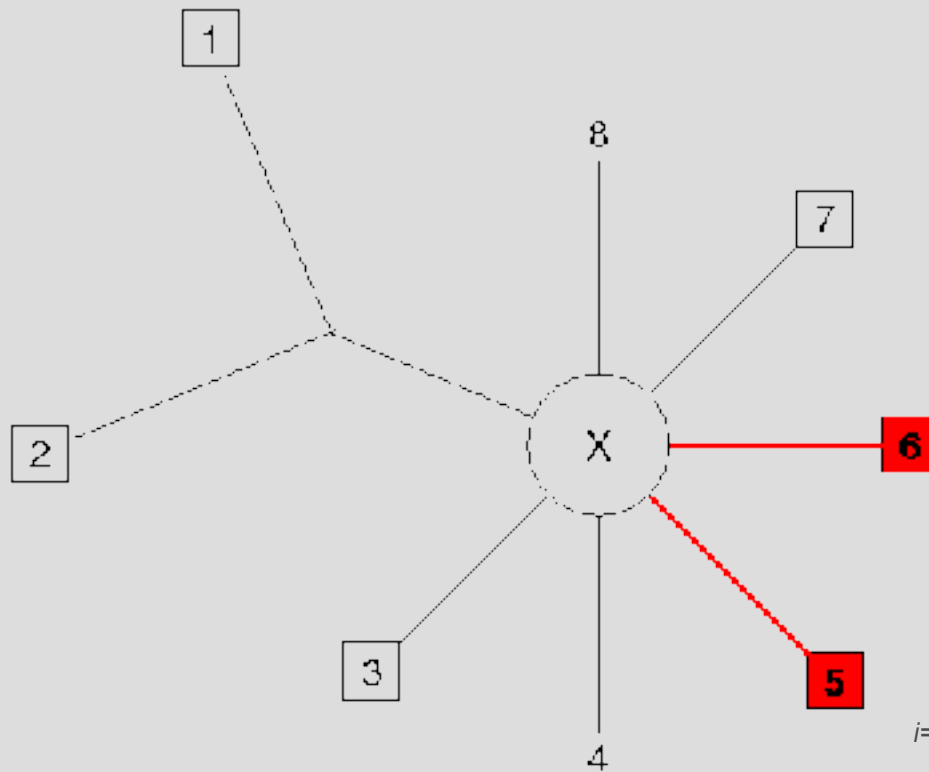


NJ: Example (SN)



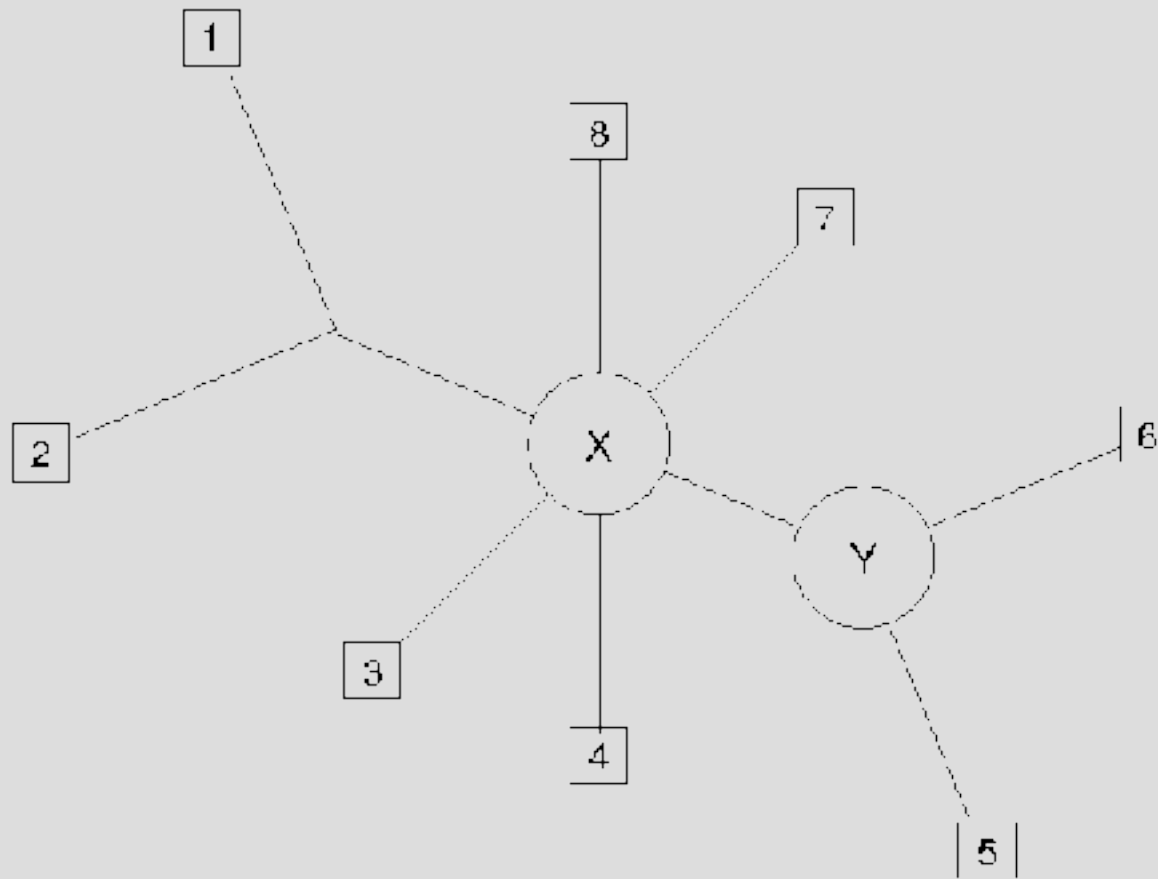
	0						
	31.50	0					
	32.30	32.30	0				
Sij=	33.90	33.90	33.70	0			
	33.90	33.90	33.70	31.30	0		
	33.70	33.70	33.50	33.10	33.10	0	
	33.70	33.70	33.50	33.10	33.10	31.90	0

NJ: Example (SN)

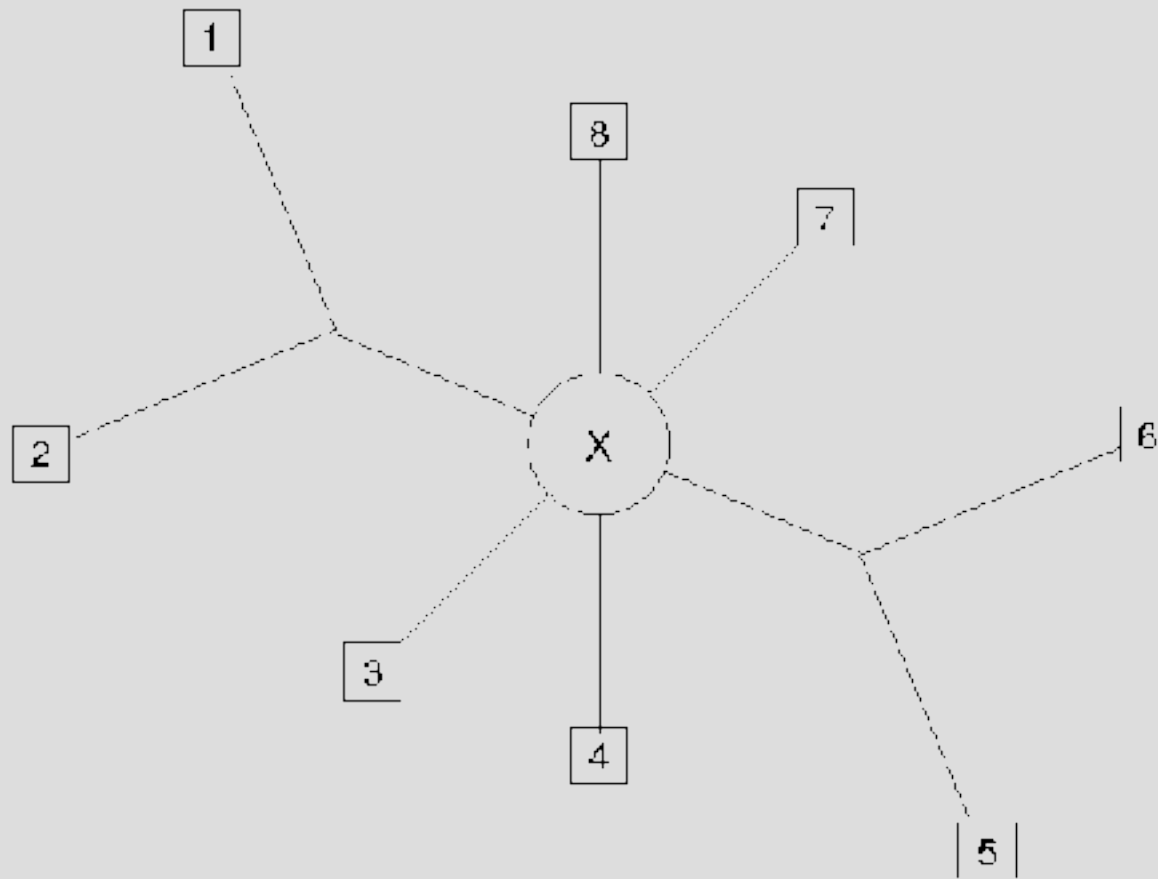


	<i>i=1_2</i>	3	4	5	6	7	8	
	0							<i>j=1_2</i>
	31.50	0						3
	32.30	32.30	0					4
<i>S_{ij}</i> =	33.90	33.90	33.70	0				5
	33.90	33.90	33.70	31.30	0			6
	33.70	33.70	33.50	33.10	33.10	0		7
	33.70	33.70	33.50	33.10	33.10	31.90	0	8

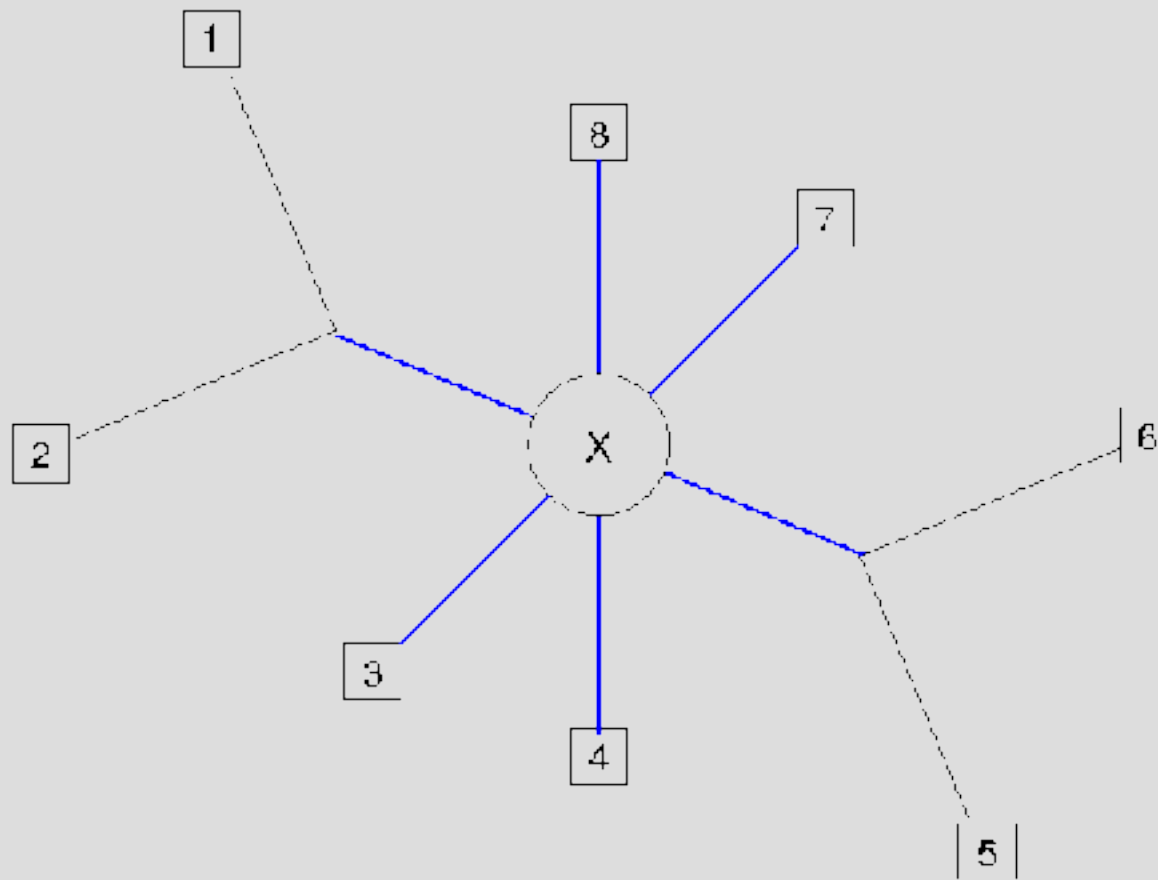
NJ: Example (SN)



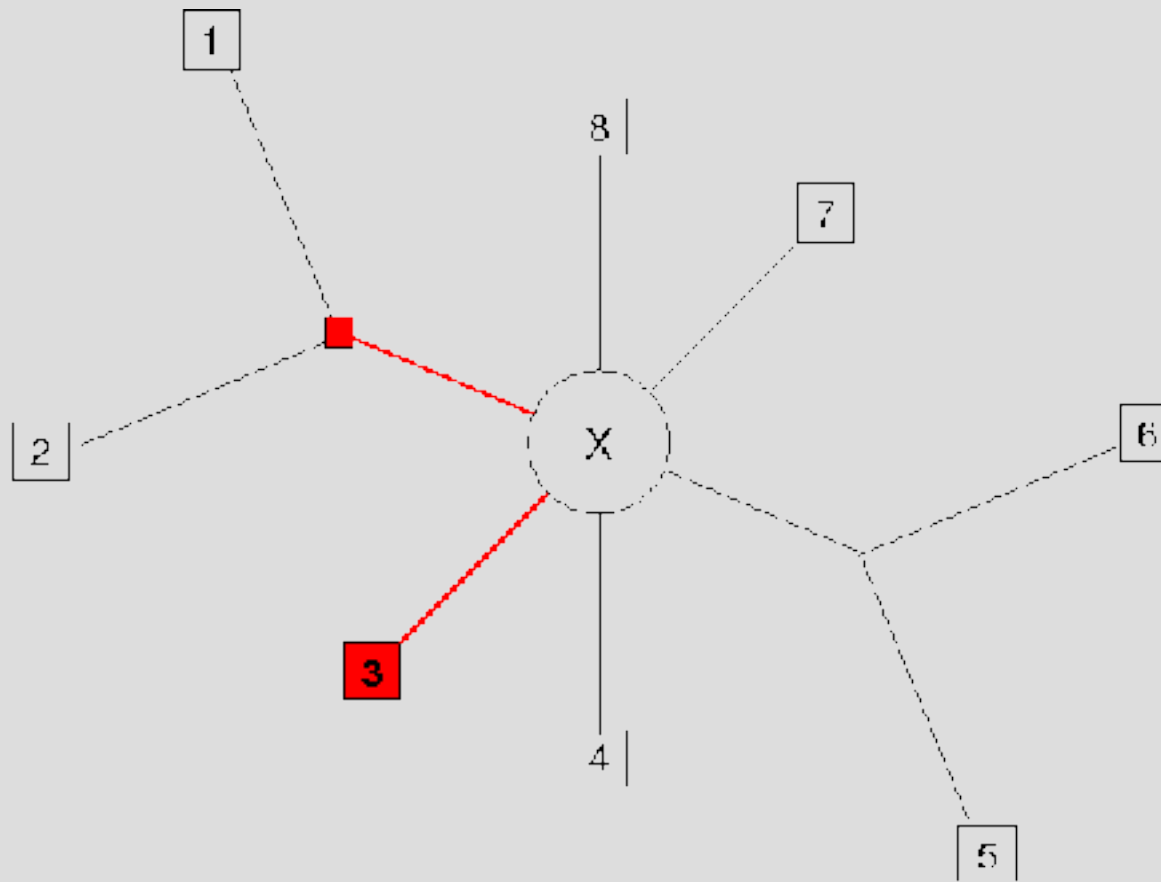
NJ: Example (SN)



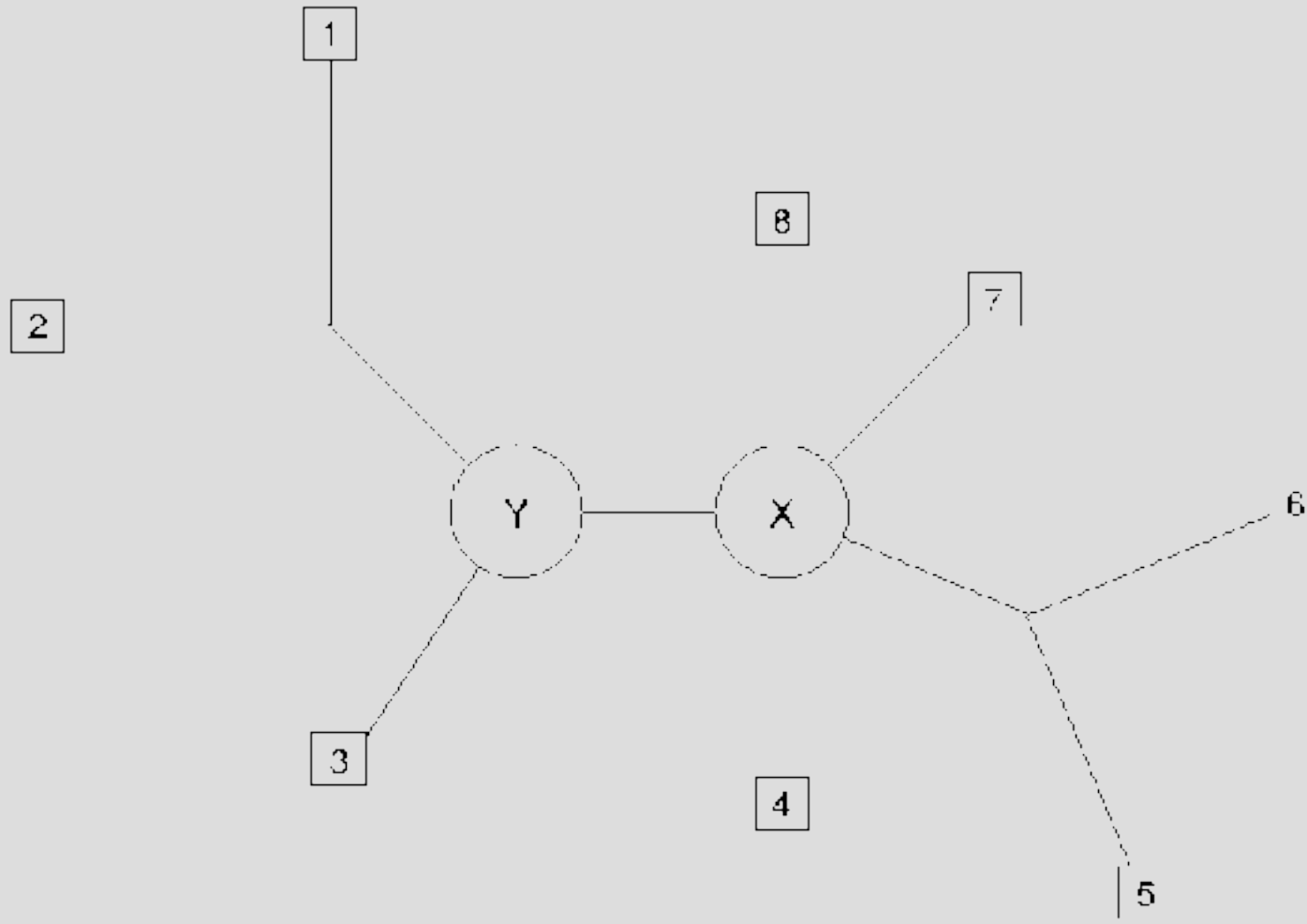
NJ: Example (SN)



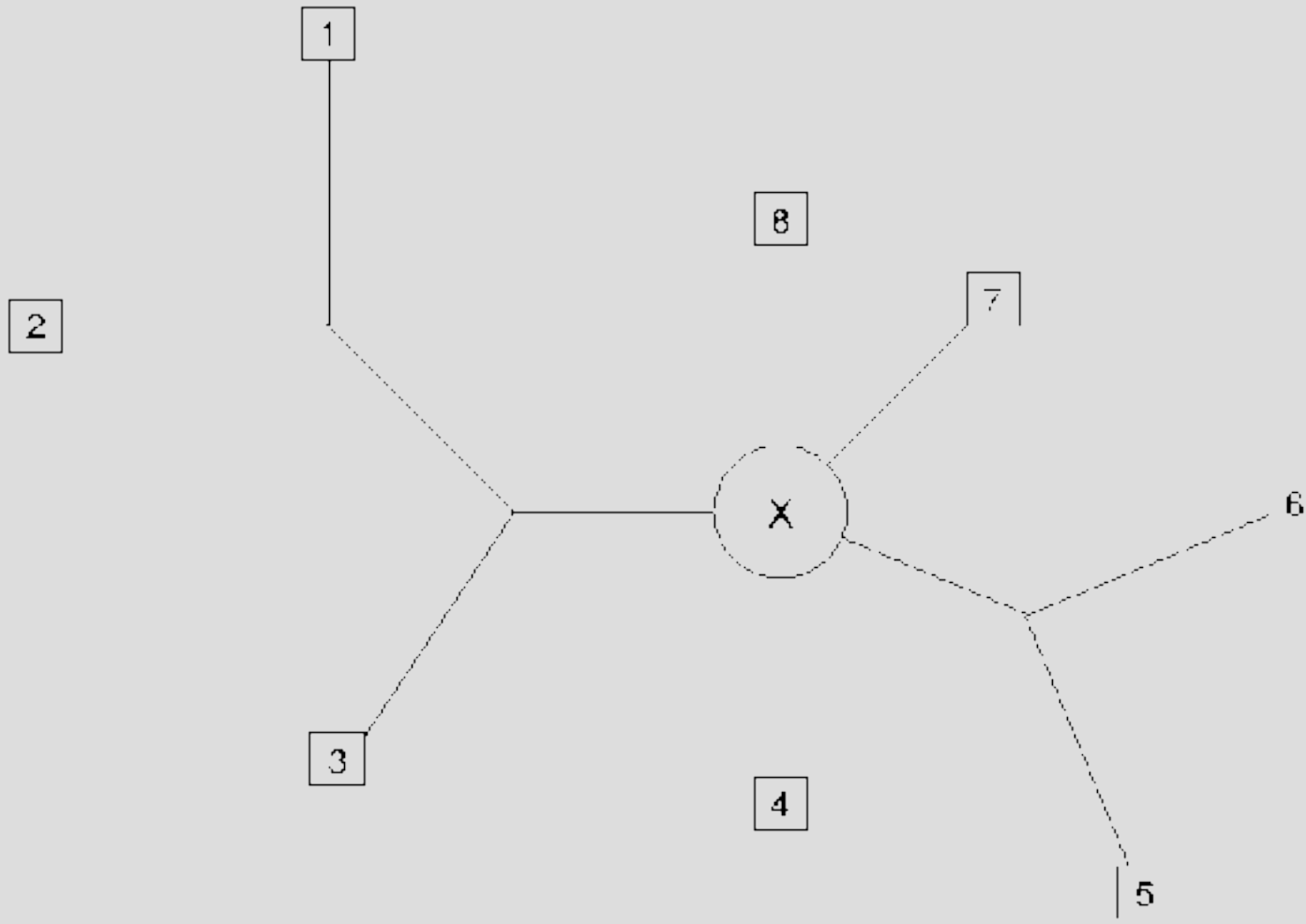
NJ: Example (SN)



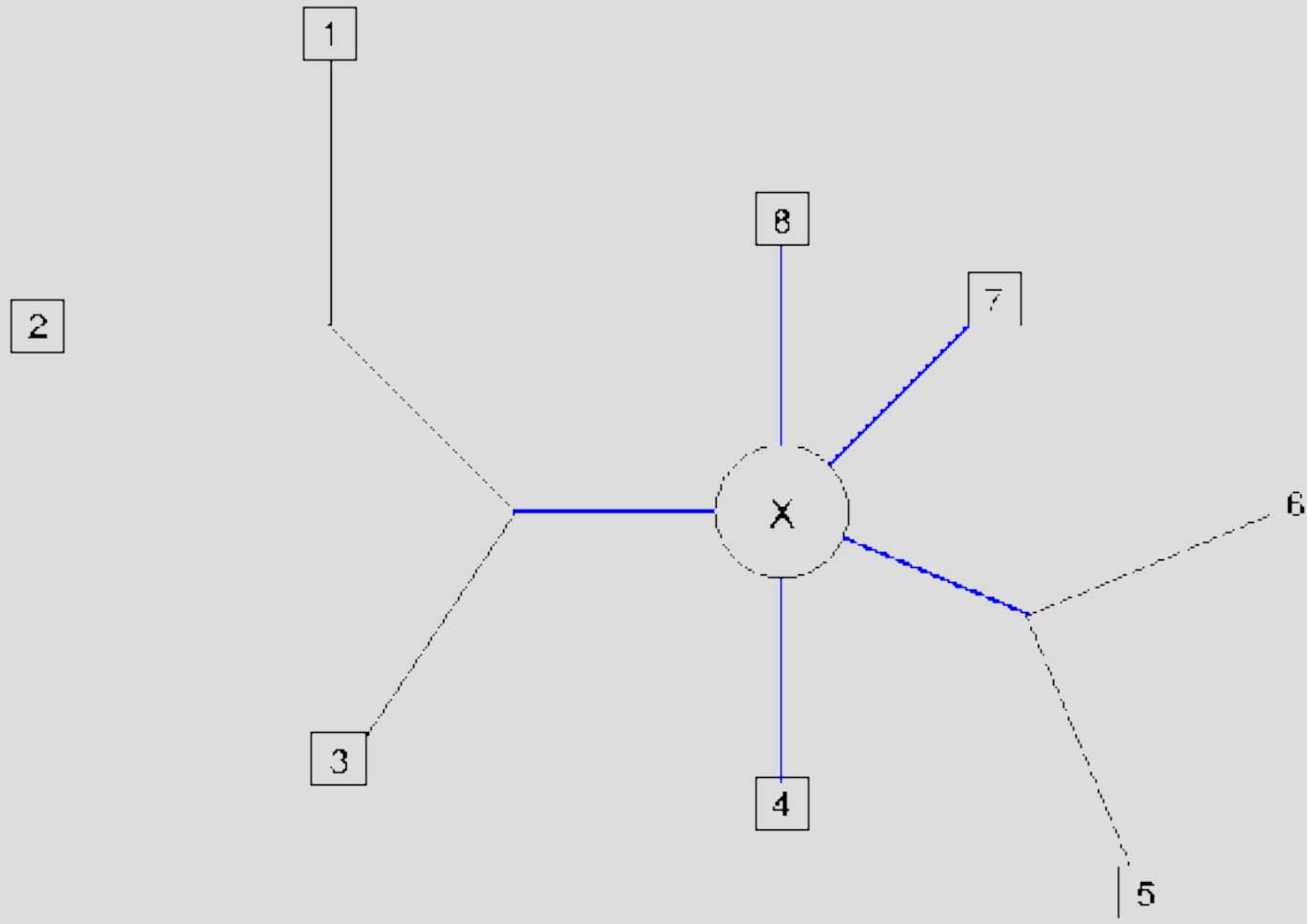
NJ: Example (SN)



NJ: Example (SN)



NJ: Example (SN)



NJ: Correctness

- Boundary conditions
 - $N > 4$ nodes
 - Purely additive tree
 - Implicates $S_{12} < S_{1i}, S_{1j}, S_{2i}, S_{2j}$
- Proof
 - 1,2 are OTU if S_{12} is min of all S_{ij}
 - Algorithm works for all purely additive trees *qu.e.d.*
- Threshold
 - 4ϵ

NJ: Conclusion

- Proof
 - *qu.e.d.* for all purely additive trees
- Complexity
 - Cubic
- Parametrisation
 - Distance metric, similitude matrix
 - Saitou & Nei, Sattath & Tversky, other
- Implementation
 - NeighborNet
 - <http://www.mcgill.ca/bryant/>

NJ: Readings

- Original paper by
Saitou M. & Nei M., Mol. Biol. Evol. 4(4):406-425, 1987
- Another fine paper by
Huson D. & Steel M., Systematic Biology, 2004 (in press)
- Performance study & proof of correctness by
Atteson K., Algorithmica (1999) 25:251-278
- NeighborNet introduced by
Bryant D. & Moulton V., Mol. Biol. Evol., 2004 (in press)