

# BMI/CS 576 – Day 15

- Today
  - Neighbor-joining and UPGMA practice
- Thursday
  - Parsimony methods

# Reminder: Midterm

- When: Wednesday, Oct 30<sup>th</sup>, 5:30-7pm
- Where: 410 Wendt commons (this space)
- What:
  - Sequence Assembly and Sequence Alignment modules
  - Paper exam (no programming)
  - Working with and reasoning about the tasks and their associated algorithms
  - No calculator/electronic devices allowed or needed
  - Two sheets of notes allowed (handwritten recommended!)
  - Old exams available on Canvas

# UW-Madison Multi-factor authentication

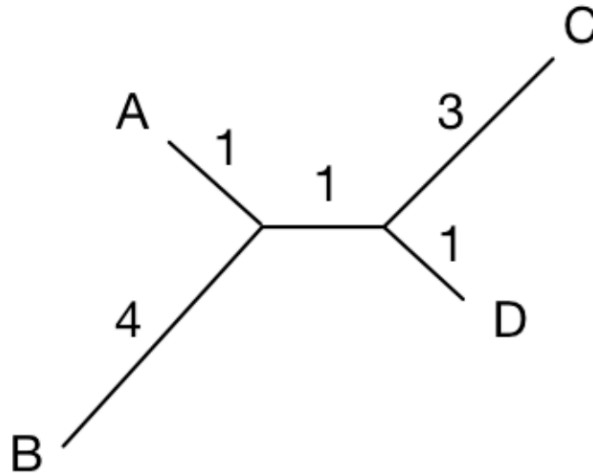
- Use secondary device (e.g., phone) to prove identity when logging in with NetID
- Required for students after Oct 31
- Enroll as soon as possible to avoid problems
- Won't be able to log in to any wisc services (including Canvas) without it!
- <https://it.wisc.edu/learn/guides/getting-started-multi-factor-authentication-students/>

# Tip: quiz for Thursday

- Only on Parsimony (first lecture)

# Quiz

Suppose you are given a pairwise distance matrix that is perfectly compatible with the unrooted tree given below. On the first iteration of the neighbor joining algorithm, which pair of leaf nodes given below could possibly be joined?

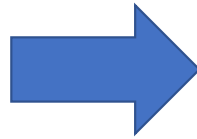
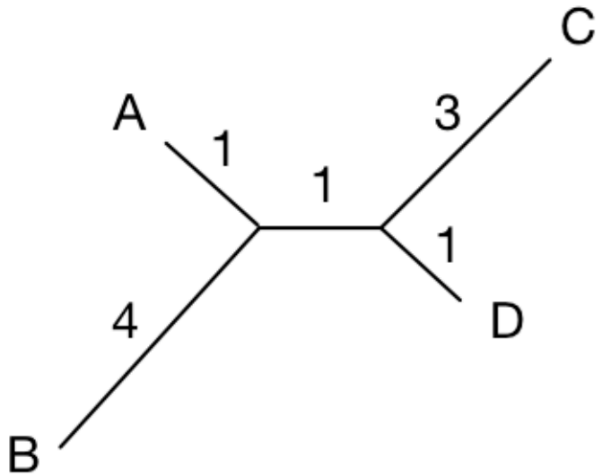


Neighbor joining is guaranteed to construct the tree if the distances are additive.

If the true tree is constructed only **A&B** or C&D can be joined on the first iteration.

Note: **UPGMA** would (incorrectly ) join A&D first because they are closest

# True tree -> Distance matrix

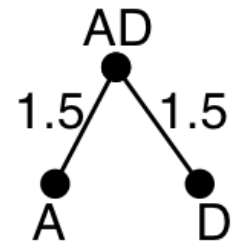


<b>d</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
<b>A</b>		5	5	3
<b>B</b>			8	6
<b>C</b>				4
<b>D</b>				

# UPGMA example

Step 1

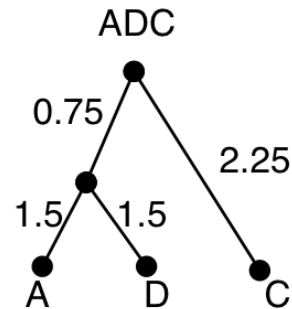
<b>d</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
<b>A</b>		5	5	3
<b>B</b>			8	6
<b>C</b>				4
<b>D</b>				



# UPGMA example

Step 2

<b>d</b>	<b>AD</b>	<b>B</b>	<b>C</b>
<b>AD</b>		5.5	4.5
<b>B</b>			8
<b>C</b>			



$$d(\text{ADC}, B) = (5.5 \times 2 + 8 \times 1) / (2 + 1) = 19 / 3 = 6.3333\dots$$

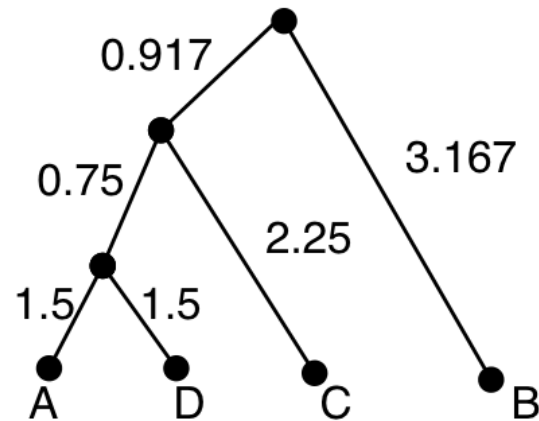
$$d_{kl} = \frac{d_{il}|C_i| + d_{jl}|C_j|}{|C_i| + |C_j|}$$



# UPGMA example

Step 3

<b>d</b>	ADC	<b>B</b>
ADC		6.33
<b>B</b>		



# NJ example

Step 1

$$D_{ij} = d_{ij} - (r_i + r_j)$$

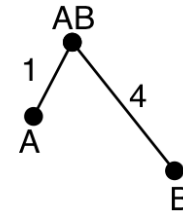
$$r_i = \frac{1}{L-2} \sum_{1 \leq k \leq L} d_{ik}$$

d	A	B	C	D
A		5	5	3
B			8	6
C				4
D				

D	A	B	C	D
A		-11	-10	-10
B			-10	-10
C				-11
D				

r
13/2
19/2
17/2
13/2

$$d_{ik} = \frac{1}{2}(d_{ij} + r_i - r_j)$$



(Could also have chosen to join C and D here)

# NJ example

Step 2

$$D_{ij} = d_{ij} - (r_i + r_j)$$

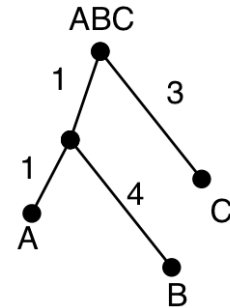
$$r_i = \frac{1}{L-2} \sum_{1 \leq k \leq L} d_{ik}$$

d	AB	C	D
AB		4	2
C			4
D			

D	AB	C	D
AB		-10	-10
C			-10
D			

r
6
8
6

$$d_{km} = \frac{d_{im} + d_{jm} - d_{ij}}{2}$$



(Three possible joins here)

# NJ example

Step 3

<b>d</b>	ABC	<b>D</b>
ABC		1
<b>D</b>		

$$d_{km} = \frac{d_{im} + d_{jm} - d_{ij}}{2}$$

