

Βιοπληροφορική

Ενότητα 11: Κατασκευή φυλογενετικών δέντρων – part II

Αν. καθηγήτης Αγγελίδης Παντελής

e-mail: paggelidis@uowm.gr

ΕΕΔΙΠ Μπέλλου Σοφία

e-mail: sbellou@uowm.gr

Τμήμα Μηχανικών Πληροφορικής και Τηλεπικοινωνιών



Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ψηφιακά Μαθήματα στο Πανεπιστήμιο Δυτικής Μακεδονίας**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



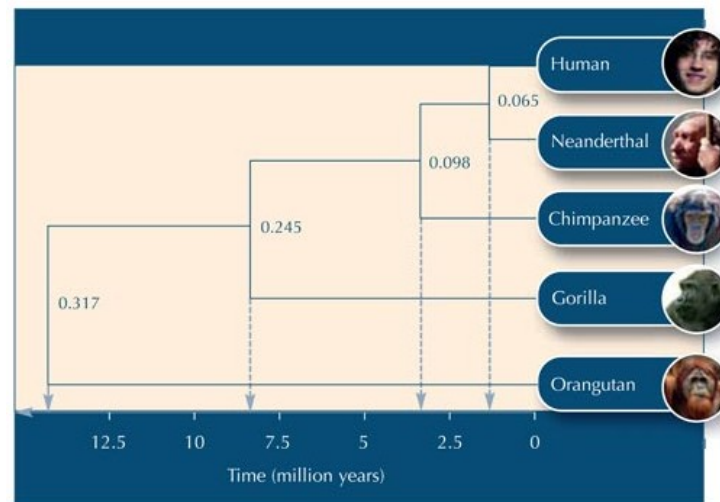
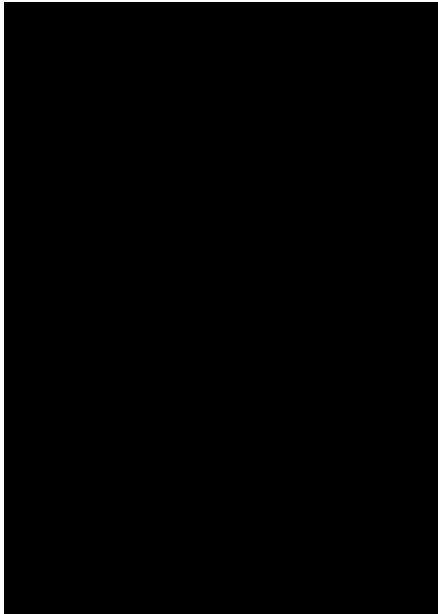
Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Κατασκευή φυλογενετικών δέντρων



Μέθοδοι κατασκευής φυλογενετικών δέντρων

- **Μέθοδοι βασισμένες σε χαρακτήρες – εξελεγκτικά μοντέλα:**
 - Μέθοδος μέγιστης συντήρησης (Maximum parsimony method).
- **Μέθοδοι βασισμένες στην απόσταση:**
 - UPGMA (Unweighted pair-group method).
 - Fitch and Margoliash Algorithm.
 - Neighbor-joining method.



Branches

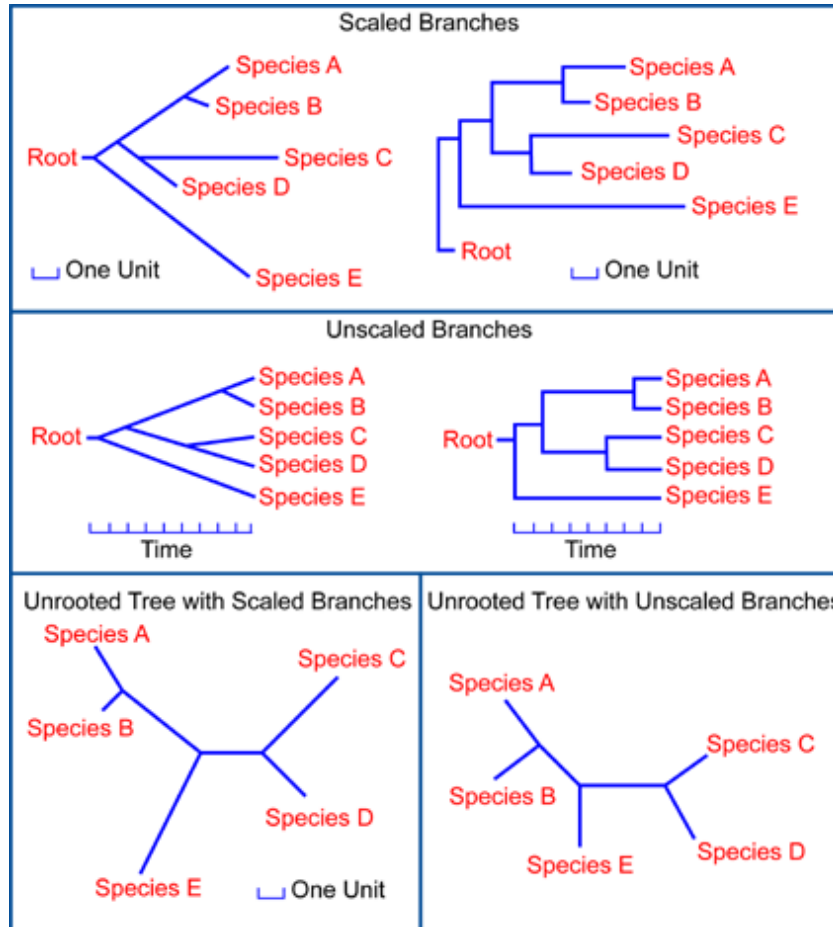


Image: <http://www.ncbi.nlm.nih.gov/About/primer/phylo.html>



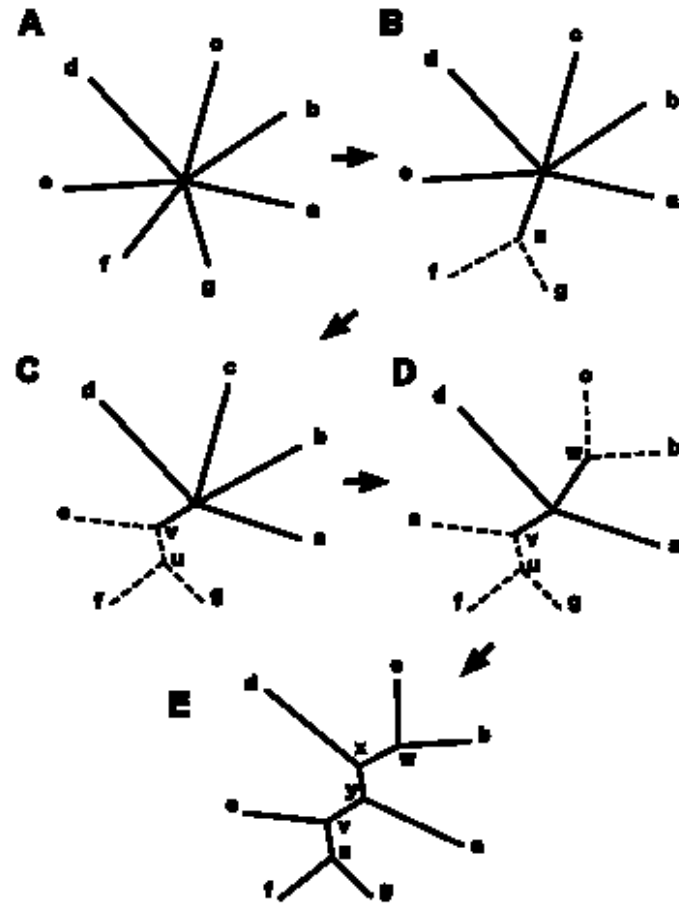
Neighbor-joining method (1/3)

- Produces unrooted tree.
- Similar to Fitch-Margoliash method.
- The choice as to which sequence to pair is determined by a different algorithm.
- Especially suitable when the rate of evolution of the separate lineages under consideration varies.
- Special case of star decomposition.
- **The pairing of sequences is based on the effect of the pairing on the sum of the branch length of the tree.**



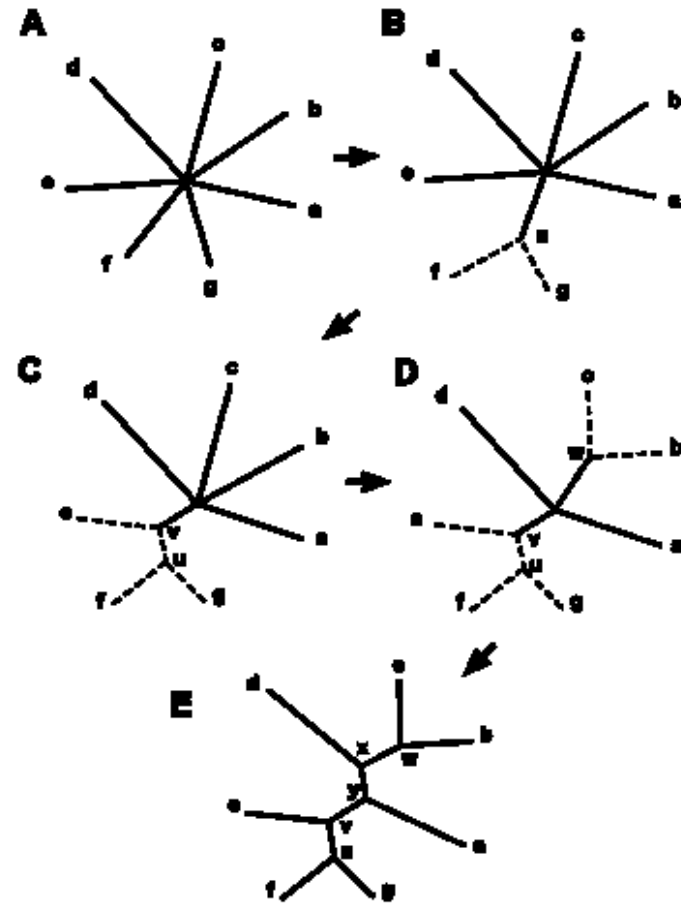
Neighbor-joining method (2/3)

1. Το αρχικό δέντρο είναι ένα δέντρο-αστέρι.
2. Υπολογίζεται η συνολική απόσταση κάθε είδους από όλα τα υπόλοιπα. Δημιουργείται ένας νέος πίνακας.
3. Υπολογίζεται η απόσταση κάθε ζεύγους από όλες τις υπόλοιπες οντότητες.
4. Επιλέγεται το ζεύγος με τη μεγαλύτερη απόσταση από τις υπόλοιπες οντότητες και ορίζεται ο κόμβος μεταξύ τους.



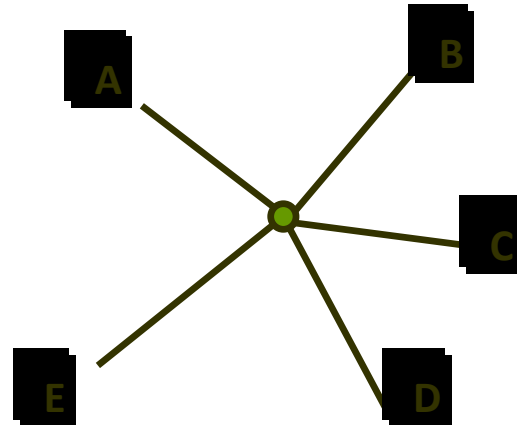
Neighbor-joining method (3/3)

5. Κατασκευάζεται καινούργιος πίνακας χρησιμοποιώντας τον κόμβο μεταξύ των δύο οντοτήτων.
6. Υπολογίζουμε αποστάσεις από τον κόμβο.
7. Επαναλαμβάνουμε τη διαδικασία μέχρι να φτάσουμε στην τελευταία οντότητα.



Neighbor-joining method – Example (1/13)

	A	B	C	D	E
A		5	4	9	8
B			5	10	9
C				7	6
D					7
E					

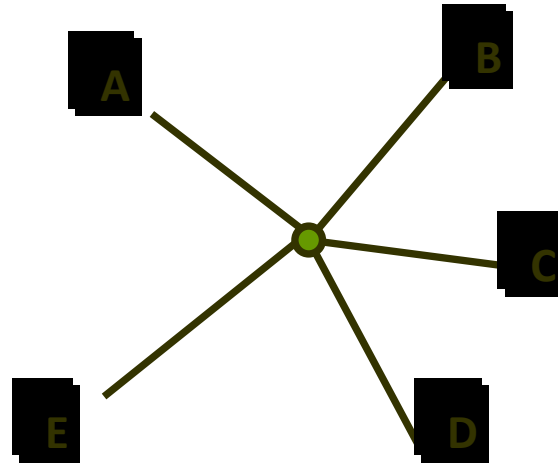


- Υπολογίζουμε τις αποστάσεις κάθε αλληλουχίας με όλες τις υπόλοιπες
A→BCDE: $5+4+9+8=26$
B→ACDE: $5+5+10+9=29$
C→ABDE: $4+5+7+6=22$
D→ABCE: $9+10+7+7=33$
E→ABCD: $8+9+6+7=30$



Neighbor-joining method – Example (2/13)

	A	B	C	D	E
A		5	4	9	8
B			5	10	9
C				7	6
D					7
E					



Υπολογίζουμε τις αποστάσεις κάθε ζεύγους από τις υπόλοιπες αλληλουχίες

$M_{nm} = (N-2)d_{nm} - R_n - R_m$, όπου

d_{nm} : απόσταση μεταξύ n, m

R_n : απόσταση n από όλες τις υπόλοιπες

R_m : απόσταση m από όλες τις υπόλοιπες



Neighbor-joining method – Example (3/13)

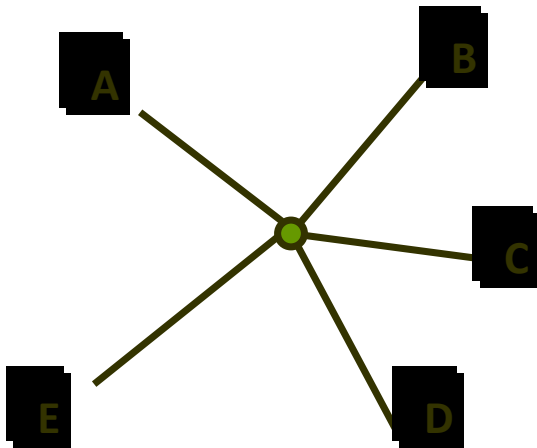
Υπολογίζουμε τις αποστάσεις κάθε ζεύγους από τις υπόλοιπες αλληλουχίες

$$M_{nm} = (N-2)d_{nm} - R_n - R_m, \text{ όπου}$$

d_{nm} : απόσταση μεταξύ n, m

R_n : απόσταση n από όλες τις υπόλοιπες

R_m : απόσταση m από όλες τις υπόλοιπες



Απόσταση του AB από το C: $AB - AC - BC$

Απόσταση του AB από το D: $AB - AD - BD$

Απόσταση του AB από το E: $AB - AE - BE$

$$M_{nm} = 3AB - (AC + AD + AE) - (BC + BD + BE)$$

$$M_{nm} = (N-2)d_{AB} - R_A - R_B$$



Neighbor-joining method – Example (4/13)

	A	B	C	D	E
A		5	4	9	8
B			5	10	9
C				7	6
D					7
E					

$$M_{nm} = (N-2)d_{nm} - R_n - R_m$$

$$R_A: A \rightarrow BCDE: 5+4+9+8=26$$

$$R_B: B \rightarrow ACDE: 5+5+10+9=29$$

$$R_C: C \rightarrow ABDE: 4+5+7+6=22$$

$$R_D: D \rightarrow ABCE: 9+10+7+7=33$$

$$R_E: E \rightarrow ABCD: 8+9+6+7=30$$

$$M_{AB} = 3 \cdot 5 - 26 - 29 = -40$$

$$M_{AC} = 3 \cdot 4 - 26 - 22 = -36$$

$$M_{AD} = 3 \cdot 9 - 26 - 33 = -32$$

$$M_{AE} = 3 \cdot 8 - 26 - 30 = -32$$

$$M_{BC} = 3 \cdot 5 - 29 - 22 = -36$$

$$M_{BD} = 3 \cdot 10 - 29 - 33 = -32$$

$$M_{BE} = 3 \cdot 9 - 29 - 30 = -32$$

$$M_{CD} = 3 \cdot 7 - 22 - 33 = -34$$

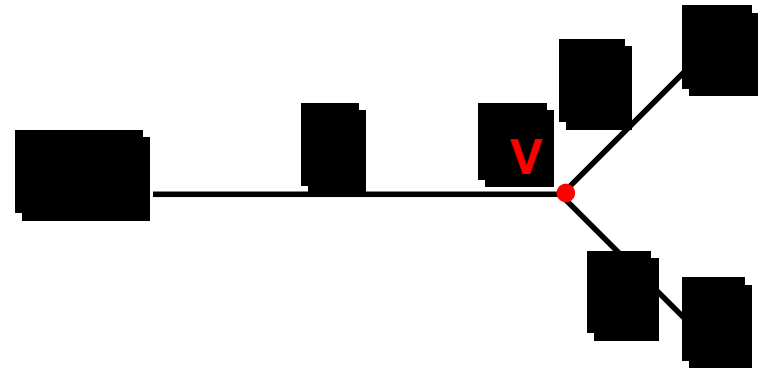
$$M_{CE} = 3 \cdot 6 - 22 - 30 = -34$$

$$M_{DE} = 3 \cdot 7 - 33 - 30 = -34$$



Neighbor-joining method – Example (5/13)

	A	B	C	D	E
A		5	4	9	8
B			5	10	9
C				7	6
D					7
E					



$$d + e = 7$$

$$D \rightarrow ABC: d+z = (DA+DB+DC)/3=8.67$$

$$E \rightarrow ABC: e+z = (EA+EB+EC)/3=7.67$$

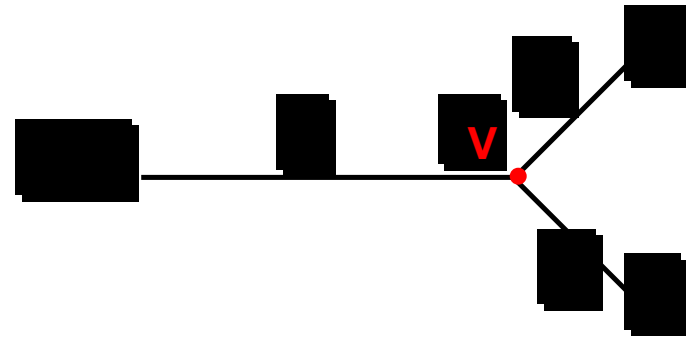
Τελικά: $d = 4, e = 3, z=4,67$

Δημιουργείται ο κόμβος V μεταξύ D & E



Neighbor-joining method – Example (6/13)

	A	B	C	D	E
A		5	4	9	8
B			5	10	9
C				7	6
D					7
E					



Υπολογίζεται η απόσταση κάθε αλληλουχίας εκτός κόμβου, από τον κόμβο, δηλ, η απόσταση των A, B, C από τον κόμβο V

$$A \rightarrow V: (AD+AE-DE)/2 = 5$$

$$B \rightarrow V: (BD+BE-DE)/2 = 6$$

$$C \rightarrow V: (CD+CE-DE)/2 = 3$$

ΠΡΟΣΟΧΗ: Η μέση απόσταση από τον κόμβο είναι $[(5+6+3)/3 = 4,67]$, ίσο με το z



Neighbor-joining method – Example (7/13)

- Κατασκευάζουμε τον καινούργιο πίνακα αποστάσεων τοποθετώντας τον κόμβο V στη θέση των D & E.

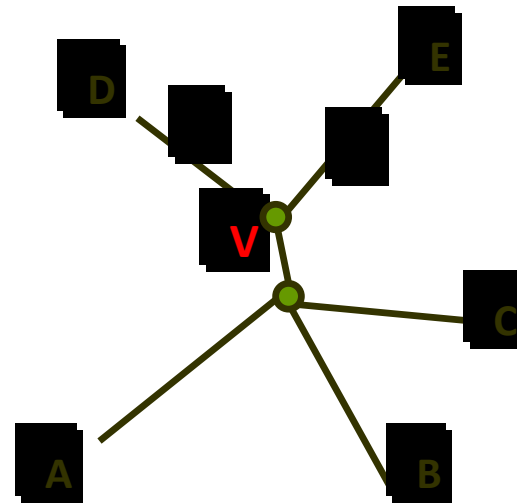
	A	B	C	D	E
A		5	4	9	8
B			5	10	9
C				7	6
D					7
E					



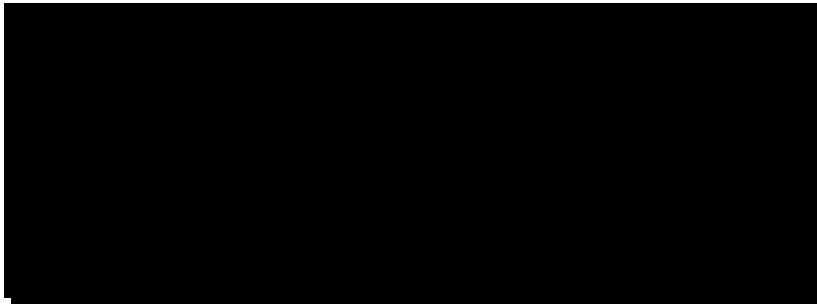
	A	B	C	V
A		5	4	5
B			5	6
C				3
V				

Neighbor-joining method – Example (8/13)

	A	B	C	V
A		5	4	5
B			5	6
C				3
V				



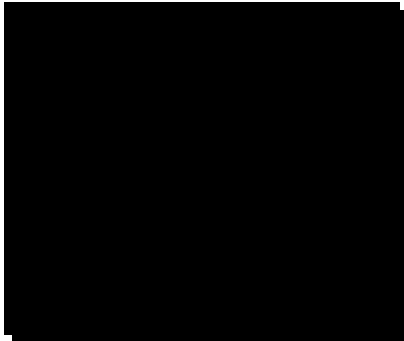
Επαναλαμβάνω την ίδια διαδικασία:



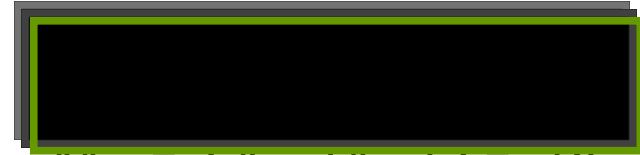
Neighbor-joining method – Example (9/13)

	A	B	C	V
A		5	4	5
B			5	6
C				3
V				

Επαναλαμβάνω την
ίδια διαδικασία:



$$M_{nm} = (N-2)d_{nm} - R_n - R_m$$



$$M_{AC} = 2 \cdot 4 - 14 - 12 = -18$$

$$M_{AV} = 2 \cdot 5 - 14 - 14 = -18$$

$$M_{BC} = 2 \cdot 5 - 16 - 12 = -18$$

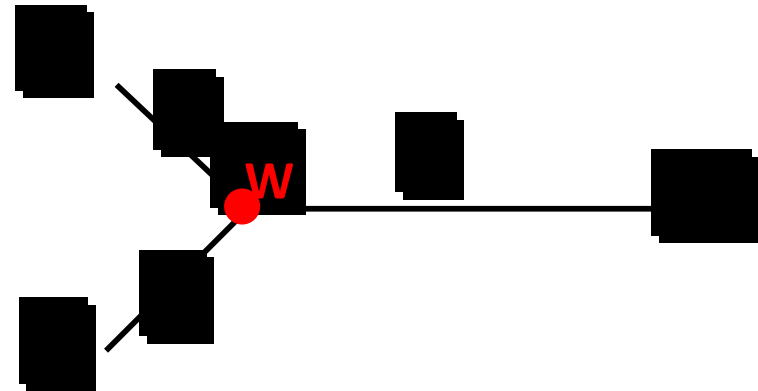
$$M_{BV} = 2 \cdot 6 - 16 - 14 = -18$$

$$M_{CV} = 2 \cdot 3 - 12 - 14 = -20$$



Neighbor-joining method – Example (10/13)

	A	B	C	V
A		5	4	5
B			5	6
C				3
V				



Υπολογίζω τις αποστάσεις a & b :

$$a + b = 5$$

$$a + \gamma = (AC + AV) / 2 = 4.5$$

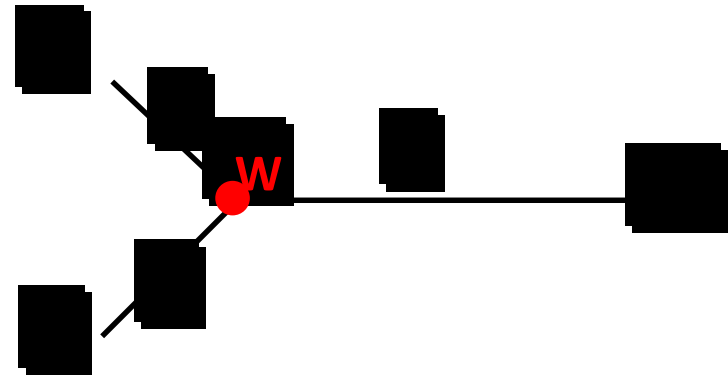
$$b + \gamma = (BC + BV) / 2 = 5.5$$

Τελικά: $a = 2$, $b = 3$



Neighbor-joining method – Example (11/13)

	A	B	C	V
A		5	4	5
B			5	6
C				3
V				



Υπολογίζεται η απόσταση κάθε αλληλουχίας εκτός κόμβου, από τον κόμβο , δηλ, η απόσταση των C, V από τον κόμβο W

$$C \rightarrow W: (CA+CB-AB)/2 = 2$$

$$V \rightarrow W: (VA+VB-AB)/2 = 3$$

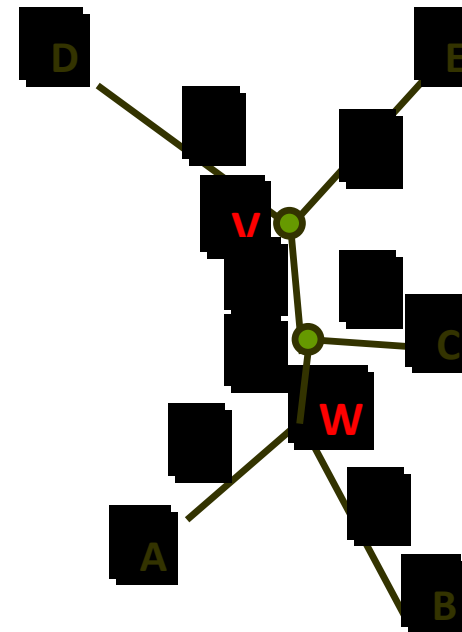
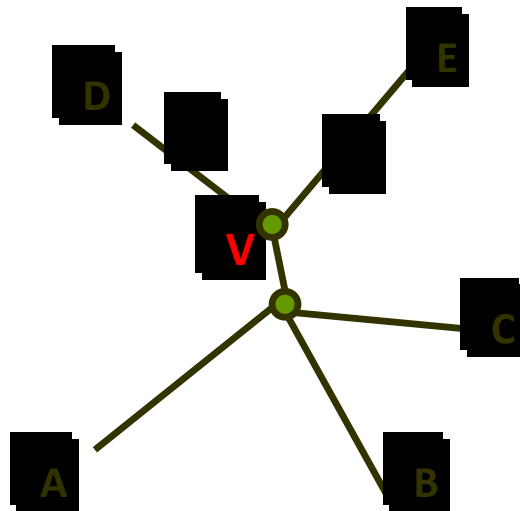


Neighbor-joining method – Example (12/13)

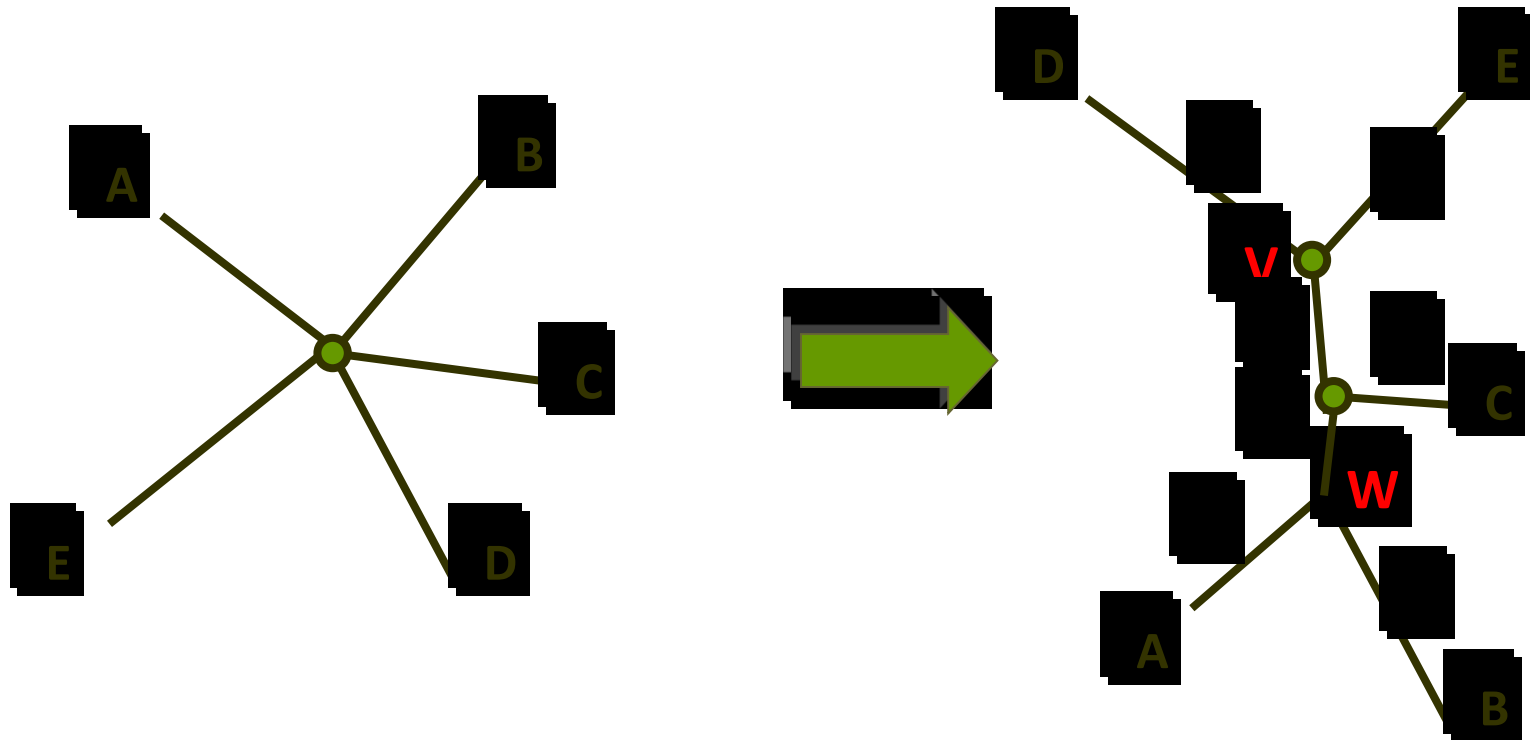
	A	B	C	V
A		5	4	5
B			5	6
C				3
V				



	W	C	V
W		2	3
C			3
V			



Neighbor-joining method – Example (13/13)



Τέλος Ενότητας



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



Σημείωμα Αναφοράς

- Copyright Πανεπιστήμιο Δυτικής Μακεδονίας, Τμήμα Μηχανικών Πληροφορικής και Τηλεπικοινωνιών, Αγγελίδης Παντελής. «**Βιοπληροφορική**». Έκδοση: 1.0. Κοζάνη 2015. Διαθέσιμο από τη δικτυακή διεύθυνση: <https://eclass.uowm.gr/courses/ICTE102/>



Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Όχι Παράγωγα Έργα Μη Εμπορική Χρήση 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Ως Μη Εμπορική ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό

Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους
υπερσυνδέσμους.

