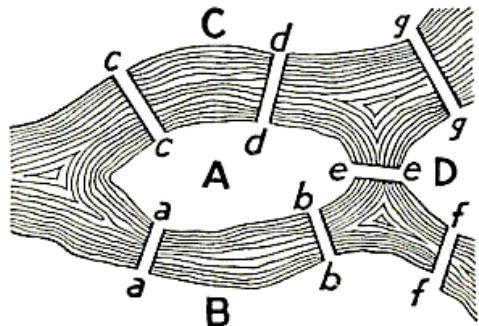
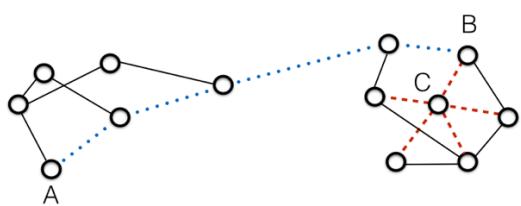


NETWORK BASICS



ABSTRACT ORGANIZATION

- In systems/mechanisms in the real world
 - Entities have distinctive properties
 - Relations between entities have distinctive properties
- Network representations abstract from these properties
 - Entities are represented as nodes
 - Relations as edges
 - undirected or directed

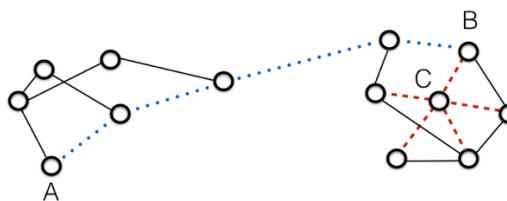


OUTLINE

1. Network measures
2. Small-world and scale-free networks
3. Connectomes
4. Motifs

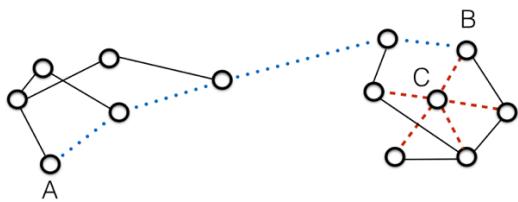
NETWORK MEASURES

- Path length measures
 - Shortest path length between two nodes is the minimum number of links that need to be followed to traverse from one to the other
 - Efficiency: Inversely related to shortest path length
 - Diameter: longest shortest path length
 - Average or characteristic path length



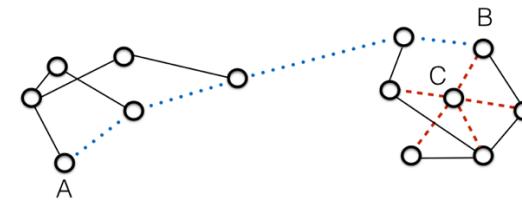
NETWORK MEASURES

- Clustering measures
 - Clustering coefficient of a node: proportion of the possible linkages between nearest neighbors that are realized
 - Clustering coefficient of a graph: average of clustering coefficients of nodes
 - Clusters or communities: subgraphs whose nodes are highly interconnected



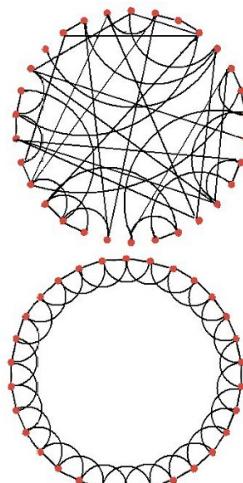
NETWORK MEASURES

- Connectivity Measures:
 - Degree or connectivity of a node k : the number of links between the node and other nodes
 - Degree distribution $P(k)$: probability that a randomly chosen node has degree k
 - Assortativity: correlation between degrees of connected nodes. Positive value indicates high-degree nodes tend to connect with each other
- Betweenness Measures
 - Centrality: Percentage of shortest paths going through a node (or link)



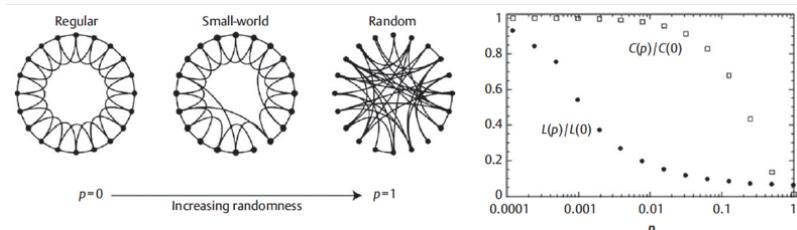
CLASSICAL FOCI: LATTICES AND RANDOM GRAPHS

- The 20th century pioneers of graph theory focused primarily on totally connected networks, random networks or regular lattices
- Characteristic path length (L): mean of the shortest paths (d) between all pairs of nodes
 - In a totally connected graph $L=1$
 - Regular lattice: L is very large
 - Random graph: L can be very small
- Clustering Coefficient (C)
 - For regular lattice: C is very high
 - For random graph: C is very low
- Apparent tradeoff between keeping path length short and maintaining high clustering



SMALL-WORLDS: BETWEEN LATTICES AND RANDOM GRAPHS

- Watts and Strogatz (1998) start with a regular lattice and gradually replace a local connection with a longer-distance one
 - Very quickly the characteristic path length drops to close to that for a random graph
 - The clustering coefficient remains high



SMALL WORLDS

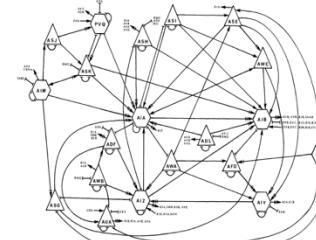
- Watts and Strogatz called these networks in the middle range *small-worlds*
- Showed that they occur widely in real world networks

Table 1 Empirical examples of small-world networks

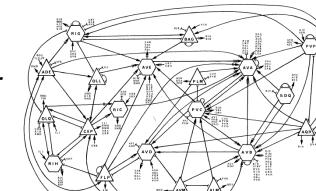
	L_{actual}	L_{random}	C_{actual}	C_{random}
Film actors	3.65	2.99	0.79	0.00027
Power grid	18.7	12.4	0.080	0.005
<i>C. elegans</i>	2.65	2.25	0.28	0.05

E. ELEGANS NEURAL NETWORK

- Based on serial electron micrographs, White et al. (1986) reconstructed the entire neural network of the hermaphrodite nematode worm *C. elegans*
 - 302 neurons
 - approx. 5000 chemical synapses
 - approx. 600 gap junctions
 - approx. 2000 neuromuscular junctions
- Watts and Strogatz showed that the *C. elegans* network constitute a small-world



• Chemoreception circuit



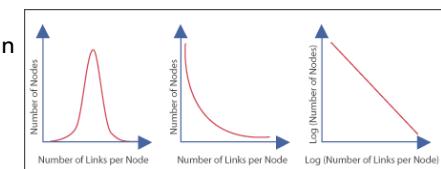
Other sensory neurons

THE FUNCTIONALITY OF SMALL-WORLDS

- "Models of dynamical systems with small-world coupling display enhanced **signal-propagation speed, computational power, and synchronizability**. In particular, infectious diseases spread more easily in small-world networks than in regular lattices." (Watts and Strogatz)

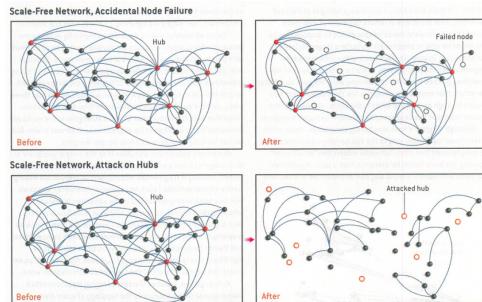
NODE DEGREE: FROM GAUSSIAN TO POWER LAW DISTRIBUTIONS

- Plausible assumption: node degree is distributed in a Gaussian fashion
 - Exponential because "the probability that a node is connected to k other sites decreases exponentially for large k ."
- Surveying webpages at U. of Notre Dame, Barabási and his colleagues were surprised to find a few nodes that had far higher degree than expected
 - Rather than exponential, they fit a power law: probability of the degree of a node was $1/k^n$
 - No peak to the distribution
 - When plotted on log-log coordinates it forms a straight line



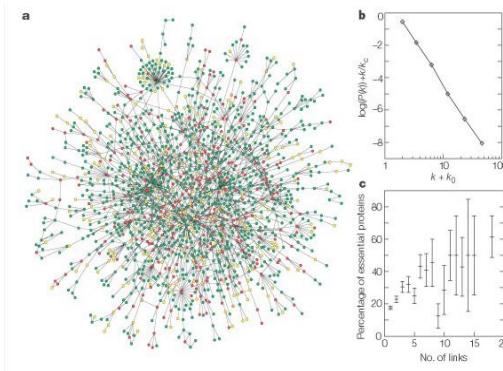
ROBUSTNESS AND VULNERABILITY

- Nodes with few connections can often be lost without compromising the integrity of the network
 - "as many as 80 percent of randomly selected Internet routers can fail and the remaining ones will still form a compact cluster in which there will still be a path between any two nodes. It is equally difficult to disrupt a cell's protein-interaction network: our measurements indicate that even after a high level of random mutations are introduced, the unaffected proteins will continue to work together."
- But loss of just a few highly-connected nodes (hubs) can be catastrophic
 - Networks fractionate



PROTEIN NETWORKS

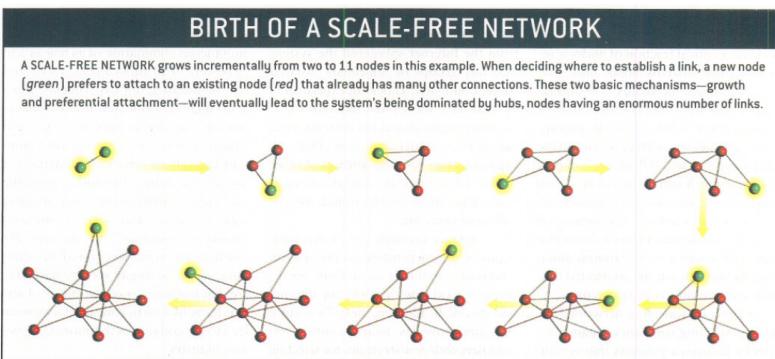
- In yeast protein networks (edges represent interactions), knocking out nodes shown in red results in death, those in orange in serious loss of function.
- Others result in far less harm



H. Jeong, et al., Nature, Vol 411, p41 (2001)

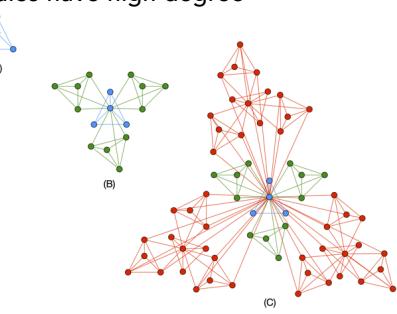
THE ORIGINS OF SCALE-FREE NETWORKS

- The rich get richer scheme—growth plus preferential attachment
 - As new nodes join a network, they are more likely to connect to those nodes that are already highly connected



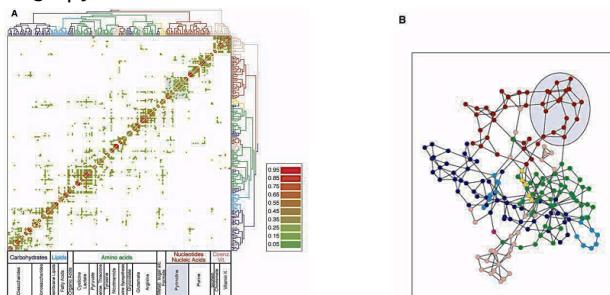
MODULAR HIERARCHICAL NETWORKS: HIGH CLUSTERING AND SCALE-FREE

- Seeing a tension between high clustering and being scale-free, Ravasz et al. (2002) explored how to construct networks that exhibit both features
 - Strategy of duplicating modules but linking them to a common node
 - Nodes at the center of modules have high degree
 - Especially true of the node at the center of the whole network
 - Nodes in the periphery exhibit high clustering



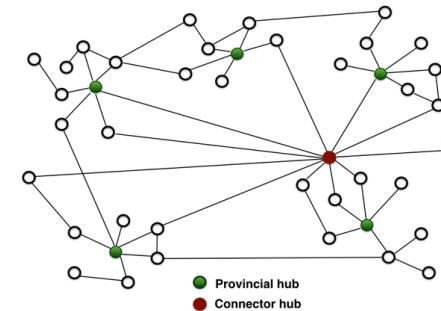
MODULAR HIERARCHICAL NETWORKS: HIGH CLUSTERING AND SCALE-FREE

- To identify modules in the metabolic network of *E. coli*, Ravasz et al. develop an overlap matrix
- Indicating the percentage of additional nodes to which two connected nodes share connections
 - Modules correspond to known metabolic grouping
 - e.g., pyrimidine in dark oval



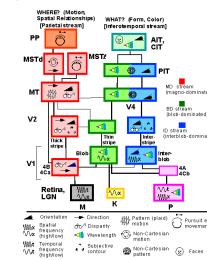
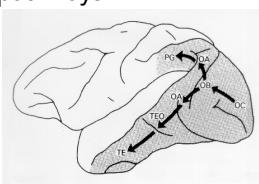
HUBS

- Nodes with particularly high degree are labeled *hubs*
- Some hubs form the core of clusters or modules: *provincial hubs*
 - These hubs have most connections to other nodes in the module
- Some hubs connect the larger network but are not members of a common cluster: *connector hubs*



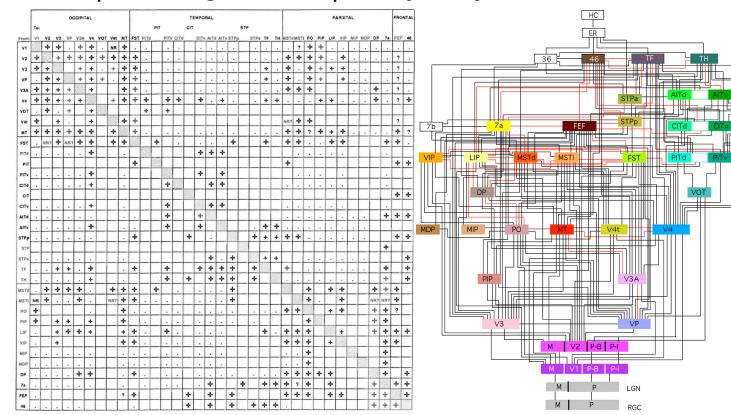
FROM BRAIN PATHWAYS TO NETWORKS: CONNECTOMES

- The investigation of brain processes involved in tasks like visual perception began by identifying early processing areas and what stimuli they responded to
 - Retina and LGN: center-surround contrast
 - V1: edges
- Then proceeded forward in the brain to identify specialized processing areas
 - Characterized as lying on one of two pathways



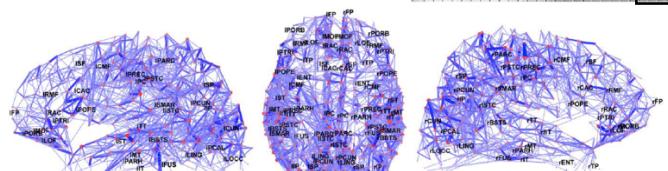
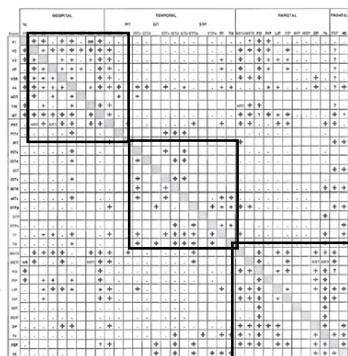
FROM BRAIN PATHWAYS TO NETWORKS: CONNECTOMES

- Already in 1991 Felleman and van Essen had advanced a more complex picture by identifying for over 30 brain regions involved in visual processing and multiple ways they were connected



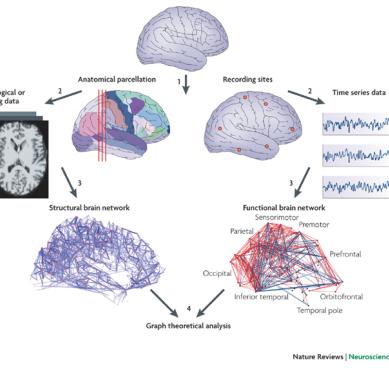
THE CONNECTOME PROJECT

- Sporns, Tononi, and Kötter (2005) introduced the term **connectome** for the connection matrix of the human brain
 - Subsequently extended to many other species (*C. elegans*, fruit fly, etc.)
- Goal is not just to identify the network but determine its properties



Linking Structural and Functional Networks

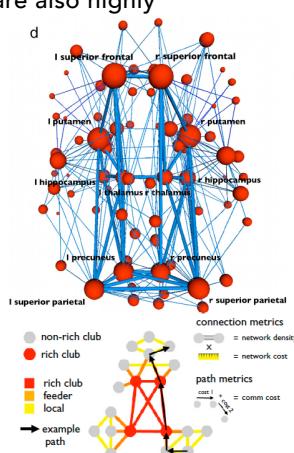
- The connectome was originally based solely on structure: neurons or brain regions and the connections between them
- Although functional neural imaging initially focused on identifying specific brain regions, more recently it has emphasized networks involved in particular types of tasks
 - With the discovery of low-frequency oscillations (< .1 Hz) detectable in fMRI, researchers have identified networks of brain regions
 - One composed of areas more active in the resting state
 - Others composed of regions more active in types of tasks
- Functional and structural networks resemble one another



Bullmore and Sporns, 2009

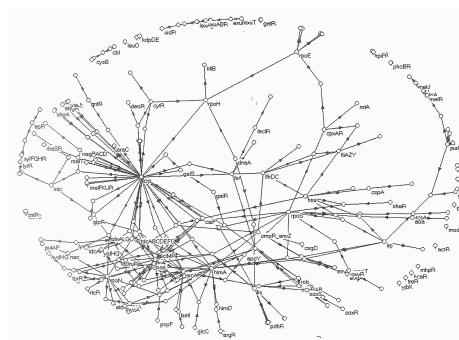
RICH HUBS

- van den Heuvel and Sporns (2011) determined that twelve of the most important hubs in the human brain are also highly interconnected with each other
 - superior parietal area, precuneus, superior frontal cortex, putamen, hippocampus, and thalamus in both hemispheres
- In computational models, van den Heuvel and Sporns showed that disruption of nodes in the rich hub severely impaired global communication
 - proposed that damage to rich hub nodes may explain generalized cognitive deficits in diseases such as Alzheimer's and schizophrenia



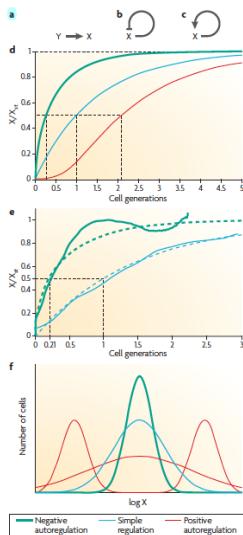
DISCOVERY OF MOTIFS

- Uri Alon and collaborators identified subgraphs (consisting of a small number of nodes) that occur repeatedly within networks such as the transcription regulation networks in *E. coli*
 - They dubbed those subgraphs that occur far more frequently than would be expected by chance **motifs**
 - Raises important question about what counts as chance
- Major claims for motifs:
 - They are building blocks of larger networks
 - They carry out specific information processing functions



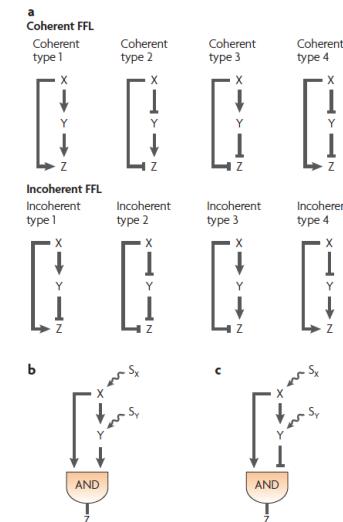
SIMPLE REGULATION

- Alon begins his analysis with single directed edges either between two nodes or from one node back onto itself
- In a transcription network, a single directed edge between two nodes will allow the first to increase the transcription of the second until it reaches steady-state
- Negative feedback allow for a faster raise in the concentration of the transcript but then leveling off due to the feedback. It can also reduce variability
- Positive feedback can yield slower rise followed by an inflection (s-shaped curve) and greater variability



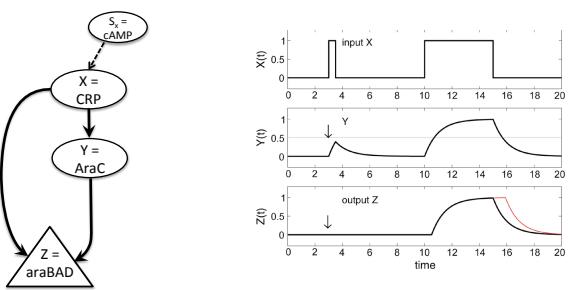
FEED-FORWARD LOOPS

- Three node subgraphs containing a direct route from X to Z and indirect route through Y
- Either route can produce activation or repression
- Coherent if both routes have the same sign
- Incoherent if opposite sign
- Treats Z as executing a Boolean function on its inputs
- Either an AND or an OR gate
- In *E. coli* and yeast transcription networks two of these appear more frequently than by chance



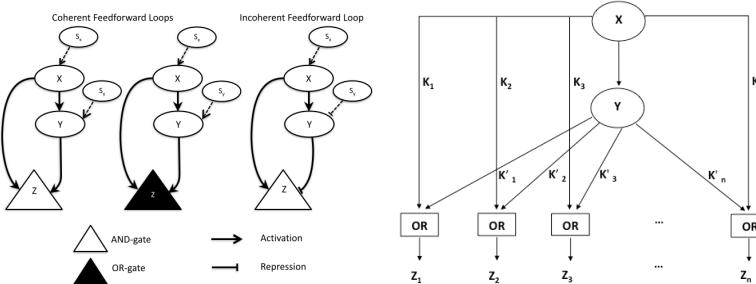
PERSISTENCE DETECTOR

- Sign-sensitive delay element: input X must persist long enough for Y to be expressed in a sufficient amount before transcription of Z begins
- Prevents transcribing Z in response to random perturbations
- Also stops transcribing Z promptly when X is inactivated



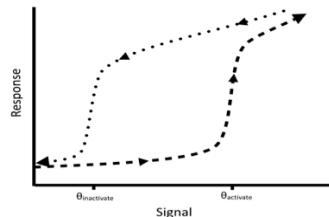
OTHER FEED-FORWARD LOOPS

- When the AND-gate is replaced by an OR-gate in a coherent feedforward loop, transcription of Z starts immediately when X is activated, but continues in the face of short disruptions due to Y
- Multiple outputs with time delays enables sequential synthesis (e.g., of a flagellum in bacteria)
- With an incoherent loop, X will initially cause transcription of Z, but as Y builds up, transcription will stop, resulting in a pulse
- enables rapid generation of Z without the risk of generating too much



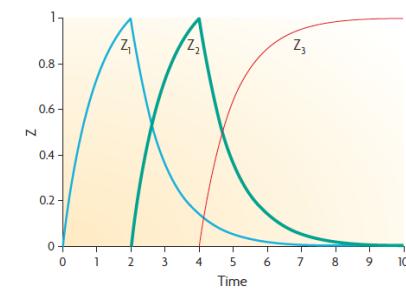
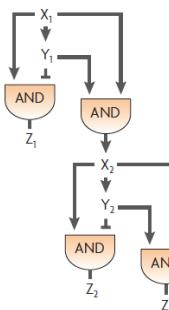
BI-STABLE SWITCH

- With appropriate parameters, a double negative feedback loop generates an especially useful property
 - With increased input to one unit, it will switch off the other unit
 - But the input to that unit must drop to a much lower value for the that unit to become active again turn it off
 - The switch will remain stable in either position, not easily reverted to the other
 - Useful in context in which it is important to maintain the sequential order of processes



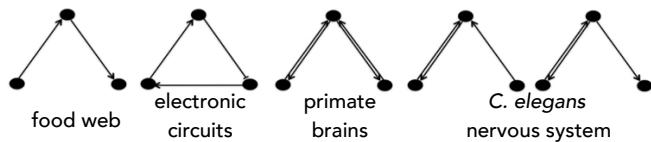
COMBINING MOTIFS

- Network that governs spore development in *Bacillus subtilis*
 - A process that bacteria undertake only in desperate circumstances
 - The outputs Z_1 , Z_2 , and Z_3 each represent large numbers of genes
 - The network combines two incoherent FFLs that each produce pulses and two coherent FFLs that generates a steady output



DIFFERENT MOTIFS IN DIFFERENT NETWORKS

- In *E. coli* transcription network only variants of the feedforward loop qualify as motifs (>10 SD above freq. in random networks)
- In other networks, other subgraphs meet the condition for being a motif
 - Alon is inclined to see motifs as adaptations—specifically selected for the roles they play in specific networks
 - This has elicited considerable controversy
 - One can, however, analyze the information processing role motifs play independently of the question of their origin



MORE COMPLEX DYNAMICS

- Negative feedback often generates oscillations, especially when it is coupled with positive feedback
- These can be useful—here used to generate circadian rhythms

