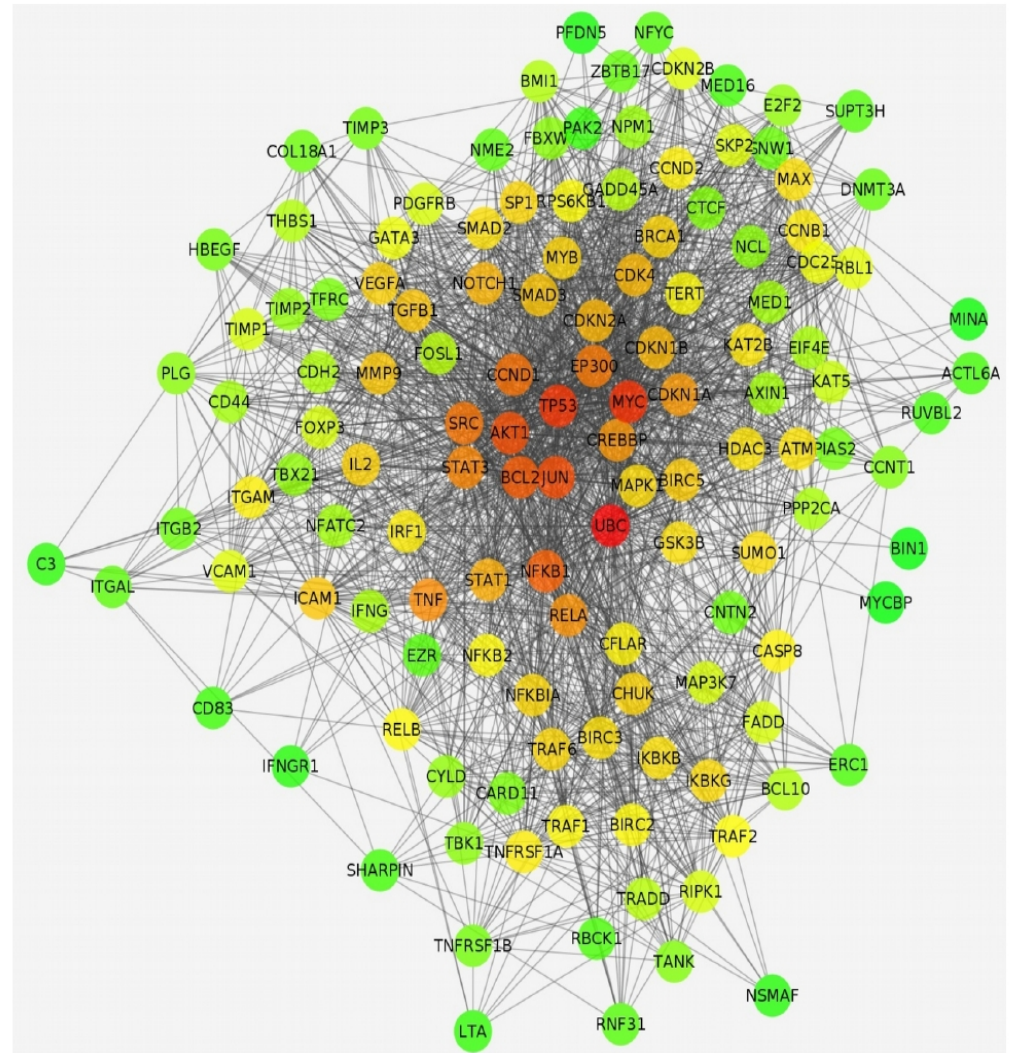


Υπολογιστική Βιολογία : BIO315
[9] Εισαγωγή στα Βιολογικά Δίκτυα

Η βιολογία στην εποχή των δικτύων

Η βιολογία των “μεγάλων δεδομένων” (big data) παρέχει τη δυνατότητα δημιουργίας συσχετίσεων σε μεγάλη κλίμακα.

Αναλύσεις μεγάλης κλίμακας σε επίπεδο ρύθμισης, συν-έκφρασης, συμμετοχής σε αντιδράσεις κλπ, οδηγούν στην ανάγκη αναπαράστασης των συσχετίσεων που παρατηρούνται.



Εικόνα 9.1: Ένα παράδειγμα δικτύου πρωτεϊνικών αλληλεπιδράσεων. Οι πρωτεΐνες αναπαρίστανται ως κόμβοι και οι μεταξύ τους σχέσεις με ακμές, που αντιστοιχούν σε σχέσεις ενεργοποίησης. Το χρώμα των κόμβων είναι ανάλογο των αριθμητικών τιμών χαρακτηριστικών ιδιοτήτων του δικτύου που θα συζητηθούν στη συνέχεια του κεφαλαίου. Η Εικόνα δημιουργήθηκε με τη χρήση του Cytoscape (Shannon et al., 2003).

Βιολογικά Δίκτυα

Δίκτυο: Μια θεωρητική, μαθηματική κατασκευή που περιγράφει τις σχέσεις μεταξύ ομοειδών ή ομάδων ομοειδών στοιχείων, αναπαριστώντας τις με σχηματικό τρόπο.

Βιολογικά στοιχεία που ικανοποιούν τον παραπάνω ορισμό:

Δίκτυα μεταγραφικής ρύθμισης: Οι σχέσεις που συνδέουν δύο στοιχεία είναι το κατά πόσο τα επίπεδα έκφρασης του ενός σχετίζονται με ή επηρεάζονται, διαμορφώνονται από το άλλο.

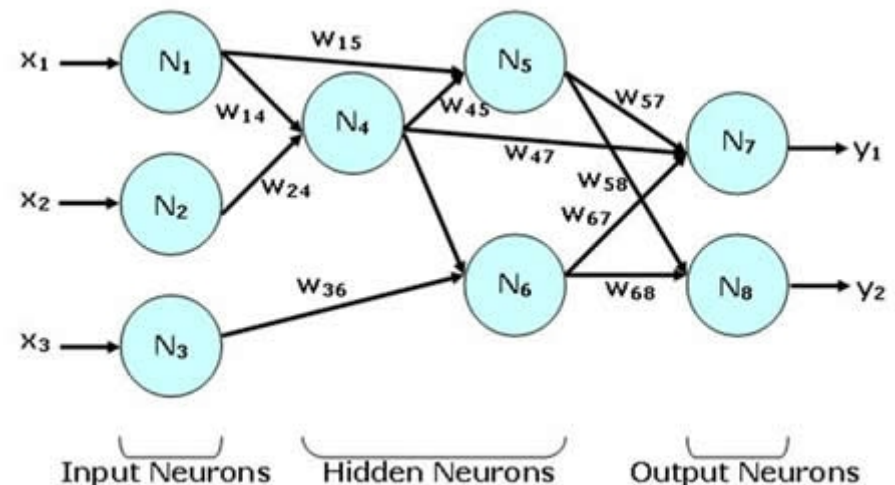
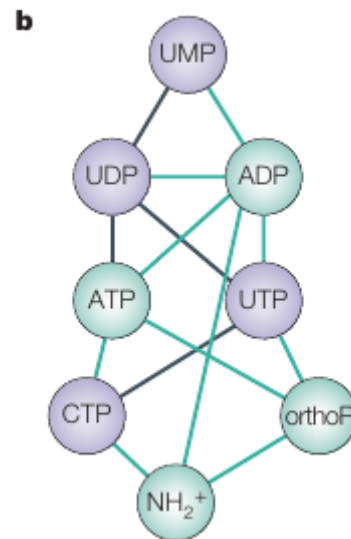
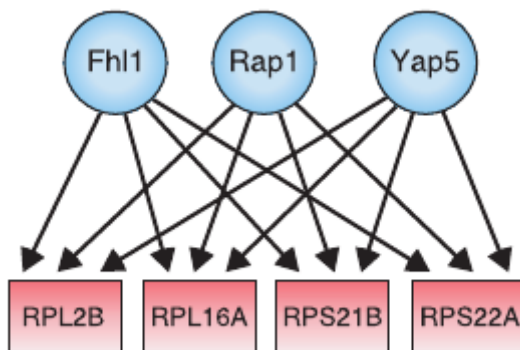
Δίκτυα Αντιδράσεων (μεταβολικά, σηματοδοτικά κλπ): Δύο στοιχεία σχετίζονται εφόσον συμμετέχουν στην ίδια αντίδραση.

Δίκτυα Φυσικών Αλληλεπιδράσεων: Δύο στοιχεία σχετίζονται όταν αποδεδειγμένα βρίσκονται σε φυσική αλληλεπίδραση (επαφή, προσκόλληση, προσρόφηση) σε κάποιο στάδιο της “ζωής” τους. (π.χ. δίκτυα νευρώνων)

Είδη δικτύων

Είναι μαθηματικές αναπαραστάσεις (graphs) στις οποίες συγκεκριμένος αριθμός στοιχείων απεικονίζεται με τη μορφή κόμβων (nodes, vertices) και οι μεταξύ τους σχέσεις με τη μορφή ακμών (edges).

Τα δίκτυα μπορεί να είναι: α) κυκλικά/μη-κυκλικά β) κατευθυνόμενα/μη-κατευθυνόμενα γ) απλά ή με φορτισμένες ακμές (βάρη)

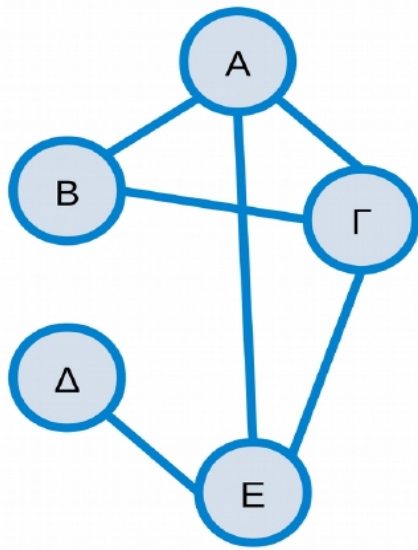


Είδη δικτύων: Ρυθμιστικά δίκτυα

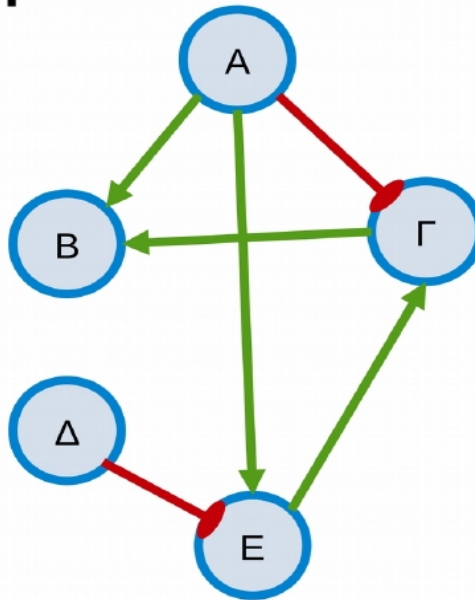
Τα στοιχεία-κόμβοι είναι γονίδια και οι σχέσεις μεταξύ τους είναι σχέσεις ρύθμισης.

Αυτές μπορεί να είναι: α) απλές (χωρίς κατεύθυνση) β) επαγωγής-καταστολής (με κατεύθυνση) γ) ποσοτικής επαγωγής-καταστολής (με κατεύθυνση ΚΑΙ βάρος)

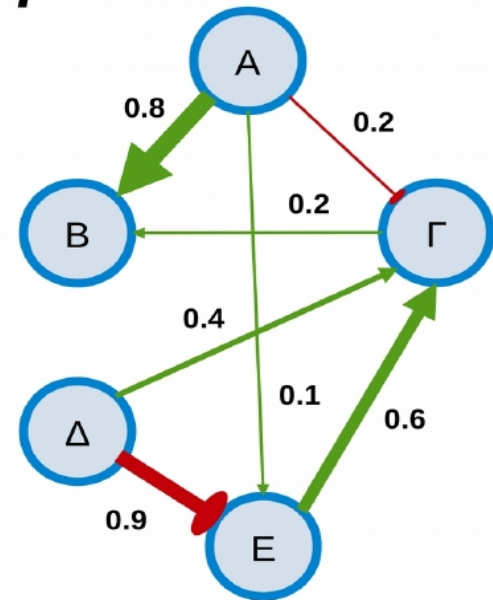
α



β



γ



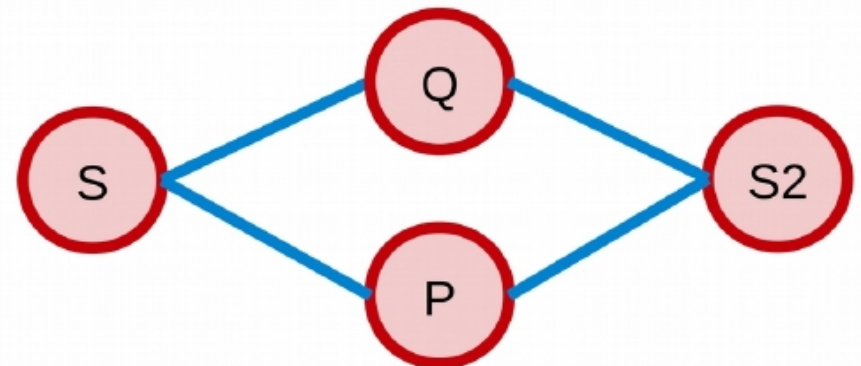
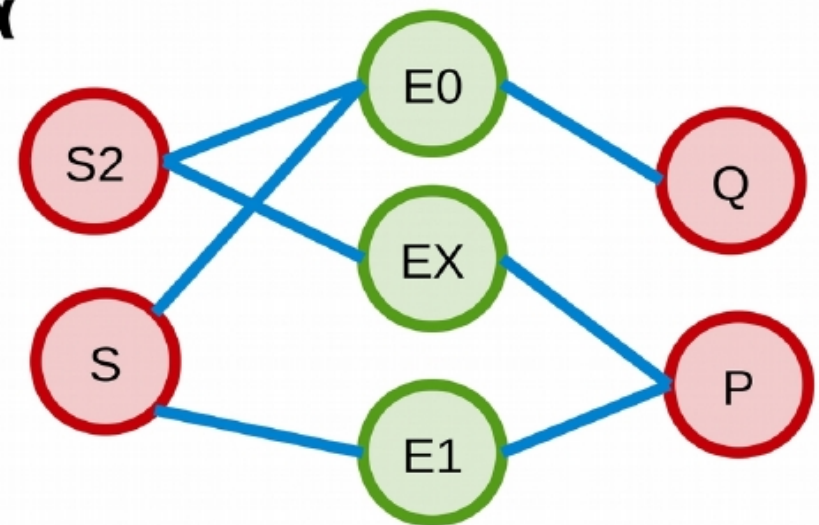
Διμερή (bipartite) Δίκτυα: Μεταβολικά Δίκτυα

Διμερή ονομάζονται τα δίκτυα που συμπεριλαμβάνουν δύο είδη οντοτήτων.

Ένα βιολογικό παράδειγμα είναι τα μεταβολικά δίκτυα που μπορεί να συμπεριλαμβάνουν τόσο τους μεταβολίτες όσο και τα ένζυμα (α).

Μια απλουστευμένη μορφή απλώς ενώνει με ακμές τους μεταβολίτες που συμμετέχουν στην ίδια αντίδραση (β)

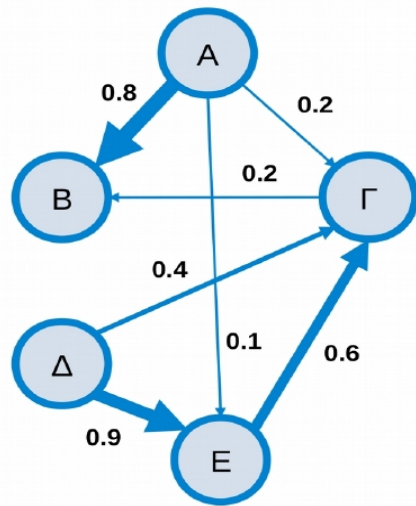
α



Αναπαραστάσεις Δικτύων

Αναπαριστούμε τα δίκτυα είτε γραφικά, είτε αλγεβρικά σαν λίστες, πίνακες αλληλεπιδράσεων (γειτνίασης) (adjacency matrices).

α



β

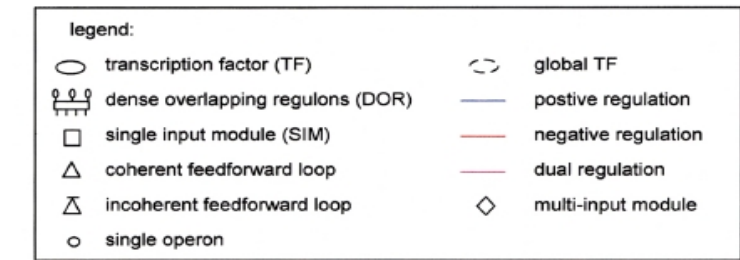
A	B	0.8
	Γ	0.2
	Ε	0.1
Γ	B	0.2
Δ	Γ	0.4
Δ	Ε	0.9
Ε	Γ	0.6

γ

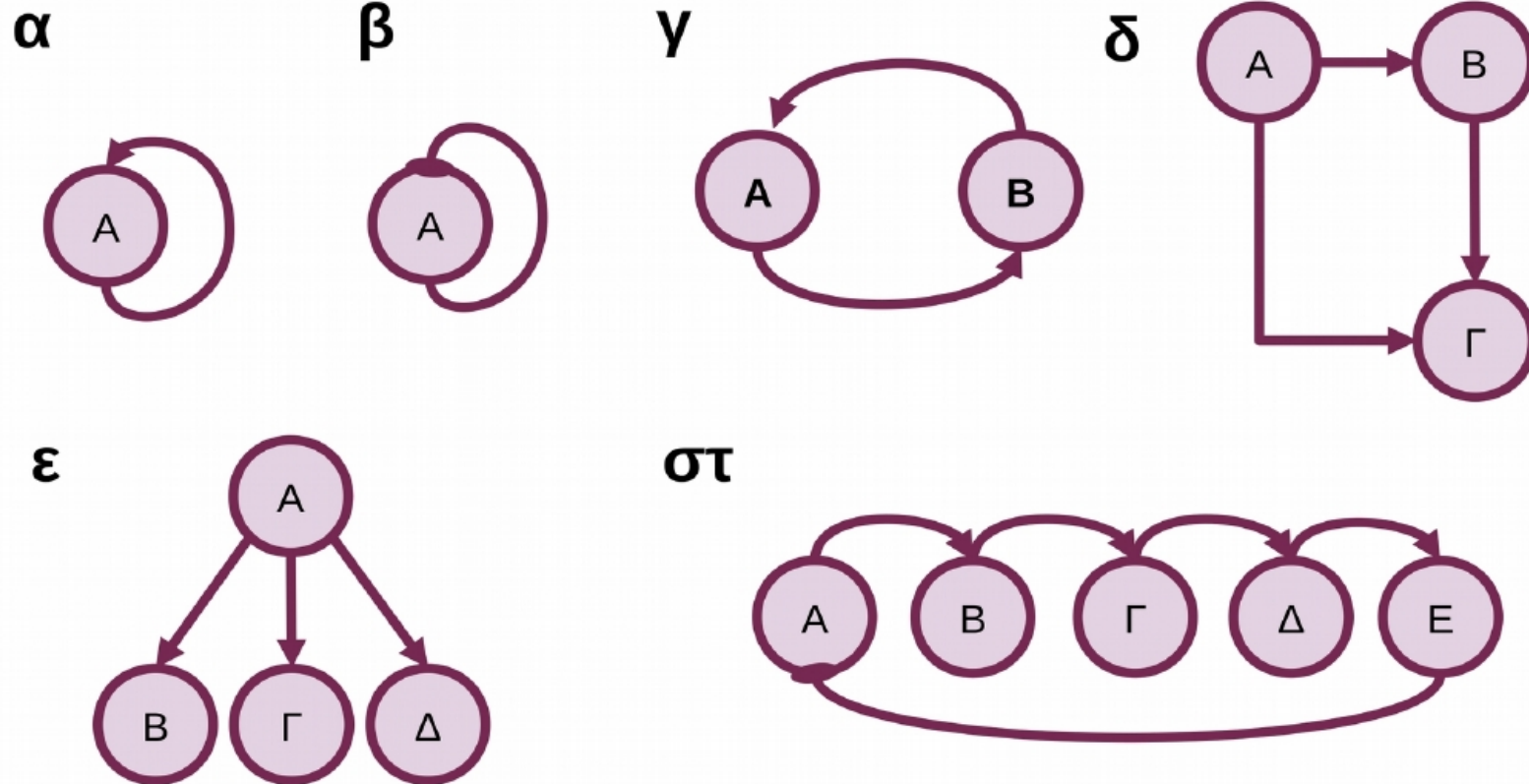
	A	B	Γ	Δ	Ε
A	0	0.8	0.2	0	0.1
B	-0.8	0	-0.2	0	0
Γ	-0.2	0.2	0	-0.4	-0.6
Δ	0	0	0.4	0	0.9
Ε	0	0	0.6	-0.9	0

- α) Απλά δίκτυα χωρίς βάρη: Αν γνωρίζουμε ότι υπάρχει μια σχέση μεταξύ A και B τότε $G[A,B] = G[B,A] = 1$, σε οποιαδήποτε άλλη περίπτωση $G[A,B] = 0$
- β) Κατευθυνόμενα δίκτυα χωρίς βάρη: Αν γνωρίζουμε ότι υπάρχει μια σχέση μεταξύ A και B τέτοια ώστε $A \rightarrow B$ τότε $G[A,B] = 1$ και $G[B,A] = -1$
- γ) Δίκτυα με βάρη: Αν γνωρίζουμε ότι υπάρχει μια σχέση μεταξύ A και B με βάρος X τότε $G[A,B] = X$

- Στην αναγνώριση συγκεκριμένων μοτίβων δικτύων
- Στην μοντελοποίηση διαδικασιών μέσω μελέτης των μοτίβων
- Στην μελέτη της οργάνωσης και της εξέλιξης σύνθετων λειτουργιών



Μοτίβα Δικτύων Μεταγραφικής Ρύθμισης



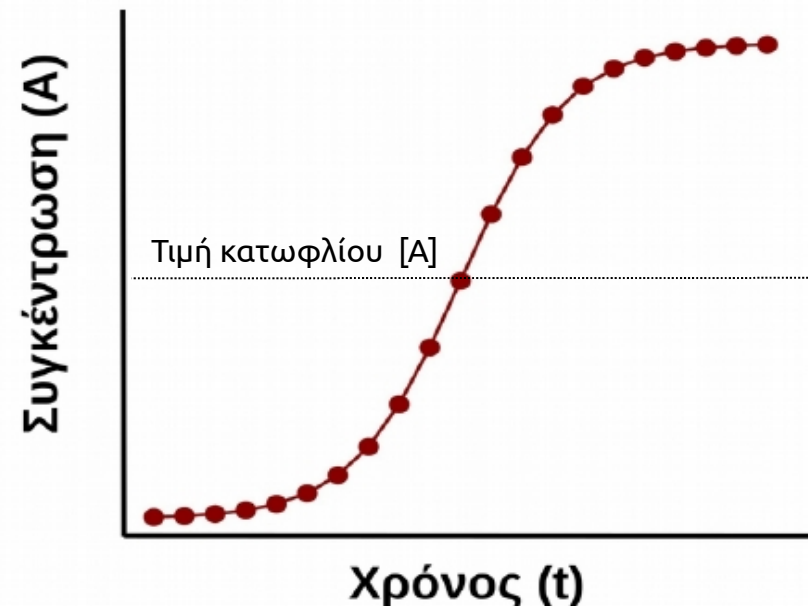
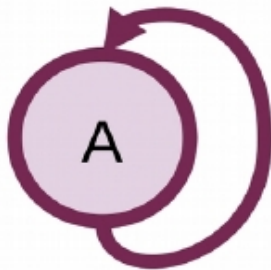
Ονομάζουμε “μοτίβο” δικτύου μια επαναλαμβανόμενη δομή του δικτύου που φαίνεται να εμφανίζεται συχνότερα από το αναμενόμενο. Σε ρυθμιστικά (κυρίως) δίκτυα τα μοτίβα έχουν συχνά ξεχωριστές ιδιότητες.

Αυτορύθμιση I: Επαγωγή

Το A επάγει τον εαυτό του. Η μεταβολή της συγκέντρωσης του A σε σχέση με το χρόνο περιγράφεται από μια σιγμοειδή καμπύλη.

Βιολογικό αποτέλεσμα: Απότομες μεταβολές συγκέντρωσης του A ώστε να επιτυγχάνεται μια on-off κατάσταση.

α

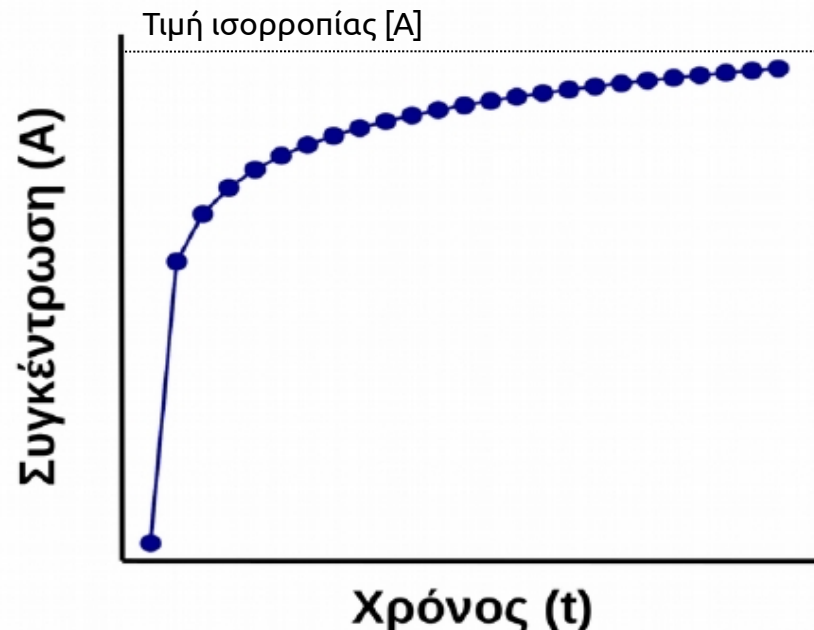
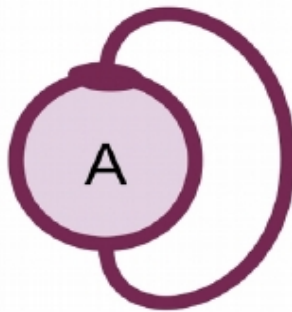


Αυτορύθμιση II: Καταστολή

Το A καταστέλει τον εαυτό του. Η μεταβολή της συγκέντρωσης του A σε σχέση με το χρόνο περιγράφεται από αντίστροφη εκθετική καμπύλη.

Βιολογικό αποτέλεσμα: Γρήγορη σταθεροποίηση της συγκέντρωσης του A σε μια τιμή ισορροπίας.

β



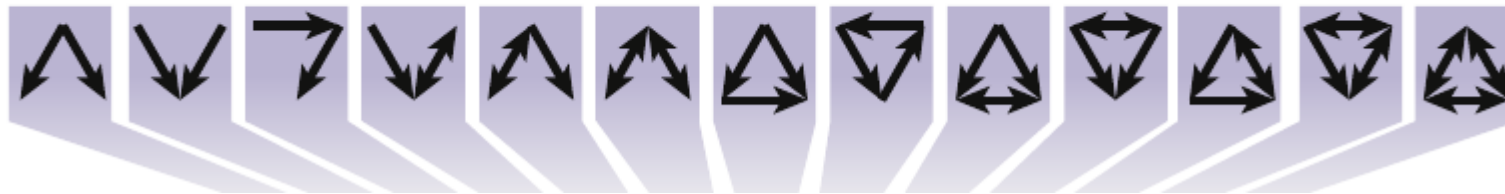
Διασυνδεδεμένες τριάδες

Αποτελούν την πιο κοινή κατηγορία μοτίβων σε βιολογικά συστήματα όλων των τύπων.

Ανάλογα με την τοπολογία τους (διάταξη των ακμών) μπορούν να προσδώσουν διαφορετικές λειτουργίες.

A

Network Motifs



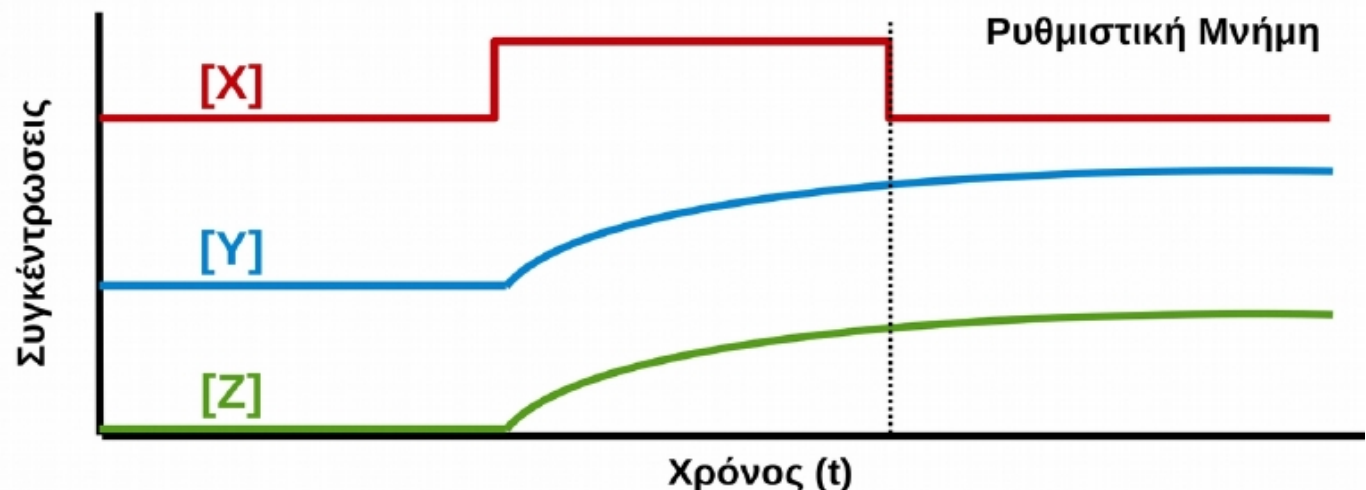
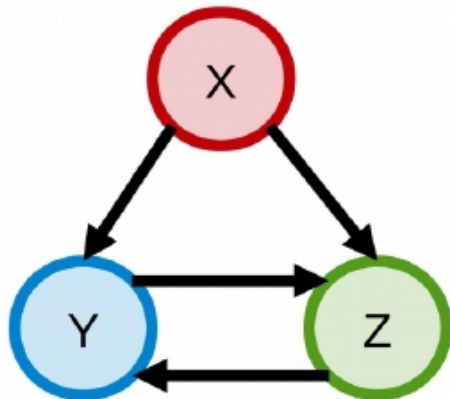
Πλήρως διασυνδεδεμένες τριάδες I: Ρυθμιστική Μνήμη

Πρόδρομη επαγωγή.

Το X επάγει τα Y,Z που μεταξύ τους αλληλο-επάγονται.

Βιολογικό αποτέλεσμα: Αρκεί μια αρχική αύξηση (παλμός) στη συγκέντρωση του X, ώστε να έχουμε διαρκή έκφραση των Y, Z.

α



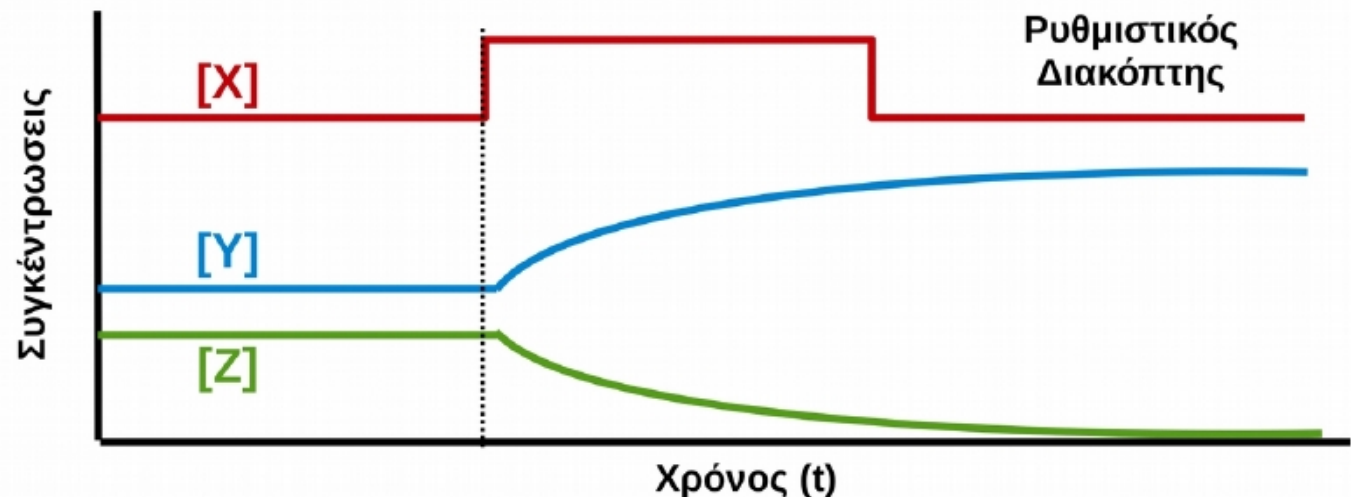
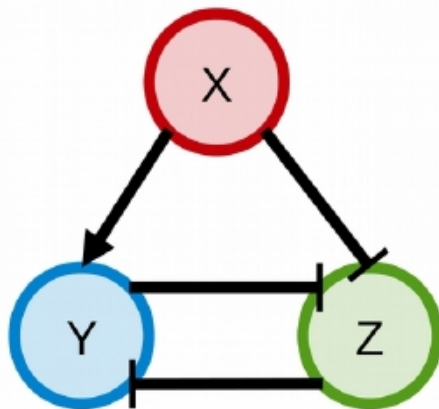
Πλήρως διασυνδεδεμένες τριάδες II: Ρυθμιστικός Διακόπτης

Διαφορική ρύθμιση.

Το X επάγει το Y αλλά καταστέλει το Z. Τα Y,Z αλληλοκαταστέλλονται.

Βιολογικό αποτέλεσμα: Αρκεί μια αρχική αύξηση (παλμός) στη συγκέντρωση του X, ώστε να αλλάξει το πρότυπο έκφρασης από Z σε Y.

β

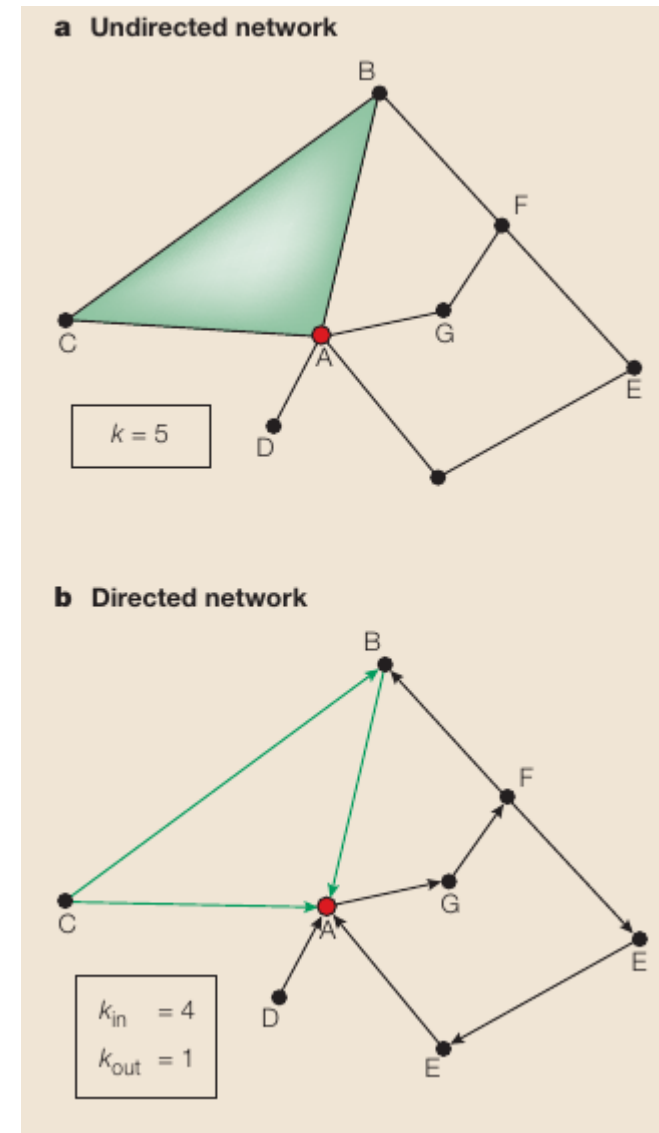


Εισαγωγή στη Θεωρία Δικτύων

Θεωρία Δικτύων είναι ο κλάδος των μαθηματικών που μελετά ιδιότητες των δικτύων με σκοπό να εξαγάγει συμπεράσματα για το φαινόμενο το οποίο αυτά περιγράφουν.

Μας ενδιαφέρει:

- Το είδος του δικτύου
- Το μέγεθος του δικτύου (αριθμός κόμβων)
- Η πυκνότητα/συνδεσιμότητα του δικτύου (σχέση του αριθμού κόμβων/ακμών)
- Η τοπολογία του. Πώς κατανέμονται οι ακμές στους κόμβους, πόσο στενά συνδεδεμένοι είναι οι κόμβοι μεταξύ τους κλπ



Μέγεθος και Πυκνότητα Δικτύου

Ονομάζουμε μέγεθος το σύνολο των κόμβων ενός δικτύου (N).

$$E_{\text{exp}}(N) = \frac{N(N-1)}{2}$$

Ορίζουμε ως πυκνότητα (d) του δικτύου τον αριθμό των ακμών που αυτό έχει σε σχέση με τον μέγιστο αριθμό δυνατών ακμών.

$$d = \frac{E_{\text{obs}}}{E_{\text{exp}}}$$

Ο μέγιστος αριθμός ακμών E_{exp} αντιστοιχεί στο πλήρως διασυνδεδεμένο δίκτυο.

Βαθμός Κόμβου (degree)

Ορίζουμε ως βαθμό ενός κόμβου τον αριθμό των ακμών του.

Αυτός μπορεί να διακρίνεται σε εισερχόμενες (incoming) και εξερχόμενες (outgoing) ακμές οπότε ορίζουμε και ανάλογα in-degree, out-degree.

Ισχύει προφανώς ότι:

$$D(n) = I(n) + O(n)$$

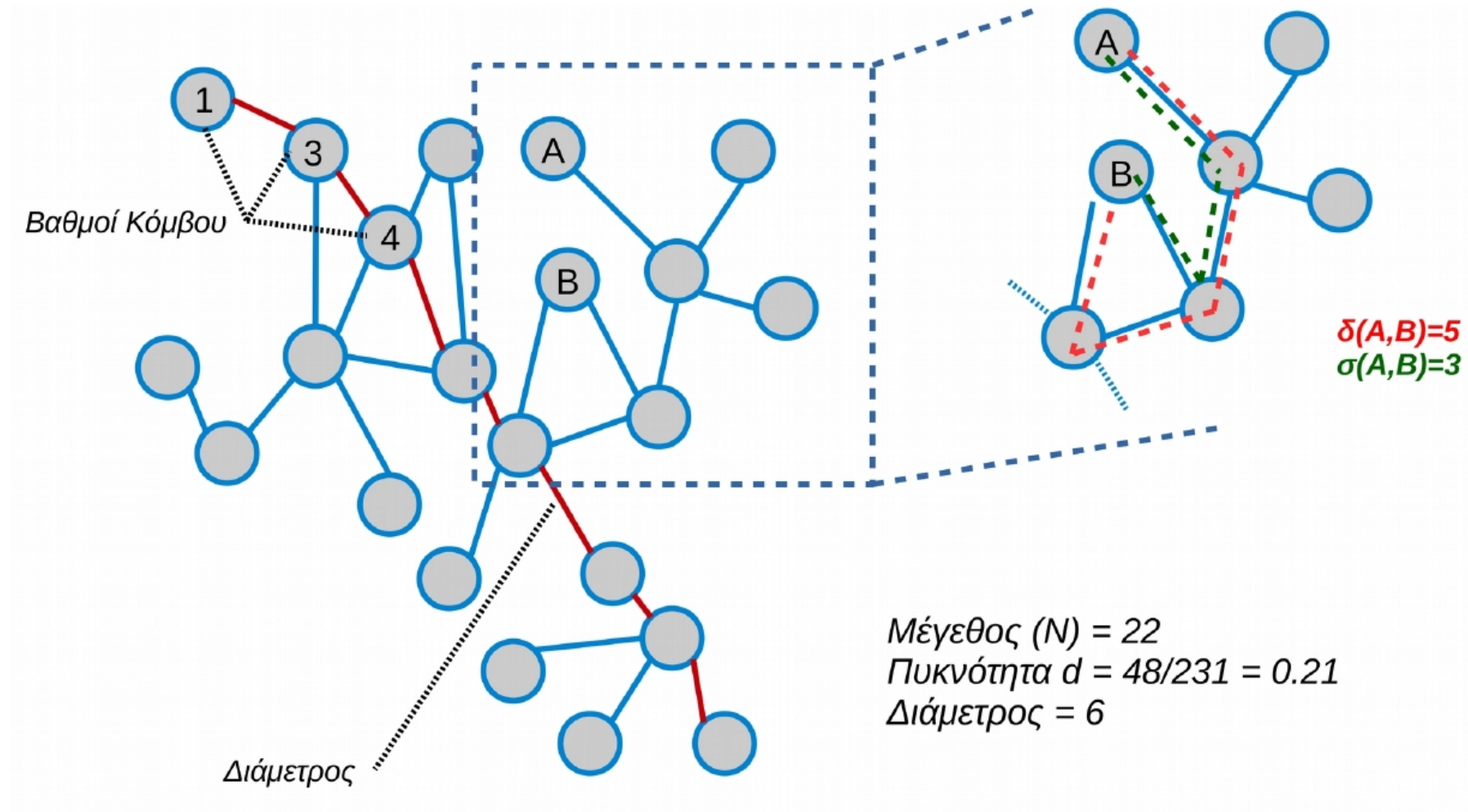
Σε περιπτώσεις δικτύων με βάρη, ο βαθμός μπορεί να σταθμίζεται με βάση τα βάρη των ακμών.

Συντομότερη Διαδρομή (shortest path) Διάμετρος Δικτύου

Ορίζουμε ως συντομότερη διαδρομή μεταξύ των κόμβων A και B τον ελάχιστο αριθμό κόμβων (σ) από τους οποίους θα πρέπει να περάσει κανείς για να φτάσει από τον κόμβο A στον B, διατρέχοντας υπάρχουσες ακμές στο δίκτυο.

Ορίζουμε διάμετρο ενός δικτύου τη μεγαλύτερη μεταξύ όλων των ελάχιστων διαδρομών του δικτύου.

Ποσοτικά Χαρακτηριστικά Δικτύων



Μέτρα Κεντρικότητας

Κεντρικότητα εγγύτητας του κόμβου A (closeness centrality) είναι το αντίστροφο του αθροίσματος των αποστάσεων του A από όλους τους κόμβους.

$$C(A) = \frac{1}{\sum_{i=1}^{N-1} \sigma_{A,i}}$$

Διακεντρικότητα (betweenness centrality) είναι ο λόγος των συντομότερων διαδρομών που περνούν από το A προς το σύνολο των συντομότερων διαδρομών

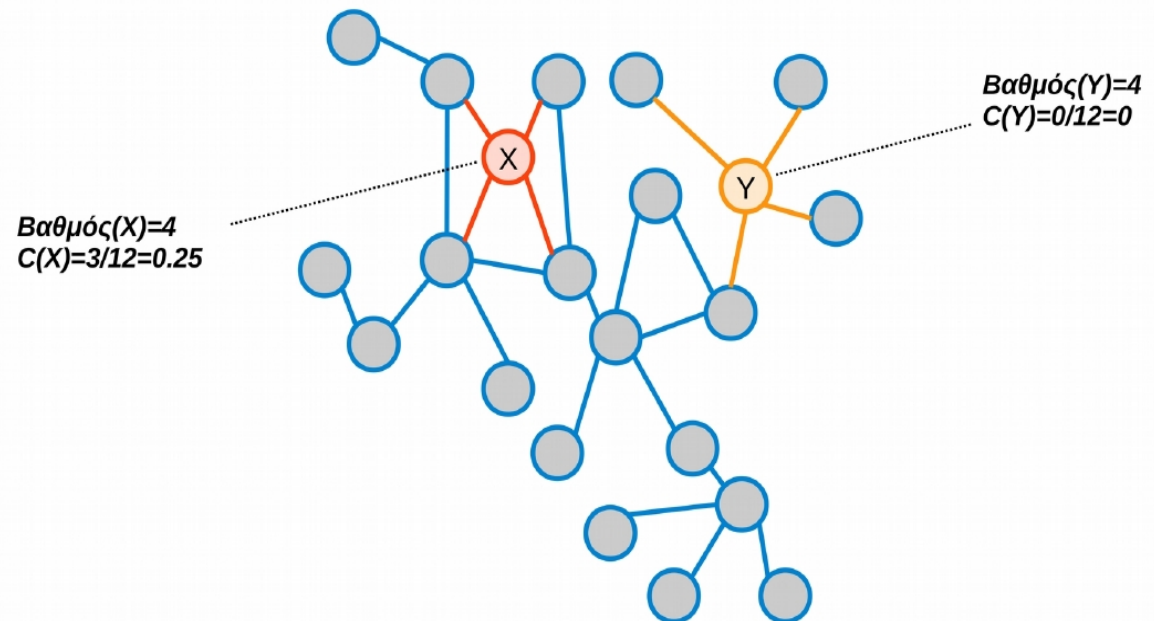
$$G(A) = \frac{\sigma_{i,j}(A)}{\sum_{i=1, j=1, i \neq j \neq A}^N \sigma_{i,j}}$$

Συντελεστής Συσσωμάτωσης και Συνδεσιμότητα (connectedness)

Είναι ένα πολύ σημαντικό τοπολογικό μέτρο για έναν κόμβο, καθώς σχετίζεται με την τάση των κόμβων να δημιουργούν τοπικά-υποδίκτυα (και κατά συνέπεια μοτίβα). Ορίζεται ως ο λόγος των άμεσα διασυνδεδεμένων πρώτων k γειτόνων του A προς τη μέγιστη τιμή των μεταξύ τους συνδέσεων.

Με τιμές από 0 έως 1, ο μέσος συντελεστής συσσωμάτωσης είναι ένα μέτρο της συνδεσιμότητας του δικτύου, πόσο δηλαδή συνεκτικό, “πυκνό” τείνει να είναι.

$$C(A) = \frac{\sum_{i=1, j=1, i \neq j}^k \text{edges}_{i,j}}{k(k-1)}$$



Συντελεστής Συσχέτισης και Καταταξιμότητα (assortativity)

Ή γιατί οι πρωτεϊνές δεν είναι celebrities;

Ο συντελεστής συσχέτισης ισούται με το συντελεστή συσχέτισης Pearson του βαθμού των κόμβων μεταξύ των άμεσα διασυνδεδεμένων κόμβων. Είναι δηλαδή μια ποσότητα που μας λέει κατά πόσο κόμβοι που συνδέονται μεταξύ τους τείνουν να έχουν παραπλήσιους βαθμούς.

Η τάση αυτή ονομάζεται καταταξιμότητα και είναι μια βασική ιδιότητα που διαφοροποιεί δίκτυα μεταξύ τους. Έτσι π.χ. κοινωνικά δίκτυα (όπως το δίκτυο των συζύγων μεταξύ των ηθοποιών του Χόλυγουντ) έχουν μεγάλη καταταξιμότητα.

Τα δίκτυα πρωτεϊνών από την άλλη έχουν μικρή καταταξιμότητα. Συνδέσεις δηλαδή αναδύονται κυρίως μεταξύ “ισχυρών” και “ασθενών” κόμβων.

Στατιστικές Ιδιότητες Δικτύων

Όλες οι παραπάνω ποσότητες μπορούν να μας βοηθήσουν να διακρίνουμε δίκτυα με βάση τις στατιστικές τους ιδιότητες. Οι ιδιότητες αυτές είναι σημαντικές γιατί αντανακλούν τα φαινόμενα που βρίσκονται πίσω από τα υπό μελέτη δίκτυα.

Τα δίκτυα γενικώς μπορούν να διακριθούν σε:

1. Τυχαία Δίκτυα (small world)
2. Δίκτυα Ανεξάρτητα Κλίμακας (scale free)
3. Ιεραρχικά-Σπονδυλωτά Δίκτυα (modular)

Από τα παραπάνω, τα δύο πρώτα μπορούν να προκύψουν από φυσικά φαινόμενα και παρατηρούνται σε βιολογικά συστήματα. Τα τελευταία είναι κυρίως αποτέλεσμα αρχιτεκτονικής και σχεδιασμού και μπορούν να προσομοιαστούν από βιολογικά φαινόμενα μόνο κατά προσέγγιση.

Τυχαία Δίκτυα

Είναι τα δίκτυα των οποίων οι συνδέσεις μπορούν να προκύψουν με μια τυχαία κατανομή. Πρακτικά μπορούμε να το φανταστούμε ως τη διαδοχική “ένωση” δύο στοιχείων επιλεγμένων τυχαία με πιθανότητα P . Ανάλογα με την τιμή της P το δίκτυο μπορεί να είναι πυκνότερο ή αραιότερο.

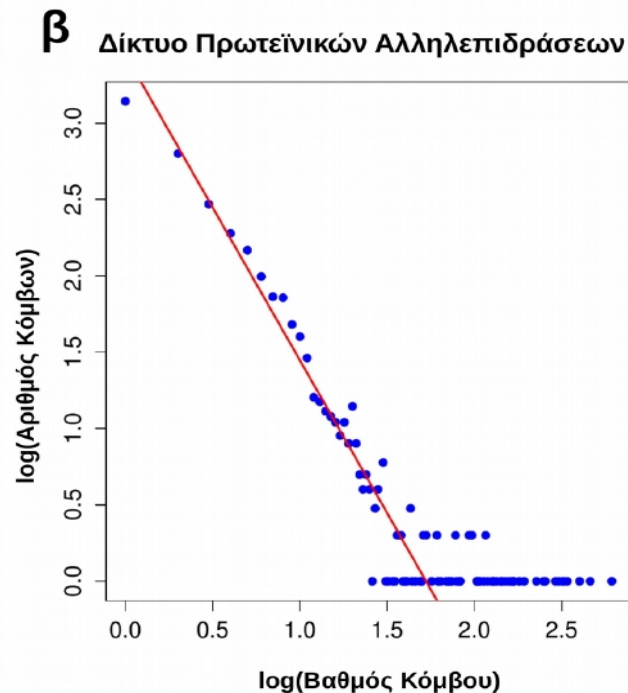
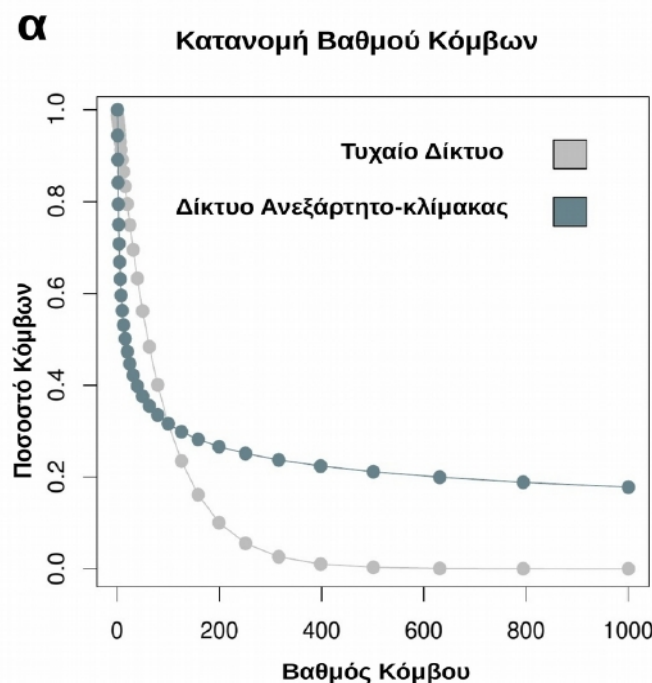
Βασική Στατιστική ιδιότητα των τυχαίων δικτύων είναι ότι η κατανομή των βαθμών των κόμβων τους ακολουθεί εκθετική φθίνουσα κατανομή.

Τυχαία δίκτυα παρατηρούνται στη βιολογία σε επιδημιολογικά συστήματα, όπου η μετάδοση ασθενείων ακολουθεί μια τυχαία πορεία. Ονομάζονται και “μικρού κόσμου” (small world) γιατί οδηγούν συχνά σε διαισθητικά μη-αναμενόμενα αποτελέσματα. Π.χ. η μετάδοση μιας ασθένειας ακόμα και με μικρή μεταδοτικότητα μπορεί να συμβεί ταχύτατα σε ένα τυχαίο δίκτυο.

Δίκτυα Ανεξάρτητα Κλίμακας

Είναι τα πιο ενδιαφέροντα από βιολογική άποψη. Βασική τους ιδιότητα είναι ότι η κατανομή του βαθμού των κόμβων τους ακολουθεί κατανομή “νόμου δύναμης” (power-law)

$$p(k) \sim k^{-\alpha} \text{ (με } 2 < \alpha < 3)$$



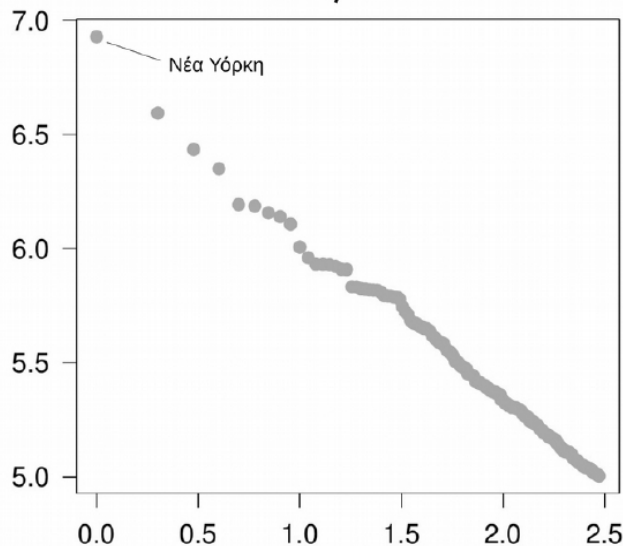
Νόμοι Δύναμης (είναι παντού!)

Η φυσική σημασία των νόμων δύναμης αναδεικνύεται από την εκπληκτική διάδοσή τους σε διάφορα δυναμικά συστήματα. Ουσιαστικά συνίσταται σε μια αρχή ομοιομορφίας σε διάφορες κλίμακες μεγέθους (scale-free, fractal ιδιότητα). Αυτό σημαίνει ότι πρακτικά η κατανομή δε διαφοροποιείται μεταξύ πολλών τάξεων μεγέθους.

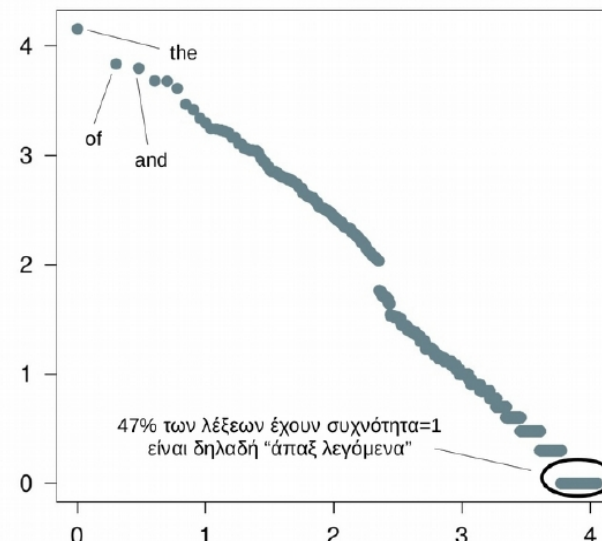
Το φυσικό αποτέλεσμα για τα δίκτυα είναι συστήματα όπου ένα πολύ μικρό ποσοστό κόμβων συγκεντρώνει ένα συντριπτικό ποσοστό των ακμών, ενώ αντίθετα η μεγάλη πλειοψηφία είναι φτωχοί κόμβοι (αρχή του 80:20, 90:10 ή της "Wall Street")

α

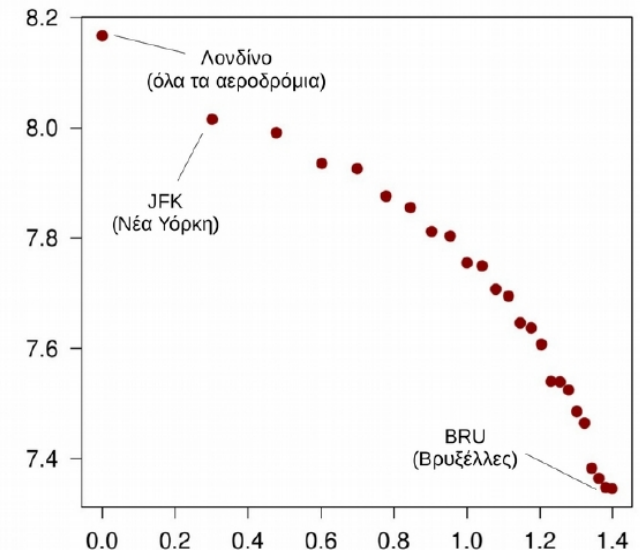
Πληθυσμός Πόλεων
στις ΗΠΑ

**β**

Συχνότητα Λέξεων
στο "Moby Dick"

**γ**

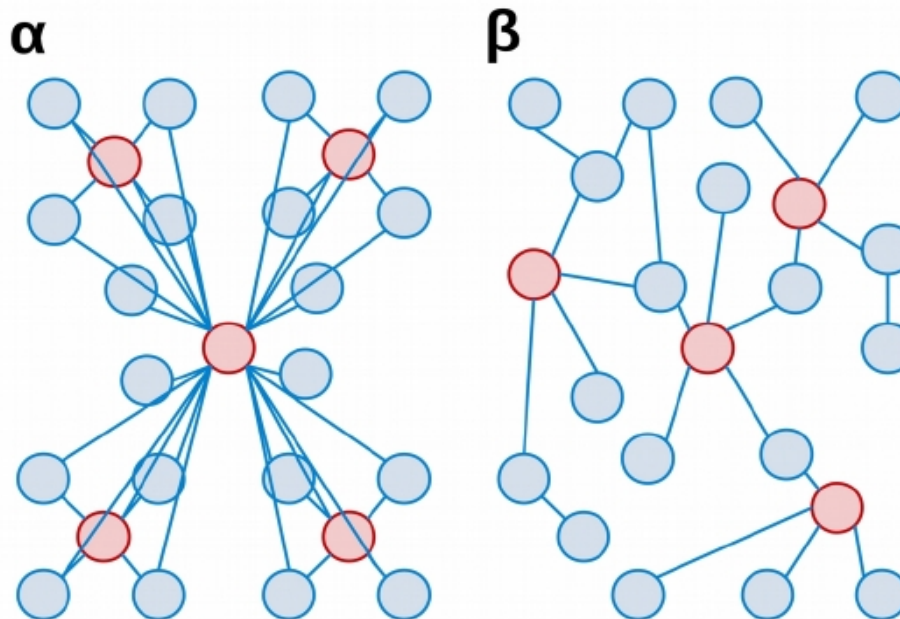
Αριθμός Επιβατών σε
διεθνή αεροδρόμια



Σπονδύλωτα Δικτύα και Δίκτυα Ανεξάρτητα Κλίμακας

Σπονδυλωτά (modular) δίκτυα είναι αυτά με ιεραρχία που είναι συνήθως αποτέλεσμα σχεδιασμού. Βασική τους στατιστική ιδιότητα ότι εκτός από το βαθμό των κόμβων ΚΑΙ ο συντελεστής συσσωμάτωσής τους ακολουθεί νόμο δύναμης.

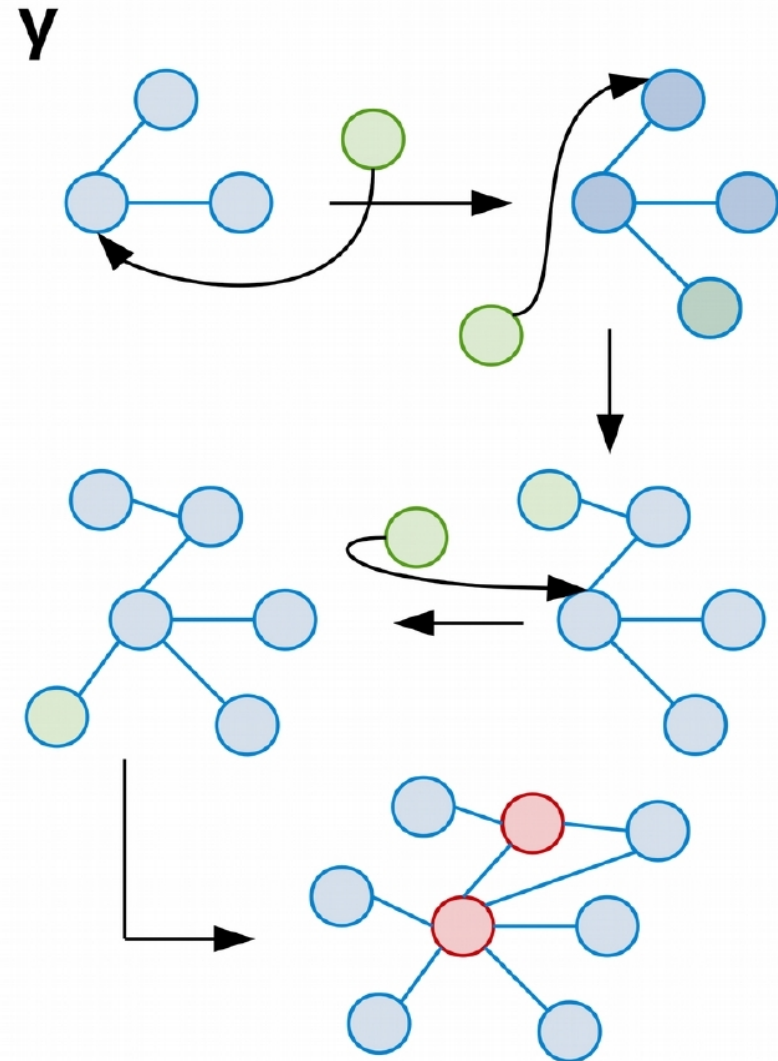
Σε βιολογικά φαινόμενα δεν παρατηρούμε modular δίκτυα αλλά κυρίως ΑΚ. Η σημασία των ΑΚ-δικτύων στη βιολογία σχετίζεται με την ισορροπία που προσφέρουν μεταξύ σταθερότητας και λειτουργικότητας. Αφενός επιτρέπουν ένα βαθμό ιεραρχίας που θα μπορούσε μόνο να προκύψει από σχεδιασμένα (και όχι τυχαία) δίκτυα, από την άλλη έχουν έναν επίπεδο πλεονασμού που τα καθιστά σταθερά σε διαταραχές (απαλοιφές κόμβων, ακμών κλπ).



Μοντέλο Προτιμησιακής Σύνδεσης

Είναι λοιπόν τα βιολογικά δίκτυα αποτέλεσμα “σχεδιασμού”; Φυσικά και όχι (όχι με την έννοια του “ευφυούς σχεδιαστή”).

Οι ιδιότητες ΑΚ μπορούν να αναδυθούν αυθόρμητα μέσω μιας δυναμικής διαδικασίας “προτιμησιακής σύνδεσης” μέσω της οποίας η πιθανότητα αύξησης του βαθμού ενός κόμβου εξαρτάται από τον ίδιο του το βαθμό.



Το φαινόμενο του Ματθαίου!

“τῷ γὰρ ἔχοντι παντὶ δοθήσεται καὶ περισσευθήσῃ
 ἔχοντος καὶ ὃ ἔχει ἀρθήσεται” $P(i) = \frac{k_i}{\sum k_i}$
 Καταματθαίου!

Καλό Πάσχα!

Διαβάστε περισσότερα

Όλα όσα θα θέλατε να ξέρατε για τα δίκτυα και δεν τολμάτε να ρωτήσετε στο βιβλίο του θεμελιωτή της μοντέρνας θεωρίας δικτύων, Albert-Laszlo Barabasi:

<http://barabasilab.neu.edu/networksciencebook/downloadPDF.html>