# Constraints and Metrics in Technical Networks

# Goals of This Class

- Metrics as indicators of system properties that may be related to structure or behavior
- Their relation to the constraints under which the systems evolved or in view of which the systems were designed
- Modeling problems

# Designed and Grown Systems

- Designed implies some degree of top-down control of the architecture
- Grown does not mean random
- Social systems
  - Designed: organizations, supply chains
  - Grown: coauthors, company directors
- Technical systems
  - Designed: assemblies, PSTN, factories, national highway network
  - Grown: regional or national electric grid, local roads outside of Northwest Territory
- Harder to classify (social?, grown?)
  - A city and its water supply or subway system

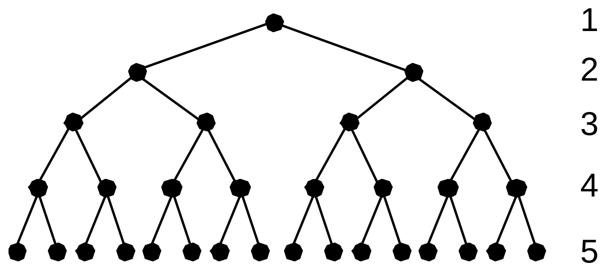
# Pearson Coefficient for Technical Systems

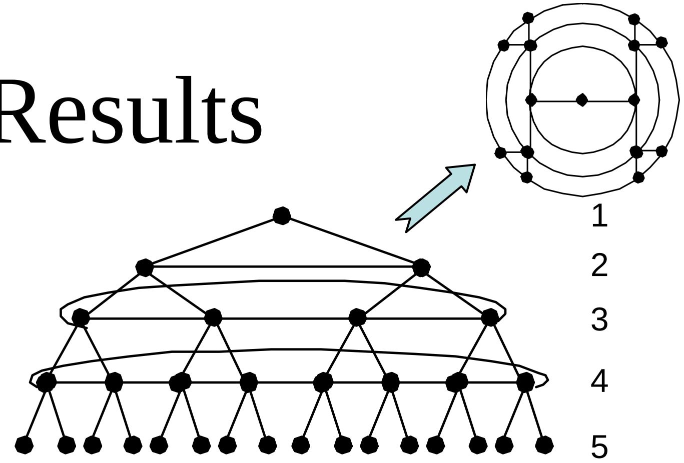
- A widely studied metric that captures some elements of structure and possibly is related to behavior
- Distribution systems - grids or stars or trees or trees with cross-links
- Mechanical assemblies - trees
- Electric and electronic circuits - should be grids
  - Computer motherboards have a few nodes with huge k
  - What do they look like inside?
  - Coarse-graining

# Pearson Coefficient for Canonical Systems

- Trees, cross-linked trees, and stars have $r < 0$
- Balanced binary trees have $r = -1/3$
- Trees with diminishing branching ratio have $r > 0$
- Trees with big branching ratios explode and r approaches -1
- Finite grids have r approaching 2/3
- Clusters with pendants at each cluster node have $r < 0$

# Closed Form Results

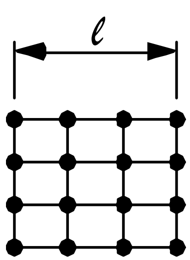



| **Property** | **Pure Binary Tree** | **Binary Tree with Cross-linking** |
|---|---|---|
| *ksum* | $2^{n+1}-4$ | $3*2^{n}-10$ |
| *ksqsum* | $10*2^{n-1}-14$ | $13*2^{n}-64$ |
| $\overline{x}$ | $\to 2.5$ as $n$ becomes large (>~ 6) | $\to \frac{13}{3}$ as $n$ becomes large (>~ 6) |
| Pearson numerator | $\sim 2^{n}(3-\overline{x})(1-\overline{x})+(ksum-2^{n})(3-\overline{x})^{2}$ | $\sim 2^{n}(5-\overline{x})(1-\overline{x})+(ksum-2^{n})(5-\overline{x})^{2}$ |
| Pearson denominator | $\sim 2^{n-1}(1-\overline{x})^{2}+(ksum-2^{n-1})(3-\overline{x})^{2}$ | $\sim 2^{n-1}(1-\overline{x})^{2}+(ksum-2^{n-1})(5-\overline{x})^{2}$ |
| $r$ | $\to -\frac{1}{3}$ as $n$ becomes large | $\to -\frac{1}{5}$ as $n$ becomes large |

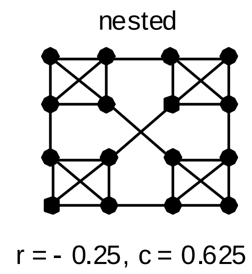
Note: Western Power Grid r = 0.0035

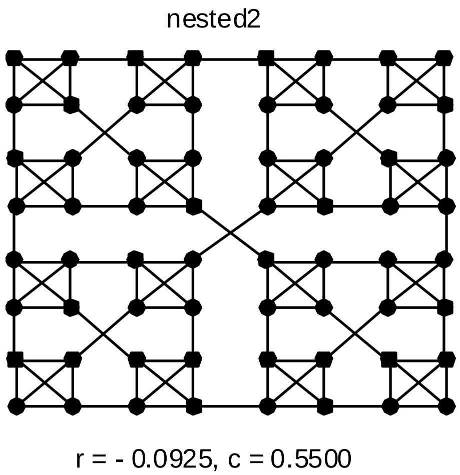
Bounded grid




$$r=\frac{16(2-\overline{x})(3-\overline{x})+8(\ell-3)(3-\overline{x})^{2}}{2(2-\overline{x})^{2}+12(\ell-2)(3-\overline{x})^{2}}\to\frac{2}{3}$$

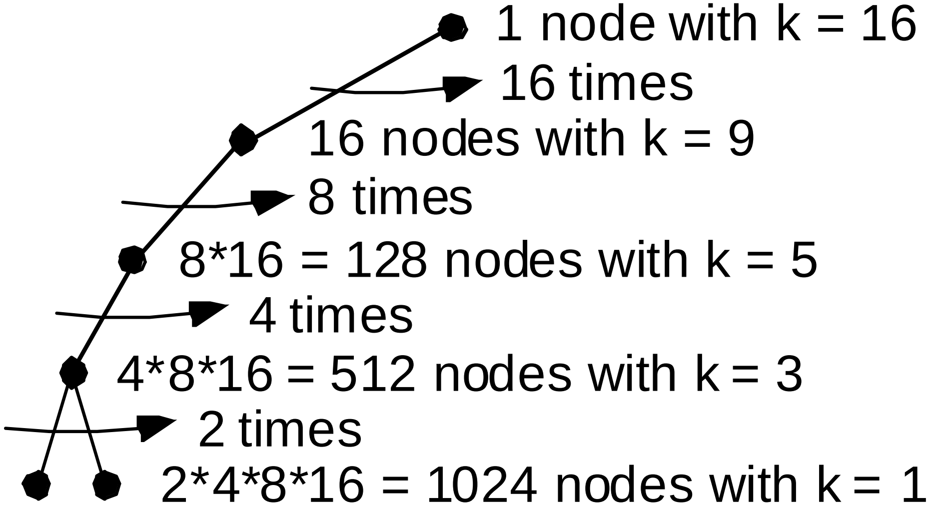
# Nested Self-Similar Networks



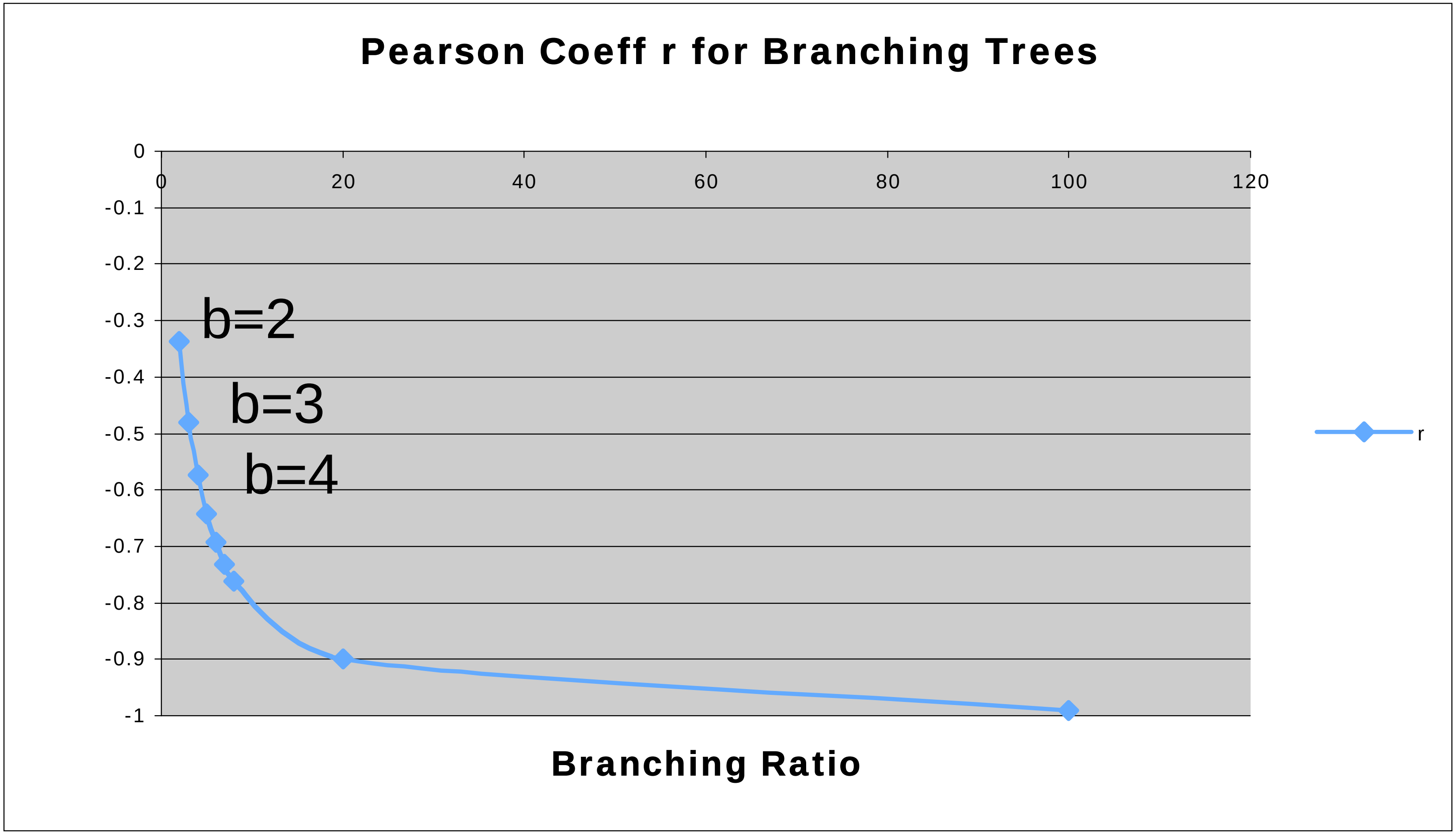

Probably, r = 0 in the limit as the network grows

# Tree with Diminishing Branching Ratio




r = 0.38166

# Trees with Branching Ratio b




Using approximate formula; tested in matlab with tree-generator written by Mo-Han Hsieh. Actual values are a bit more negative than given by approximate formula for b > 2.

# Alignment of System Type and Canonical Shape

- Assemblies seem like trees, or trees decorated with loose clusters at different hierarchical levels
- Subway systems are like clusters with interior nodes and exterior pendants
- Commuter rail lines can be trees, trees with cross-links, or grids
  - Grids arise when there is a robust intercity rail system and commuter trains can share these tracks
  - Also helps to have relatively flat ground
  - Trains must follow flat ground or cost a lot for tunnels
  - Flat ground is associated with water courses, as are locations of towns that need train service

# Network Models of Technical Systems

- Need to carefully define what is a node and what is a link
- Examples:
  - Assemblies: node = part; link = joint between two parts tat constrains at least one degree of freedom
  - Rail lines: node = rail junction or place where people can change train lines; link = rail
  - Electric circuit: node = circuit element (R, C, IC); link = wire
  - Distribution infrastructure: node = branch point, load, or sink; link = conductor ( pipe, wire)
  - Food web: node = species; link (directed) what eats what (can include cannibalism)

# Design of Distribution Systems

- Fundamental need is to “fill space” in some sense
- Scaling issues
- Cost per unit of capability or capacity
- Levels of service: speed, choice of destination, equity
- Context, legacy
  - Ability to run commuter trains on inter-city tracks
  - Ability to exceed service of legacy system

# Spatial Distribution Networks -1

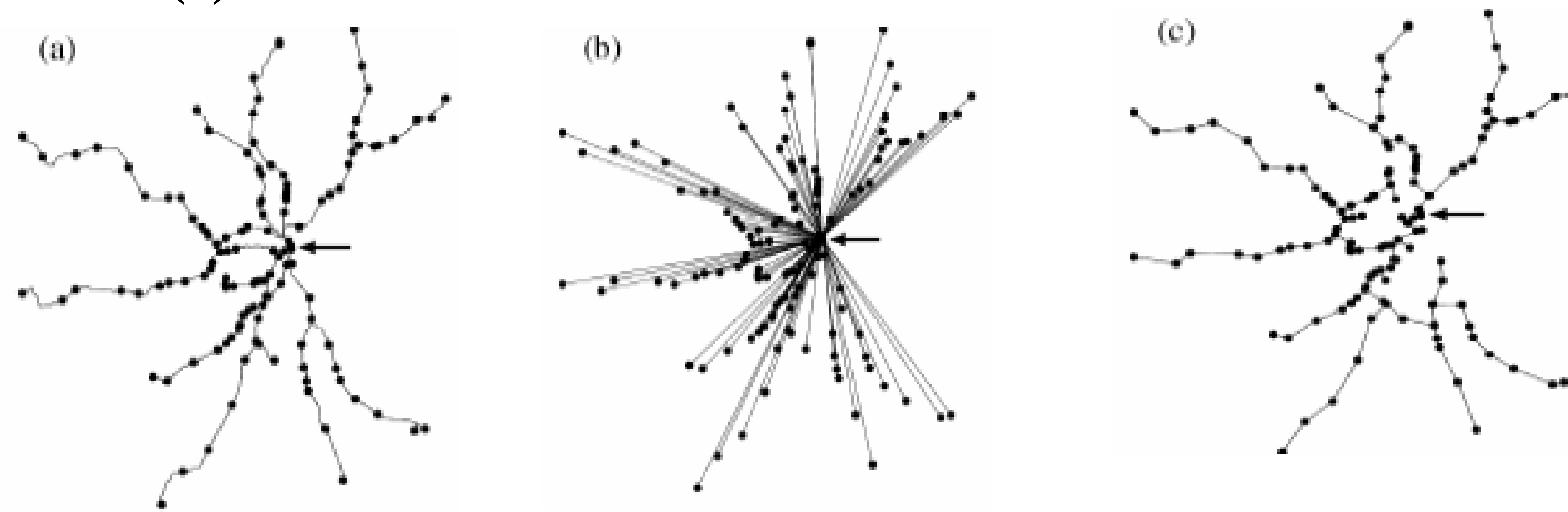
- Blood, water, sewer, newspapers are one to many or many to one
- Urban rail systems can be like this or can be many to many in the core and many to one in the periphery
- The phone system is many to many in the core and one to many in the periphery
- One to many = tree
- Many to many = grid or hierarchy with level-skipping

# Spatial distribution networks (Magee slides)

- Gastner and Newman analyzed the case where the distribution system has a “root node” which is the sole source or sink for the commodity being distributed.
- Additional design factors considered
  – Additional node locations (constraint)
  – Total link length (smaller lowers cost )
  – Shortest path length between two nodes (smaller lowers transport time)
- Tradeoff in last two factors is the design/architecting problem
  – Look at ideal solutions for each criterion
  – Examine how real networks compare on the tradeoffs
  – Build growth model to derive pattern and look for consistency

# Spatial distribution networks -2

- Minimum total edge length including paths to the root node is given by a Minimum Spanning Tree (c) while obtaining shortest paths from each node to the root node is optimized by a star graph (b). Actual system is (a)




From Gastner, M. T., and M. E. J. Newman. "Shape and efficiency in spatial distribution networks." *J Stat Mech* (2006): P01015. (Fig. 1) Courtesy the Institute of Physics. Used with permission.

# Spatial distribution networks - 3

- From transportation research, a route factor $q$ is where $\ell_{io}$ is the shortest actual path length and and $d_{io}$ is the shortest Euclidean distance from node $i$ to the origin $o$

$$q = \frac{1}{n}\sum_{i=1}^{n}\frac{\ell_{io}}{d_{io}} \geq 1$$

***q is equal to 1 for a star graph, which has maximum total edge length***

- For three real technological system networks,

| network | $n$ | route factor actual | route factor MST | edge length (km) actual | edge length (km) MST | edge length (km) star |
|---|---|---|---|---|---|---|
| sewer system | 23 922 | 1.59 | 2.93 | 498 | 421 | 102 998 |
| gas (WA) | 226 | 1.13 | 1.82 | 5 578 | 4 374 | 245 034 |
| gas (IL) | 490 | 1.48 | 2.42 | 6 547 | 4 009 | 59 595 |

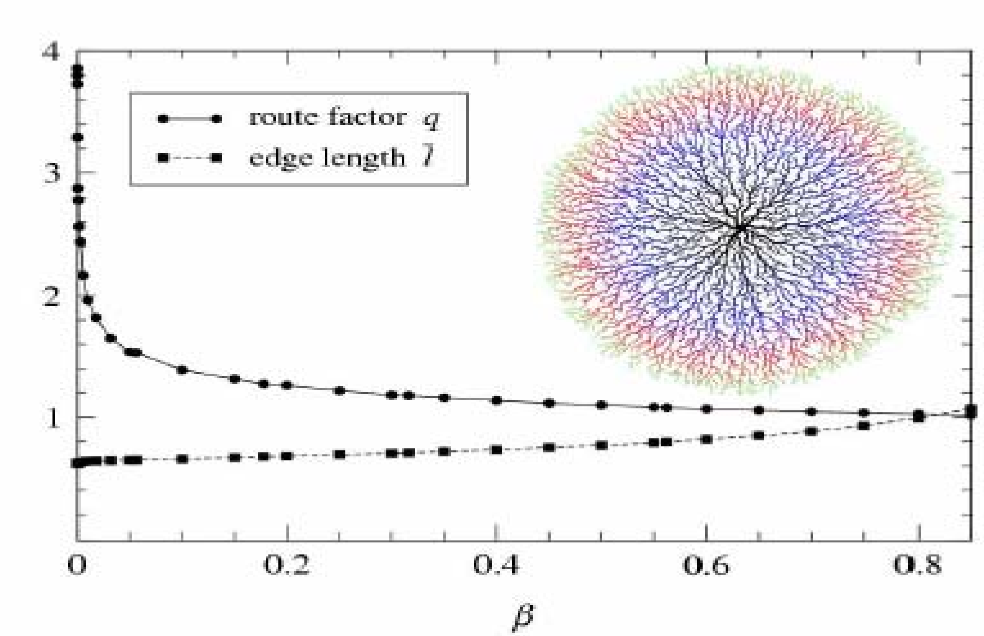
From Gastner, M. T., and M. E. J. Newman. "Shape and efficiency in spatial distribution networks." *J Stat Mech* (2006): P01015. (Table 1) Courtesy the Institute of Physics. Used with permission.

- The systems favor minimum edge length but have route factors considerably superior to MST optimums indicating *effective tradeoff in the two criteria.*
- A simple growth model is used to explain this result

# Spatial distribution networks -4

- The growth model assumes that the systems evolve from the root node by linking in new (but already existing) nodes using a greedy optimization criterion that adds unconnected node, i, to an already connected node, j with the weighting factor given by $w'_{ij} = d_{ij} + \beta l_{j0}.$
- Simulations using these model assumptions yield

showing small tradeoffs in total link length give large improvements in path length




From Gastner, M. T., and M. E. J. Newman. "Shape and efficiency in spatial distribution networks." J Stat Mech (2006): P01015. (Fig. 3) Courtesy the Institute of Physics. Used with permission.

# Spatial distribution networks 5

- What is **missing** from these studies of spatial distribution networks from your perspective? What ***future research*** do these studies suggest?
- Consideration of other network properties
- Consideration of constraints like geography
- Development of more broadly applicable models
  - More than one source/sink node
  - Coordination with other networks (subway and commuter or intercity rail)
- Development of other rules/protocols for growth that achieve the key properties well
- Consideration of top-down vs. evolved systems

# London Underground

$$r = 0.0997$$

Image removed for copyright reasons.
Map of the London subway system.
See: http://www.tfl.gov.uk/tfl/pdfdocs/colourmap.pdf

# Tokyo JR East Lines and Subways

$$r_{\text{regional rail}} = -0.0134$$

$$r_{\text{regional rail plus subways}} = 0.0425$$

Image removed for copyright reasons.
Map of the Tokyo railroad and subway systems.
See: http://www.jreast.co.jp/e/info/map_a4ol.pdf

# Tokyo Subways

Image removed for copyright reasons.
Map of the Tokyo subway system.
See: http://www.deutsch-japanischer-kulturverein.de/Images/Karte%20Tokyo%20Subway2.jpg

# Moscow Metro

Image removed for copyright reasons.
Map of the Moscow subway system.
See: http://meta.metro.ru/moskva/moscow-metro.gif

# Moscow Regional Rail

Images removed for copyright reasons.
Map of the Moscow Regional Rail.

# Moscow Metro and Regional Rail

Images removed for copyright reasons.
Map of the Moscow Metro and Regional Rail.

$$r_{\text{subway + rail}} = 0.2601$$

$$r_{\text{subway}} = 0.1846$$

# Berlin U-bahn and S-bahn

Images removed for copyright reasons.
Map of the Berlin subway and train.
See: http://www.lodging-germany.com/info/Berlin/berlin-7citymapubahn.htm



$r = 0.0957$

# Munich U-bahn and S-bahn

Images removed for copyright reasons.
Map of the Munich subway and train.
See: http://www.munich-info.de/images/mvv.jpg

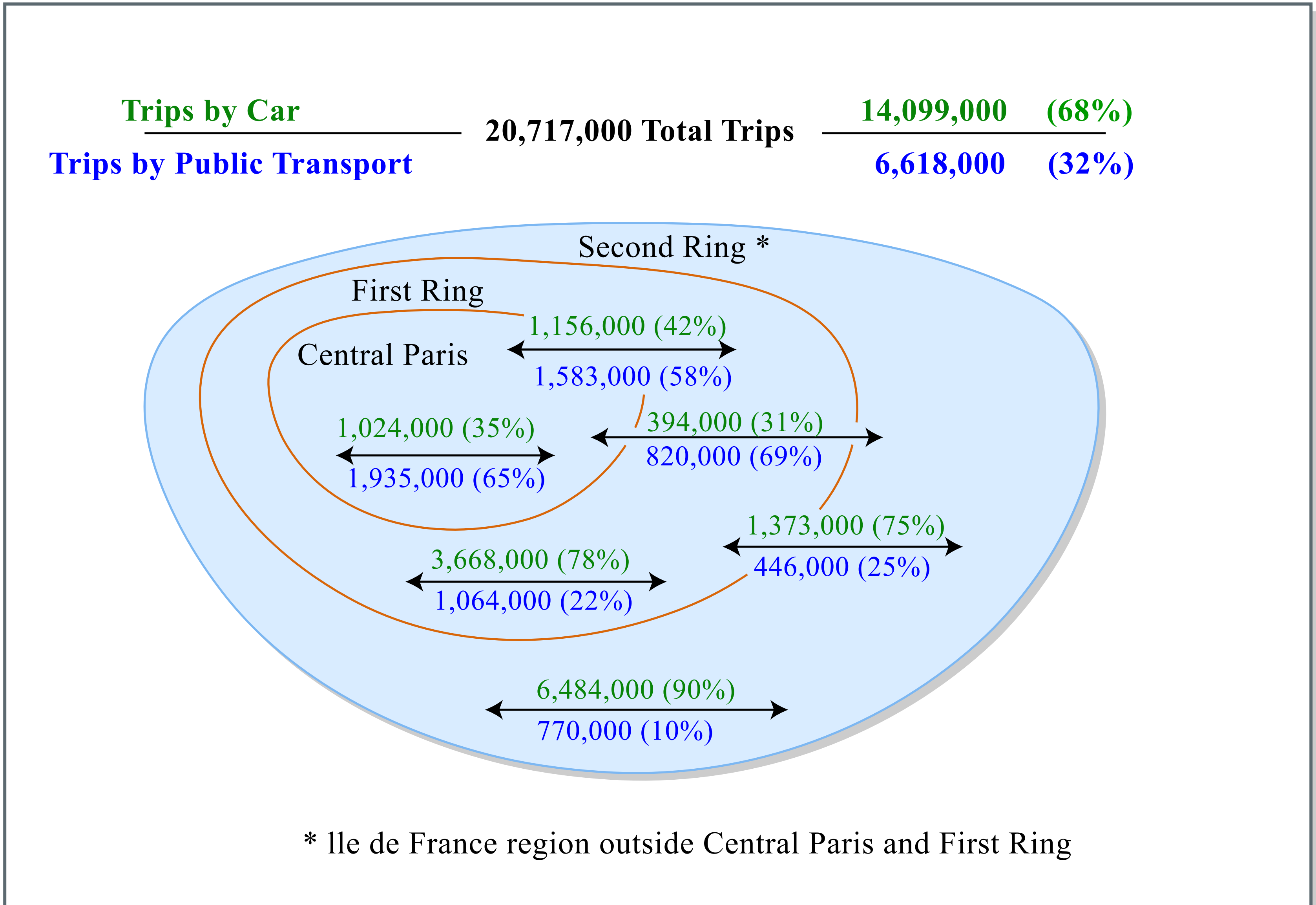
$r$ = -0.0317

# Paris Metro and RER

Image removed for copyright reasons.
Map of the Paris Metro.
See: http://www.kigoobe.com/parishotel/img/carte.gif

Image removed for copyright reasons.
Map of the Paris RER.
See: http://www.paris.org/Metro/gifs/rer01.map.jpg

# Car and Train Traffic, Paris



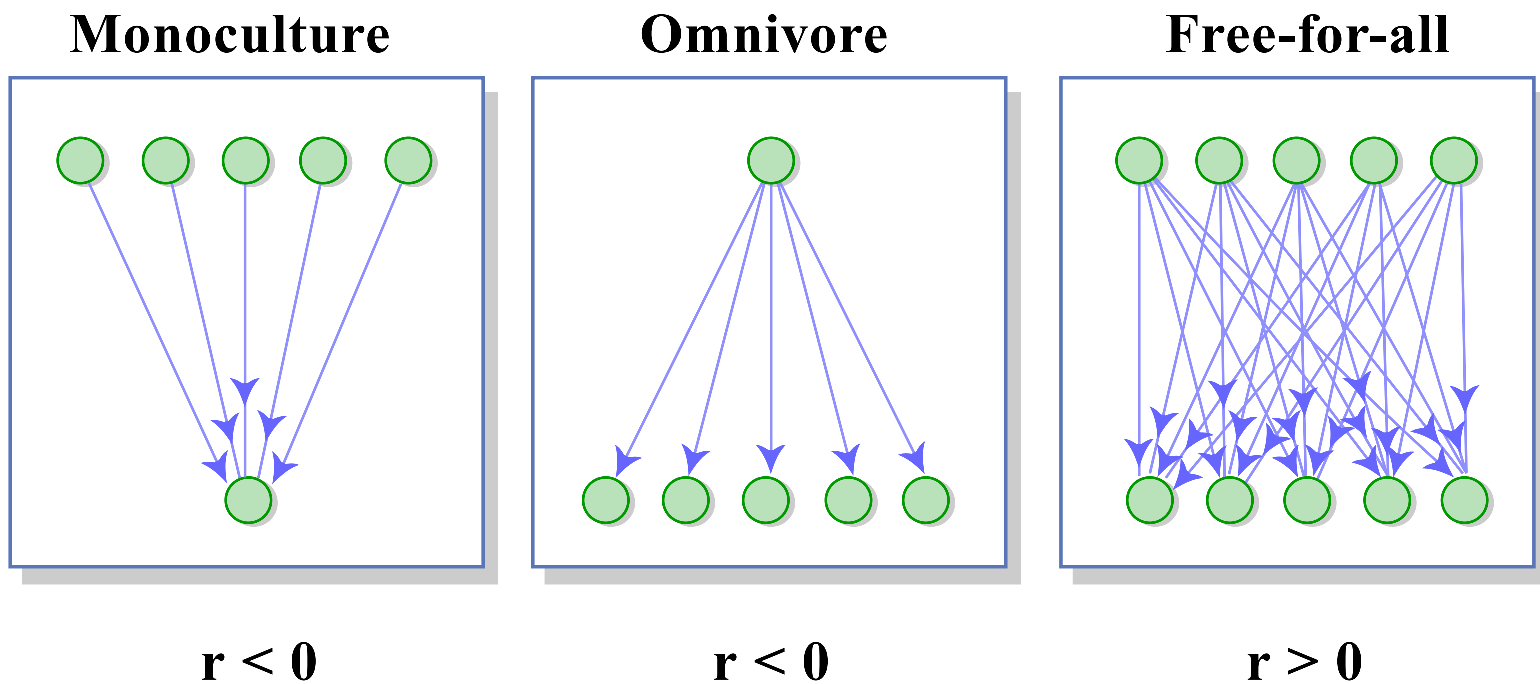

Daily Trips by Mode in the Paris Region

Figure by MIT OCW. After figure by Renault Cars UK.

# Food Webs

- These are directed graphs of what eats what
- They are classic hierarchies with typically three levels
  - Top - have no predators
  - Middle - are both prey and predators
  - Bottom - have no prey
- Trophic species are those with the same predators and prey, so that they look alike from the point of view of the web
- Some food webs are quoted in the network literature with or without trophic species condensation, without noting this difference
- Data in following slides are from Jennifer Dunne, Santa Fe Institute

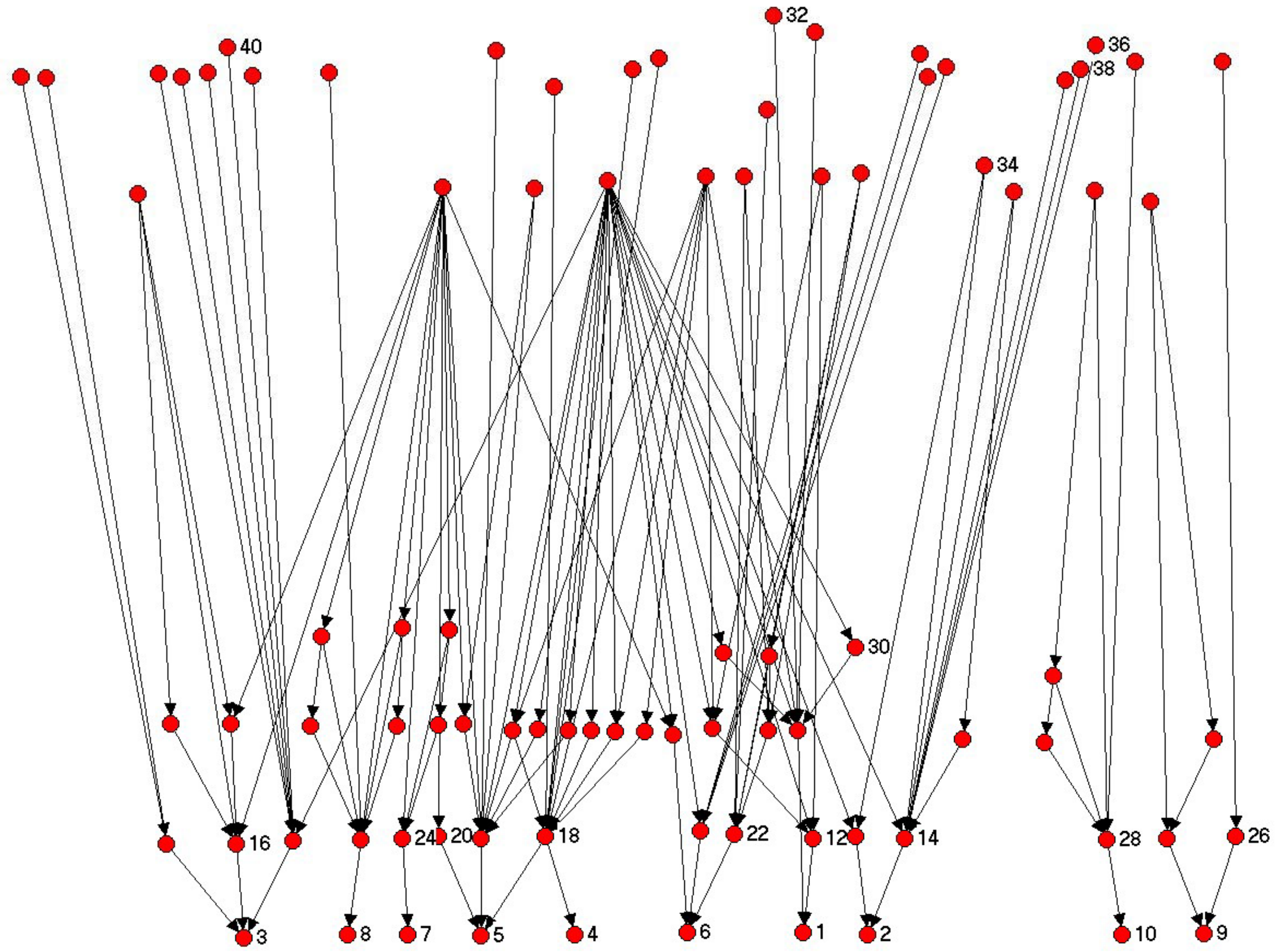
# Canonical Forms in Food Webs




Figures by MIT OCW.

$$r_{dir} = \frac{\sum (k_{out} - \overline{x}_{out})(k_{in} - \overline{x}_{in})}{\sqrt{\sum (k_{out} - \overline{x}_{out})^2 \sum (k_{in} - \overline{x}_{in})^2}}$$

# UK Grassland




r = - 0.1724

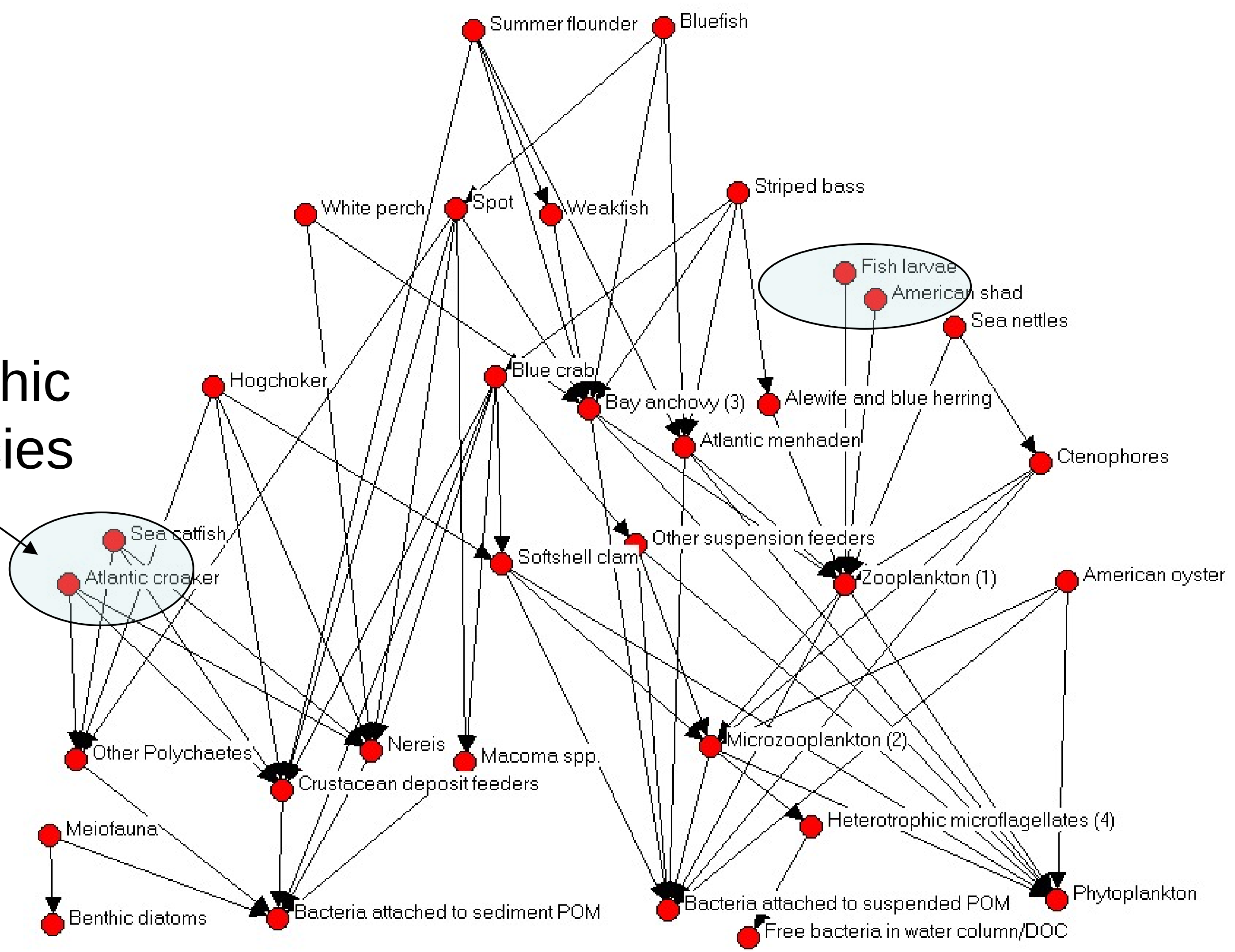
Parasitic insects

Herbivore insects

Grasses

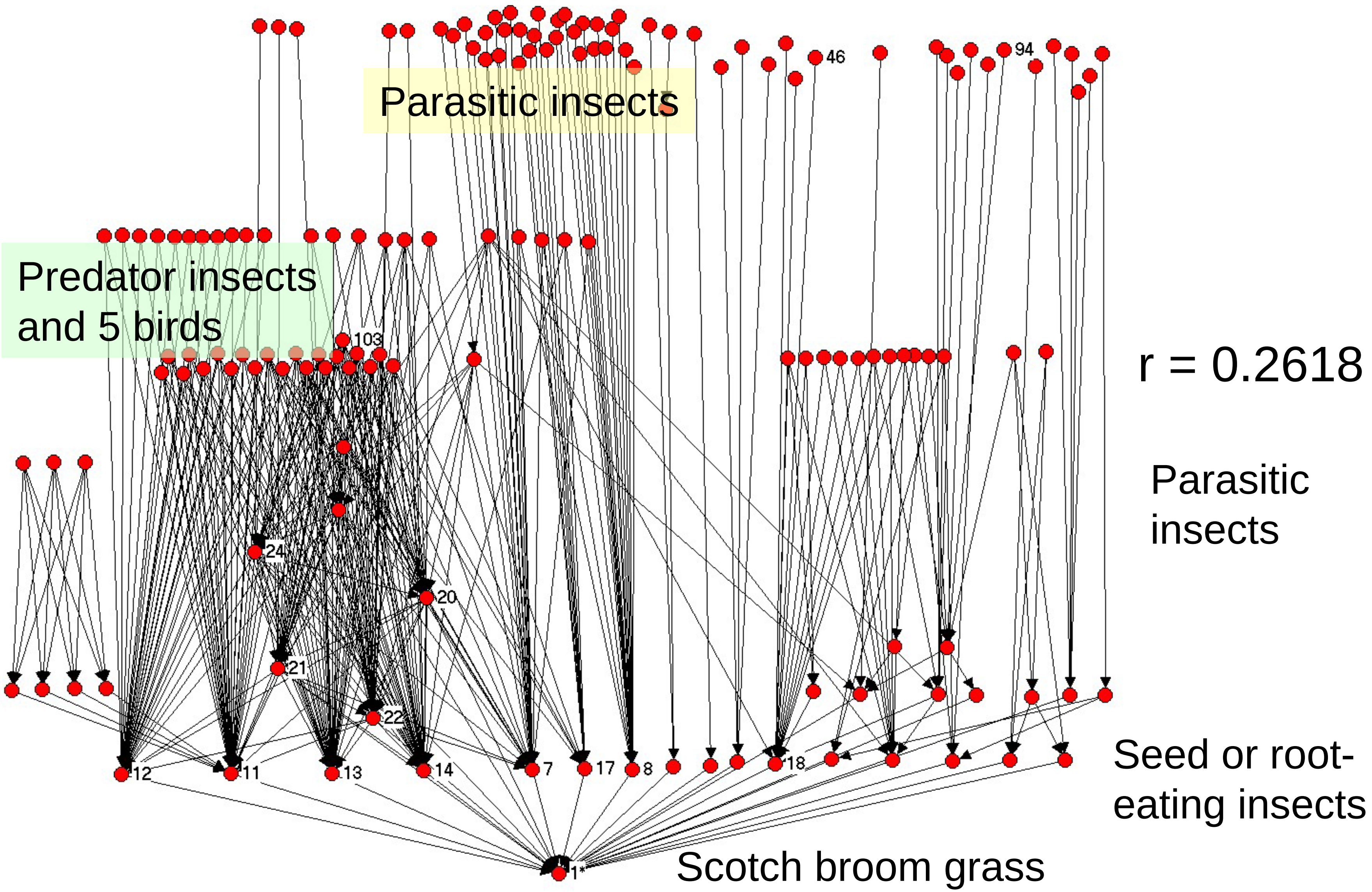
# Chesapeake Bay

Trophic species

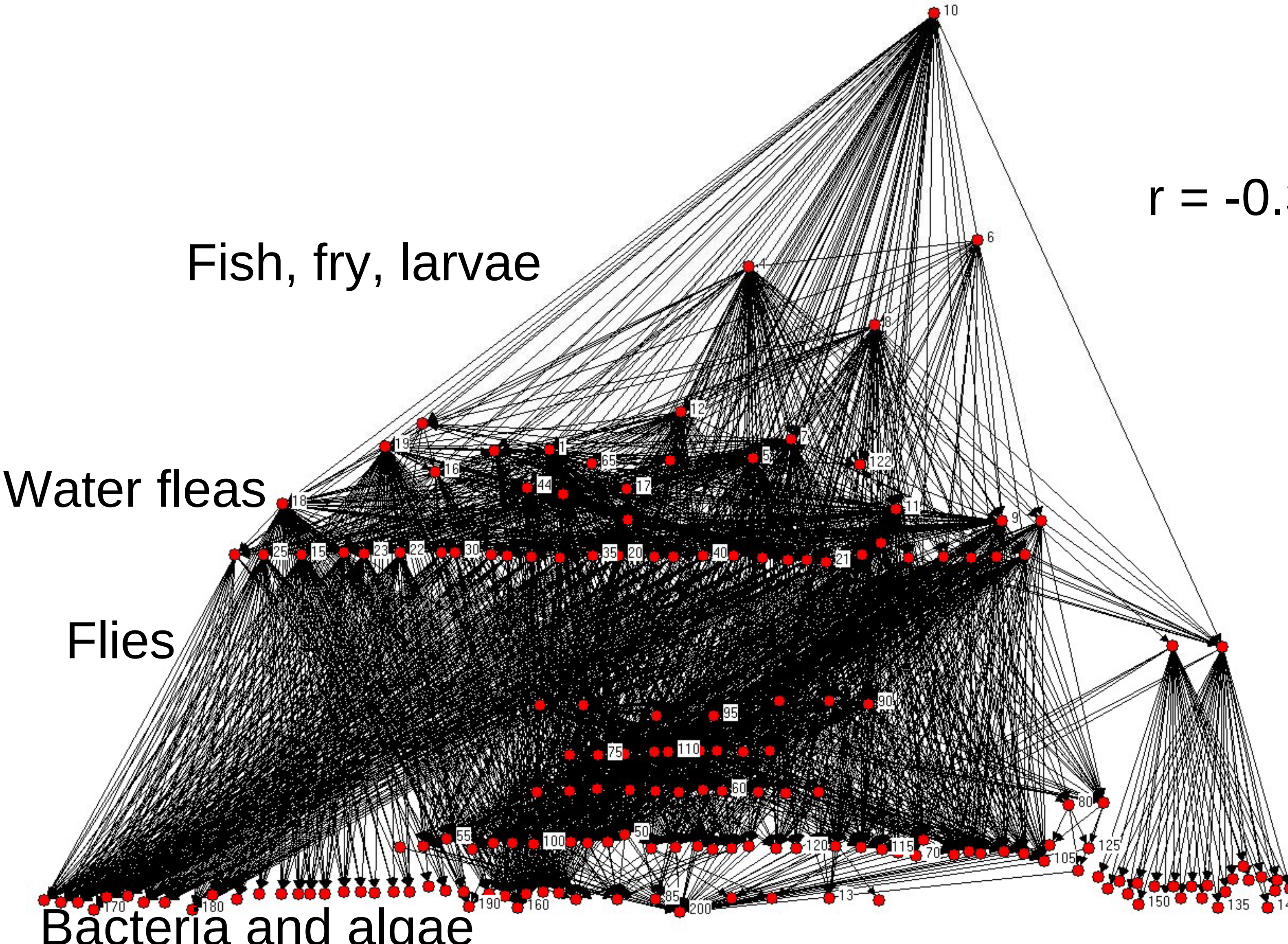



r = -0.2942

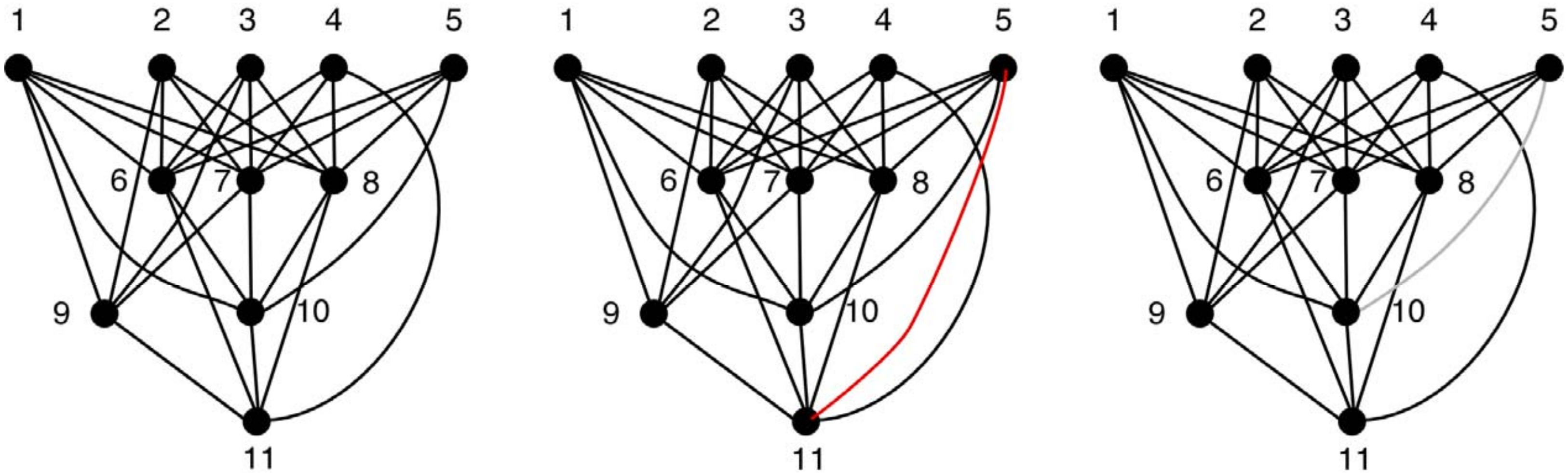
# Broom (Scotch Broom Grass)



# Little Rock Lake, WI



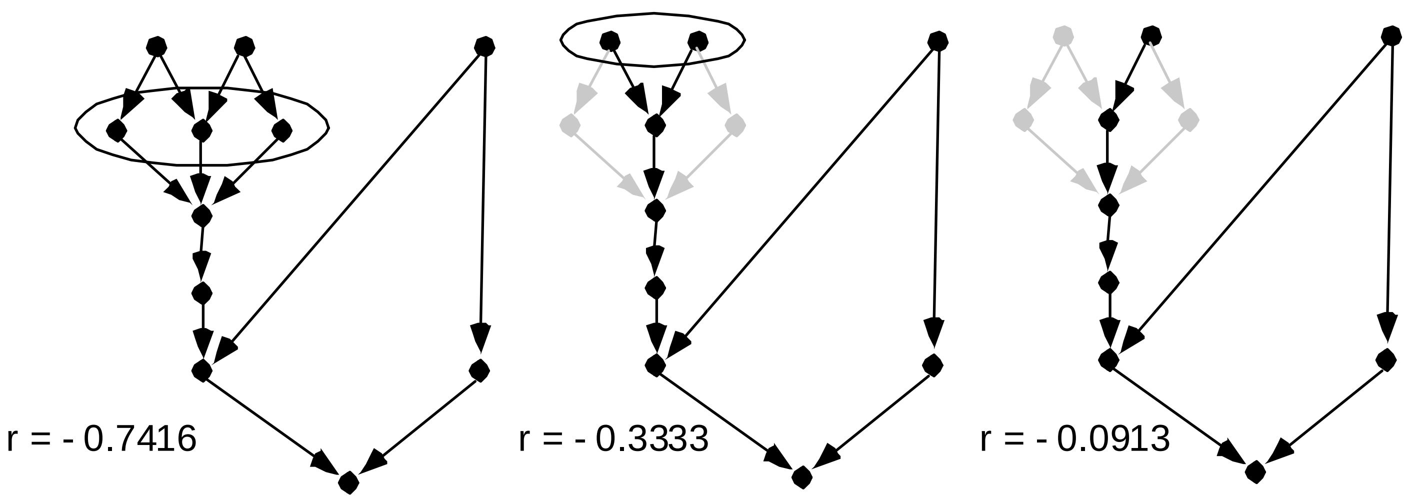
# r for Food Webs: Toy Example Seeking the Reason for ± r




r = 0/0 | r = - 0.448 | r = 0.1704

# Trophic Species Affect r

Trophic species = Trophic species are functional groups that contain organ- isms that appear to eat and be eaten by the exact same species within a food web.

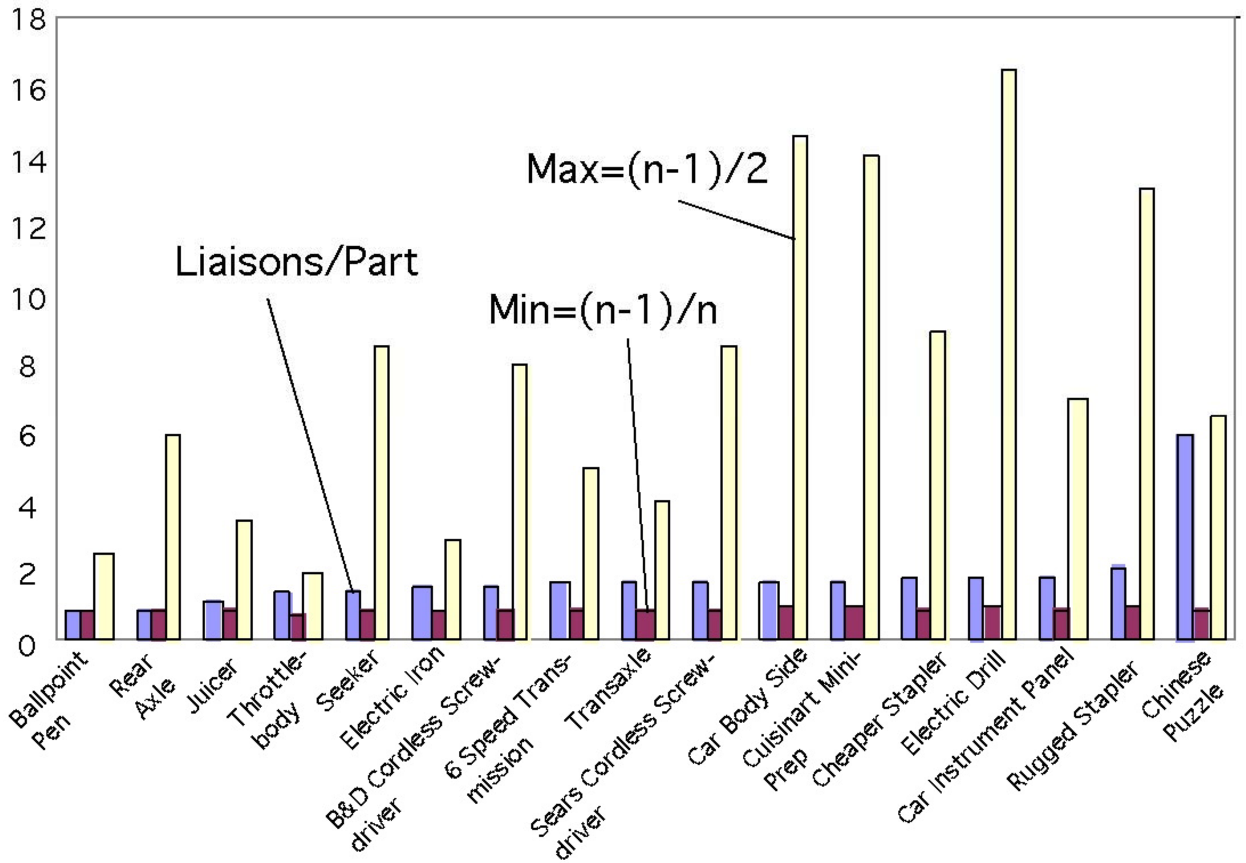



In graph theory this is called structural equivalence

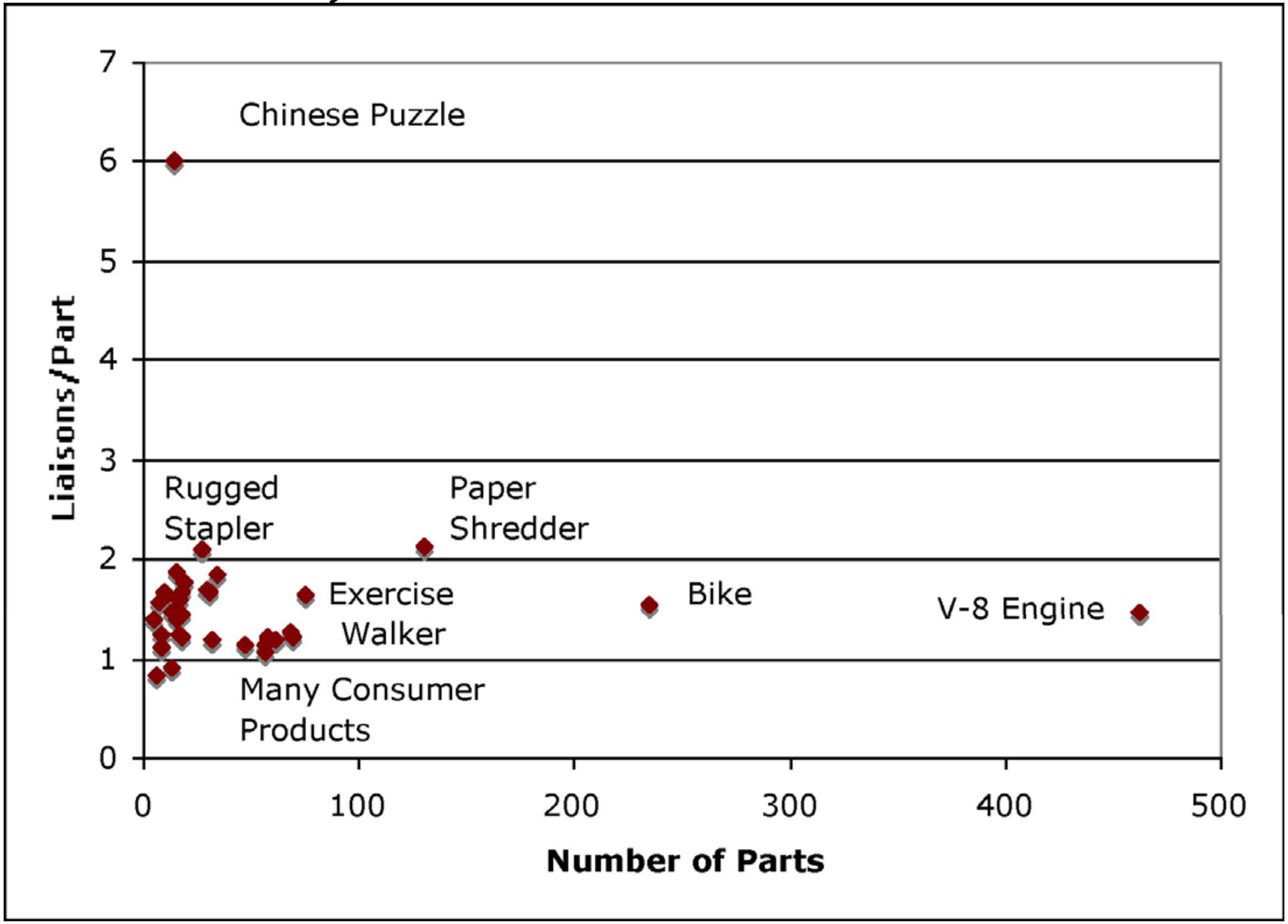
# Mechanical Assemblies

- All analyzed so far have $r < 0$.
- There is usually one or a few foundation parts that support important operating loads and have many parts attached to them
- Other parts group into subassemblies but these do not have high clustering coefficient or big differences in nodal degree
- The reason is avoidance of mechanical over-constraint

# Average Nodal Index for Mechanical Assemblies ≈ 1.5*2



# Average Nodal Index Does Not Grow with Network Size, Unlike Most Networks: Why?



# Constraint as a Limit on Connectivity

- Each time two parts are joined, each part subtracts some number of degrees of freedom from the other.
- An unconnected part has 6 dof
- Adding more parts may increase or decrease their mutual dof
- The chicken comes home to roost when a loop closes and the dof arithmetic has to be summed around the loop
- The result could be negative dof!
- This is called over-constraint
- The Kutzbach criterion formalizes this for simple assemblies
- Screw Theory is the definitive method

# Constraint and <k>

$$M = 3(n - g - 1) + \sum \text{joint freedoms } f_i$$

where

$n$ = number of parts

$g$ = number of joints

$f_i$ = degrees of freedom of joint $i$

$M > 0$: under-constrained

$M = 0$: properly (exactly) constrained

$M < 0$: over-constrained

$$\alpha = g/n$$

$$<k> = 2\alpha$$

$$\beta = \frac{\sum \text{joint freedoms } f_i}{\text{number of joints } g}$$

For $M = 0$:

$$\alpha = \frac{3 - 3n}{n(\beta - 3)} \rightarrow \frac{3}{3 - \beta} \text{ as } n \text{ gets large}$$

For spatial mechanisms, replace "3" with "6"

| $\beta$ | $\alpha$ **planar** | $\alpha$ **spatial** |
|---|---|---|
| 0 | 1 | 1 |
| 1 | 1.5 | 1.2 |
| 2 | 3 | 1.5 |

# What Matters in Assembly Networks

- It's chains and loops, not clusters
- Well, maybe it's a combination of chains, loops, and clusters
- Next few slides give a first pass at this
- Main functions are implemented by loops that appear to pass through major hubs
- For a class of assemblies:
  - Functions appear to apply loads
  - Hubs appear to be load-accumulators, balancers, distributors, shedders

# Example Assembly: Exercise Walker

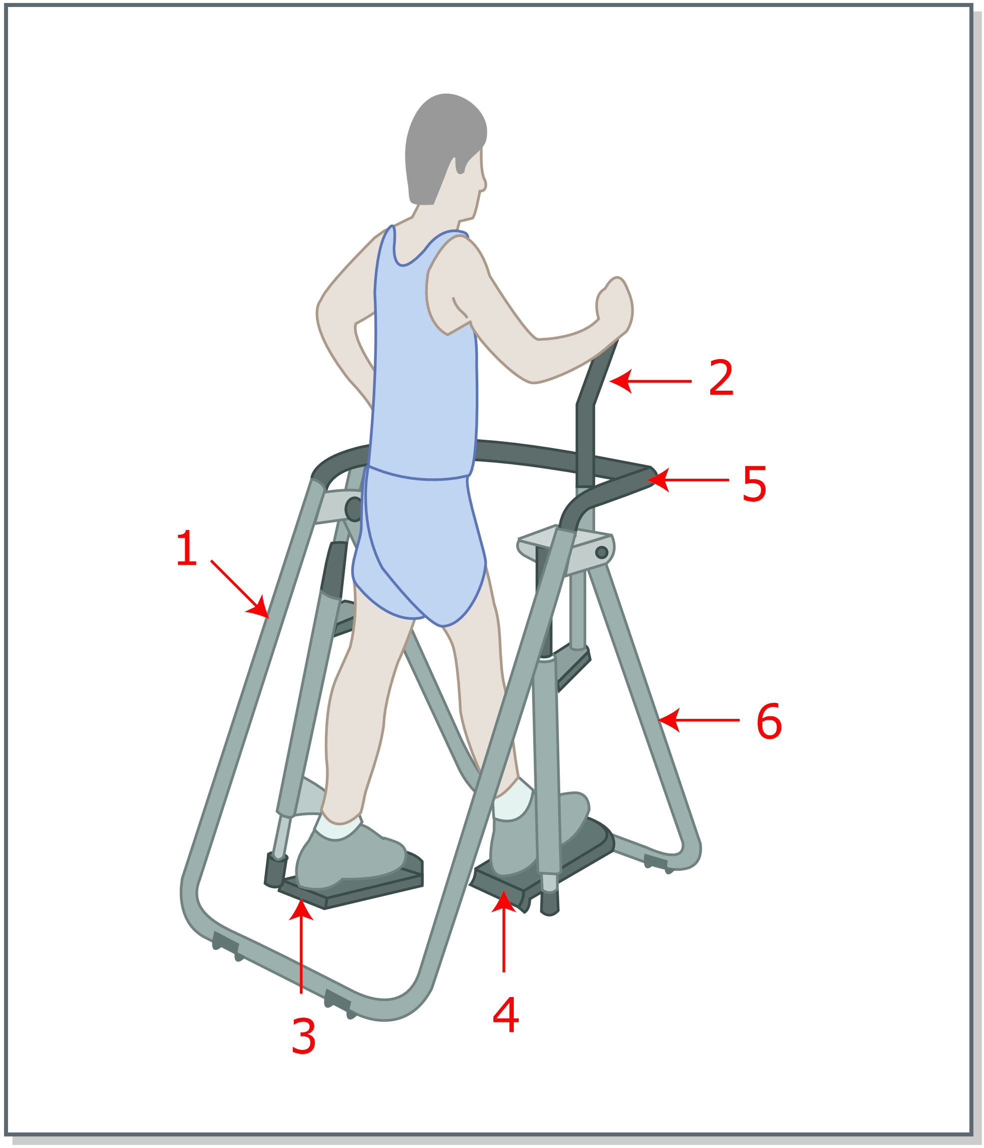
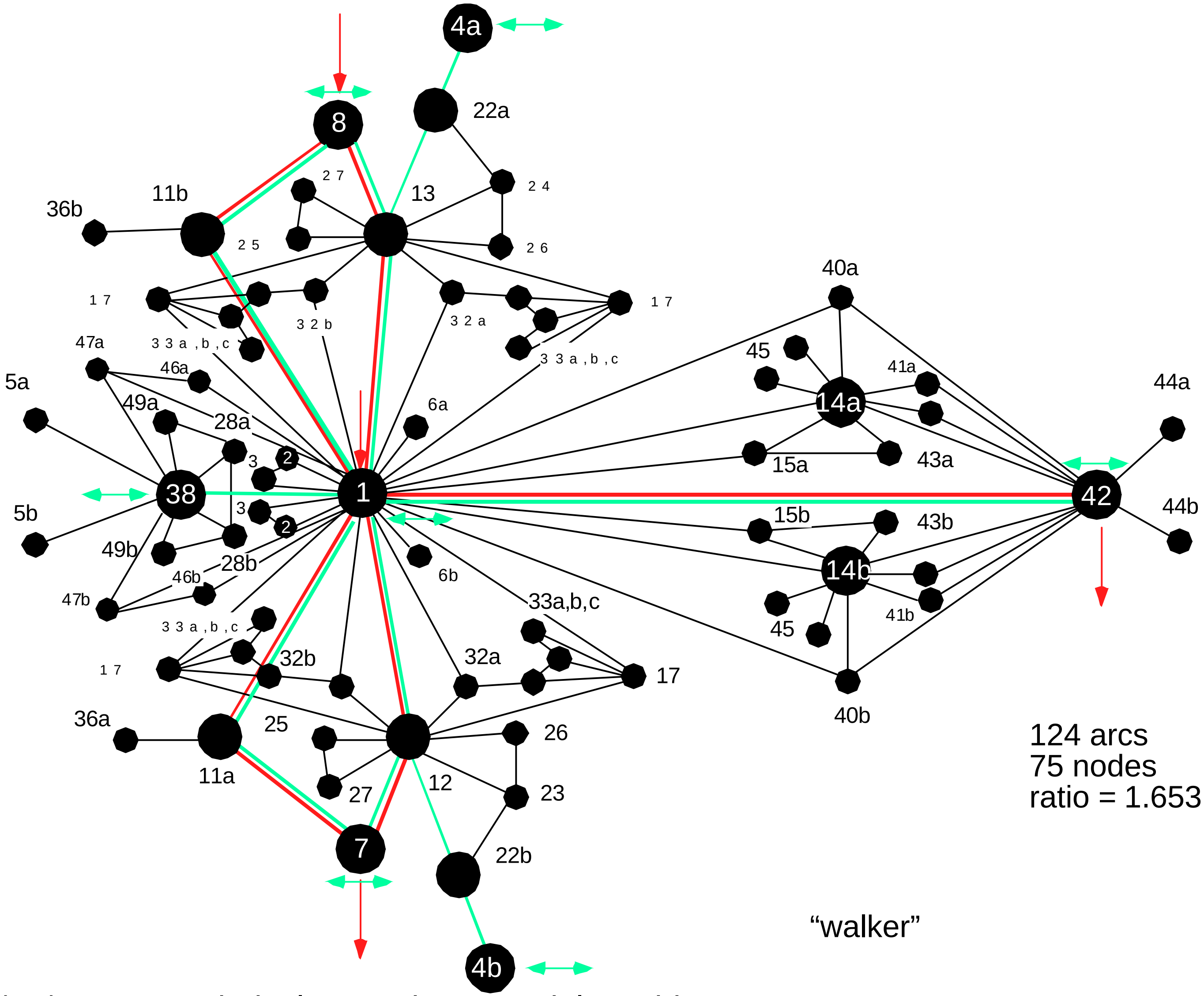



Figure by MIT OCW.

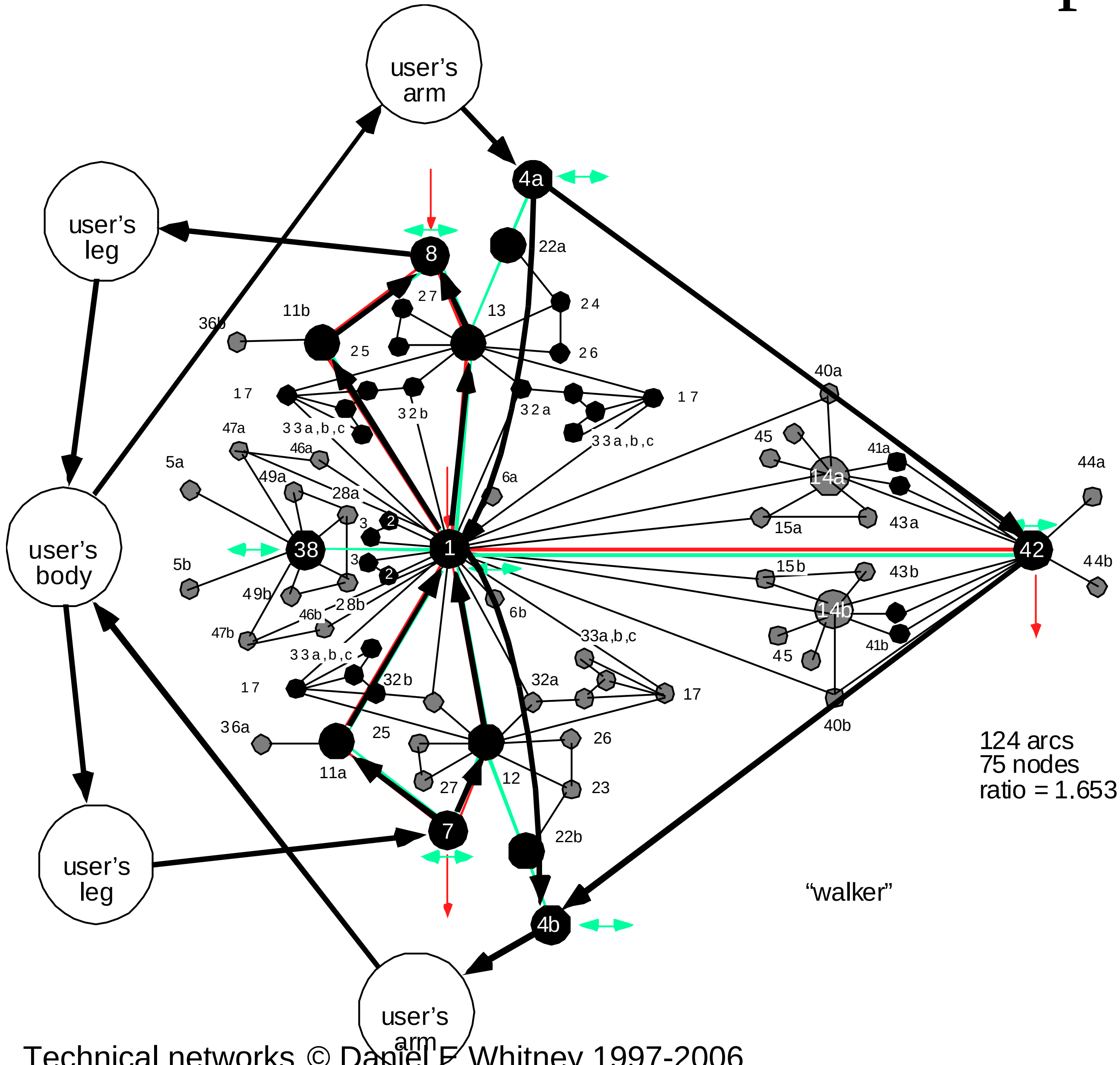
# Network Diagram of Walker



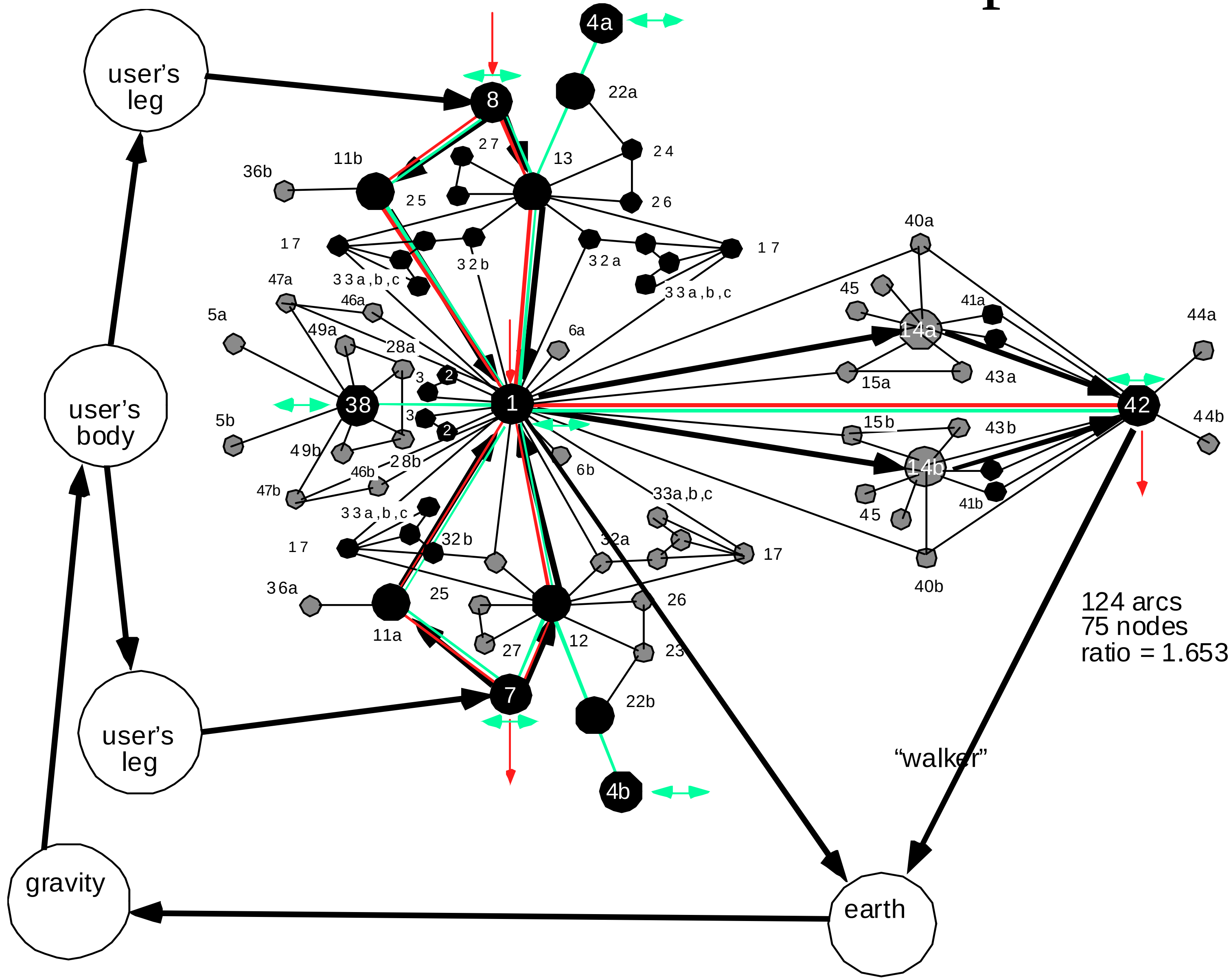

124 arcs
75 nodes
ratio = 1.653

“walker”

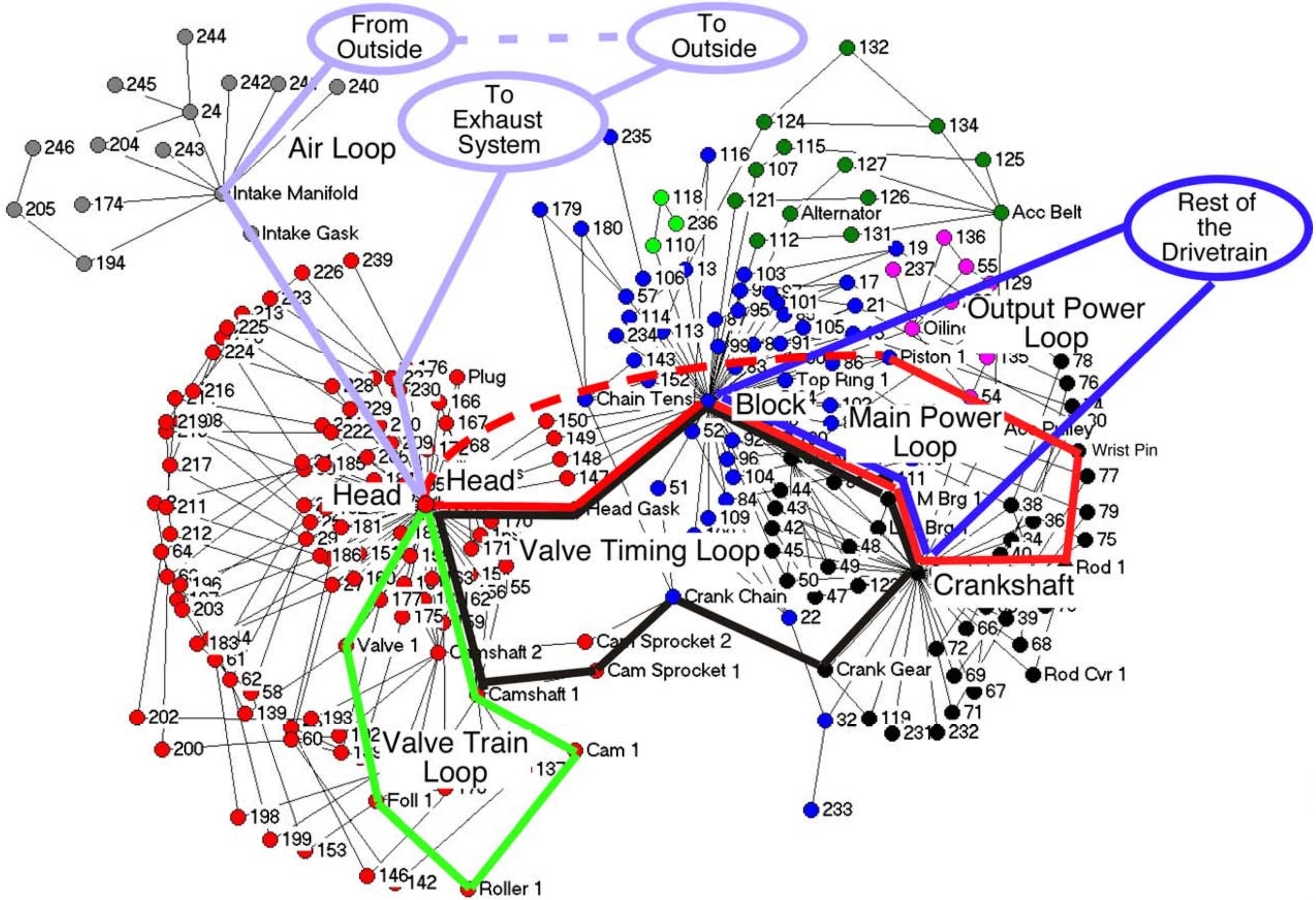
# Walker Horizontal Functional Loops



# Walker Vertical Functional Loops



# V-8 Engine Functional Loops



# Summary Properties of Several Big Networks (Newman)

Figure by MIT OCW.

| Network | Type | n | m | z | l | α | $C^{(1)}$ | $C^{(2)}$ | r |
|---|---|---|---|---|---|---|---|---|---|
| **SOCIAL** | | | | | | | | | |
| Film actors | undirected | 449913 | 25516482 | 113.43 | 3.48 | 2.3 | 0.20 | 0.78 | 0.208 |
| Company directors | undirected | 7673 | 55392 | 14.44 | 4.60 | - | 0.59 | 0.88 | 0.276 |
| Math coauthorship | undirected | 253339 | 496489 | 3.92 | 7.57 | - | 0.15 | 0.34 | 0.120 |
| Physics coauthorship | undirected | 52909 | 245300 | 9.27 | 6.19 | - | 0.45 | 0.56 | 0.363 |
| Biology coauthorship | undirected | 1520251 | 11803064 | 15.53 | 4.92 | - | 0.088 | 0.60 | 0.127 |
| Telephone call graph | undirected | 47000000 | 80000000 | 3.16 | | 2.1 | | | |
| E-mail messages | directed | 59912 | 86300 | 1.44 | 4.95 | 1.5/2.0 | | 0.16 | |
| E-mail address books | directed | 16881 | 57029 | 3.38 | 5.22 | - | 0.17 | 0.13 | 0.092 |
| Student relationships | undirected | 573 | 477 | 1.66 | 16.01 | - | 0.005 | 0.001 | -0.029 |
| Sexual contacts | undirected | 2810 | | | | 3.2 | | | |
| **INFORMATION** | | | | | | | | | |
| WWW nd.edu | directed | 269504 | 1497135 | 5.55 | 11.27 | 2.1/2.4 | 0.11 | 0.29 | -0.067 |
| WWW Altavista | directed | 203549046 | 2130000000 | 10.46 | 16.18 | 2.1/2.7 | | | |
| Citation network | directed | 783339 | 6716198 | 8.57 | | 3.0/- | | | |
| Roget's Thesaurus | directed | 1022 | 5103 | 4.99 | 4.87 | - | 0.13 | 0.15 | 0.157 |
| Word co-occurrence | undirected | 460902 | 17000000 | 70.13 | | 2.7 | | 0.44 | |
| **TECHNOLOGICAL** | | | | | | | | | |
| Internet | undirected | 10697 | 31992 | 5.98 | 3.31 | 2.5 | 0.035 | 0.39 | -0.189 |
| Power grid | undirected | 4941 | 6594 | 2.67 | 18.99 | - | 0.10 | 0.080 | -0.003 |
| Train routes | undirected | 587 | 19603 | 66.79 | 2.16 | - | | 0.69 | -0.033 |
| Software packages | directed | 1439 | 1723 | 1.20 | 2.42 | 1.6/1.4 | 0.070 | 0.082 | -0.016 |
| Software classes | directed | 1377 | 2213 | 1.61 | 1.51 | - | 0.033 | 0.012 | -0.119 |
| Electronic circuits | undirected | 24097 | 53248 | 4.34 | 11.05 | 3.0 | 0.010 | 0.030 | -0.154 |
| Peer-to-peer network | undirected | 880 | 1296 | 1.47 | 4.28 | 2.1 | 0.012 | 0.011 | -0.366 |
| **BIOLOGICAL** | | | | | | | | | |
| Metabolic network | undirected | 765 | 3686 | 9.64 | 2.56 | 2.2 | 0.090 | 0.67 | -0.240 |
| Protein interactions | undirected | 2115 | 2240 | 2.12 | 6.80 | 2.4 | 0.072 | 0.071 | -0.156 |
| Marine food web | directed | 135 | 598 | 4.43 | 2.05 | - | 0.16 | 0.23 | -0.263 |
| Freshwater food web | directed | 92 | 997 | 10.84 | 1.90 | - | 0.20 | 0.087 | -0.326 |
| Neural network | directed | 307 | 2359 | 7.68 | 3.97 | - | 0.18 | 0.28 | -0.226 |

Basic statistics for a number of published networks. The properties measured are: type of graph, directed or undirected; total number of vertices n; total number of edges m; mean degree z; mean vertex-vertex distance l; exponent α of degree distribution if the distribution follows a power law (or "-" if not; in/out-degree exponents are given for directed graphs); clustering coefficient $C^{(1)}$; clustering coefficient $C^{(2)}$; and degree correlation coefficient r. Blank entries indicate unavailable data.

# Additional Networks

Social
Biological
Assemblies (Technological)
Transport (Technological)
Software
Electric (Technological)

| Network | n | m | <k> | r |
|---|---|---|---|---|
| Karate Club | 34 | 78 | 4.5882 | -0.4756 |
| Little Rock Lake Food Web | 200 | 1159 | 11.59 | -0.3398 |
| Santa Fe Coauthors | 118 | 198 | 3.3559 | -0.2916 |
| V8 engine | 243 | 367 | 3.01 | -0.269 |
| Exercise walker | 82 | 116 | 2.8293 | -0.256 |
| Bike | 131 | 208 | 3.1756 | -0.2018 |
| Six speed automatic transmission | 143 | 244 | 3.4126 | -0.1833 |
| "HOT" | 1000 | 1049 | 2.098 | -0.1707 |
| St. Martin Food Web | 44 | 132 | 6 | -0.1097 |
| Tokyo Regional Rail | 147 | 204 | 2.775 | -0.0911 |
| Mozilla19980331 | 811 | 4077 | 5.0271 | -0.0499 |
| Mozilla all comp | 1187 | 4129 | 3.4785 | -0.0393 |
| Munich Schnellbahn | 50 | 65 | 2.6 | -0.0317 |
| Western Power Grid | 4941 | 6594 | 2.6691 | 0.0035 |
| Physics coauthors | 145 | 346 | 4.7724 | 0.0159 |
| Tokyo Regional Rail plus Subways | 191 | 300 | 3.1414 | 0.0425 |
| London Underground | 92 | 139 | 3.02 | 0.0997 |
| Moscow Subways | 51 | 82 | 3.216 | 0.1846 |
| Company directors | 6731 | 50775 | 15.09 | 0.2386 |
| Moscow Subways and Regional Rail | 129 | 204 | 3 | 0.2601 |
| Scottish Broom Grass Food Web | 154 | 185 | 2.406 | 0.2618 |

Definitely negative

Probably Indifferent From zero

Definitely positive