Lecture 3: Aggregate network properties

Noshir Contractor

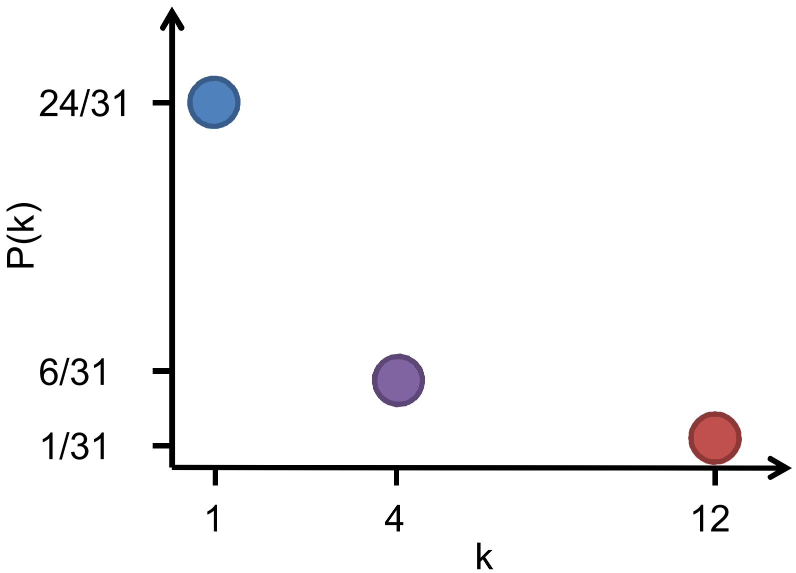
Jane S. & William J. White Professor of Behavioral Sciences

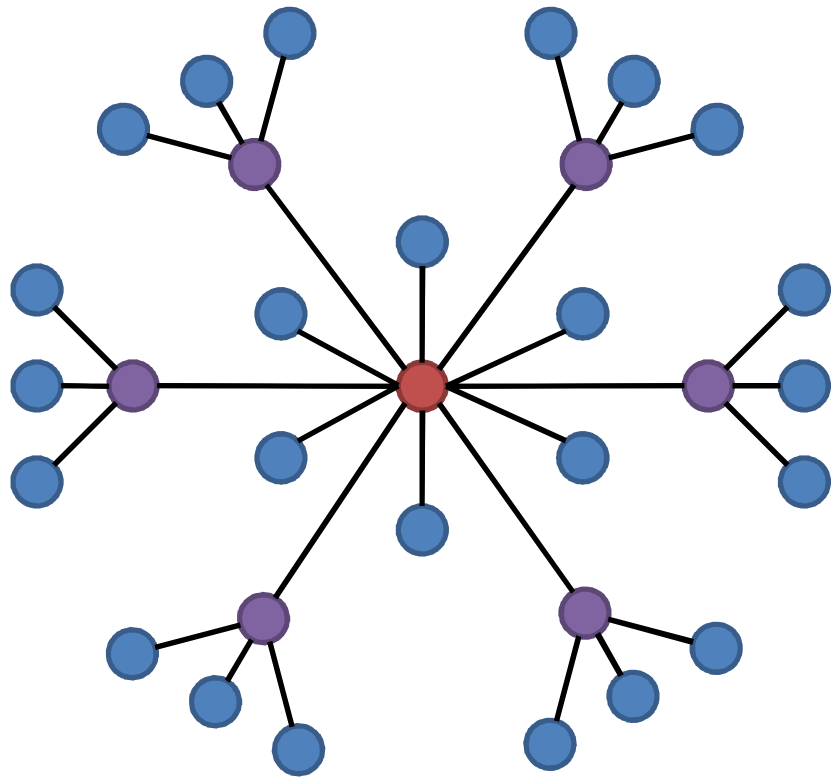
Professor of Ind. Eng. & Mgmt. Sciences, McCormick School of Engineering
Professor of Communication Studies, School of Communication &
Professor of Management & Organizations, Kellogg School of Management,
Director, Science of Networks in Communities (SONIC) Research Laboratory
nosh@northwestern.edu

NORTHWESTERN UNIVERSITY

Degree distributions

- What's the probability *P*(*k*) of randomly selecting a node with degree *k* in this network?



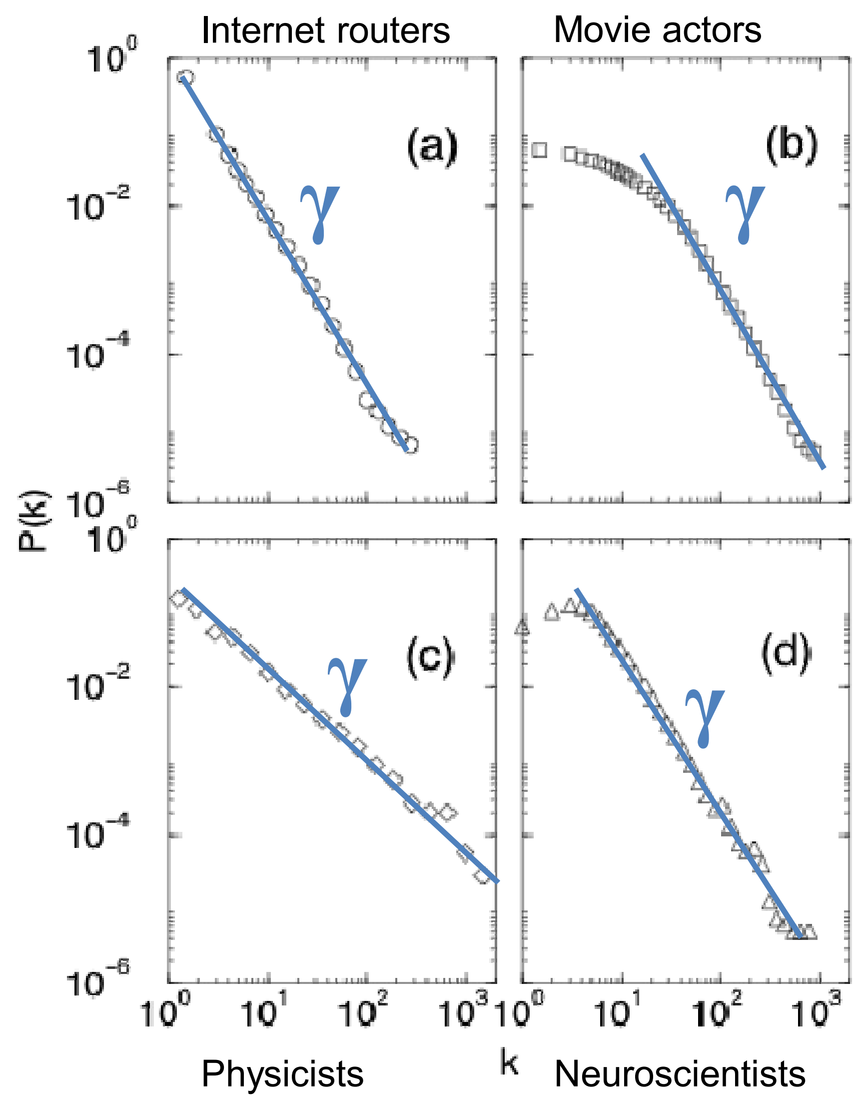


Power laws

- Large networks can have degree distributions that span several orders of magnitude
- Many real world networks follow a **power law degree distribution**
  - Scale free networks, 80/20 rule, Pareto principle, Zipf's Law, long tail, etc.

$$P(k) \sim k^{-\gamma}$$

- How do you generate scale free networks?
  - Check back in week 7!

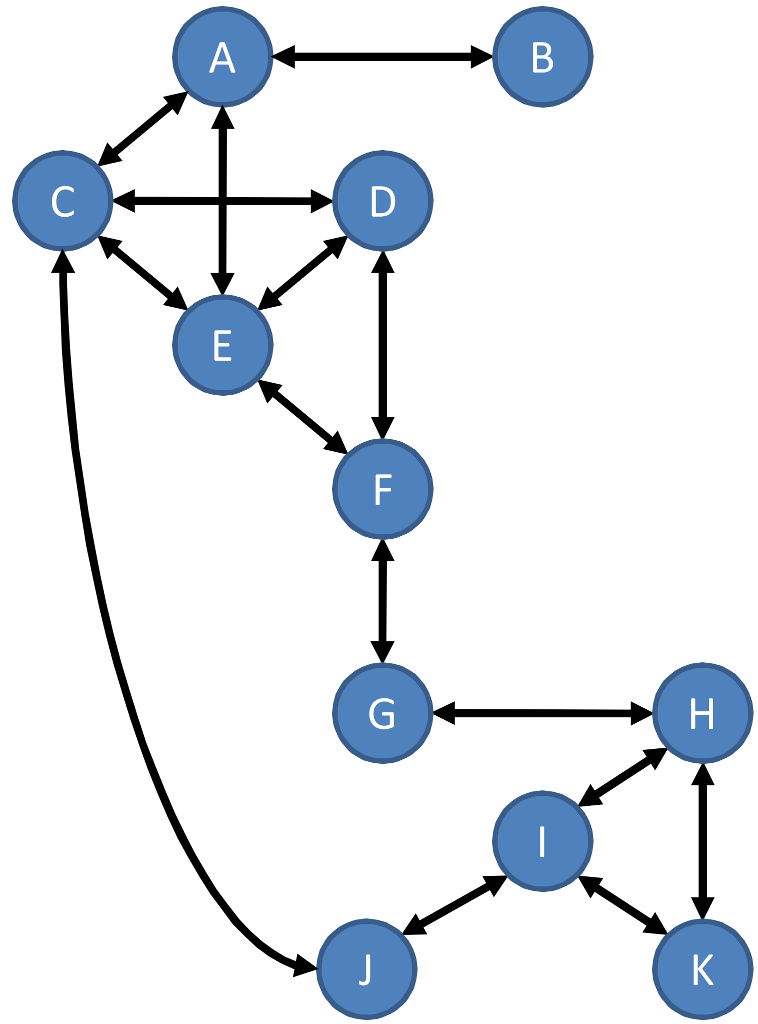


Deg distributions across networks

Network	Size	$\langle k \rangle$	κ	ℓ_{real}	ℓ_{rand}	ℓ_{pow}	Reference
WWW	$325,729$	4.51	900	11.2	8.32	4.77	Albert, Jeong, Barabási 1999
WWW	4×10^7	7					Kumar *et al.* 1999
WWW	2×10^8	7.5	$4,000$	16	8.85	7.61	Broder *et al.* 2000
WWW, site	$260,000$						Huberman, Adamic 2000
Internet, domain*	$3,015$ - $4,389$	3.42 - 3.76	$30 - 40$	4	6.3	5.2	Faloutsos 1999
Internet, router*	$3,888$	2.57	30	12.15	8.75	7.67	Faloutsos 1999
Internet, router*	$150,000$	2.66	60	11	12.8	7.47	Govindan 2000
Movie actors*	$212,250$	28.78	900	4.54	3.65	4.01	Barabási, Albert 1999
Coauthors, SPIRES*	$56,627$	173	$1,100$	4	2.12	1.95	Newman 2001b,c
Coauthors, neuro.*	$209,293$	11.54	400	6	5.01	3.86	Barabási *et al.* 2001
Coauthors, math*	$70,975$	3.9	120	9.5	8.2	6.53	Barabási *et al.* 2001
Sexual contacts*	2810						Liljeros *et al.* 2001
Metabolic, E. coli	778	7.4	110	3.2	3.32	2.89	Jeong *et al.* 2000
Protein, S. cerev.*	1870	2.39					Mason *et al.* 2000
Ythan estuary*	134	8.7	35	2.43	2.26	1.71	Montoya, Solé 2000
Silwood park*	154	4.75	27	3.4	3.23	2	Montoya, Solé 2000
Citation	$783,339$	8.57					Redner 1998
Phone-call	53×10^6	3.16					Aiello *et al.* 2000
Words, cooccurence*	$460,902$	70.13					Cancho, Solé 2001
Words, synonyms*	$22,311$	13.48					Yook *et al.* 2001

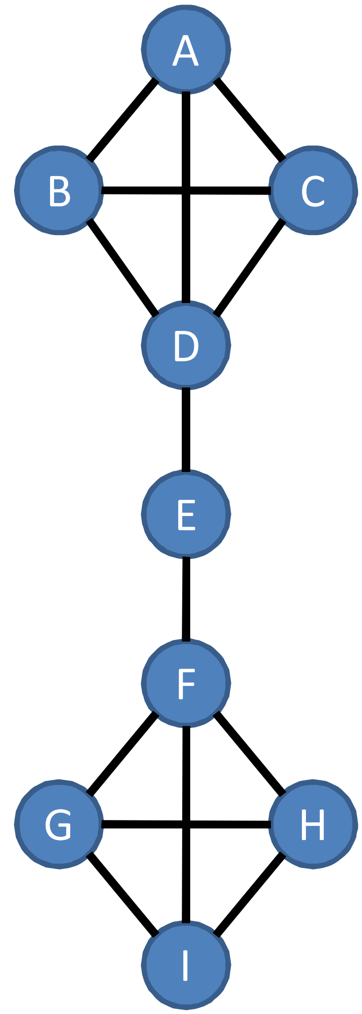
Path length

- **Path length**: number of links between two nodes (degrees of separation)
  - BACDE = 4
- **Geodesic**: Shortest path length between two nodes
  - BAE = 2
- **Eccentricity**: Each actor's longest geodesic
- **Diameter**: Network's largest geodesic\eccenctricity
  - BAEFGH\BACJIH
- **Shortcut**: paths that bypass clusters
  - CJ saves traveling across 6 links



Density, clustering, centralization

- **Density**
  - Observed edges in network / maximum possible edges
- **Clustering**
  - Count ties among alters, removing ego and ties to ego
  - Observed ties in actor's ego network / maximum possible ties in ego network
- **Network centralization**
  - Variation in individual actors' centralities
  - High centralization when few actors possess higher centrality than average
  - Low centralization when actors all have similar centralities

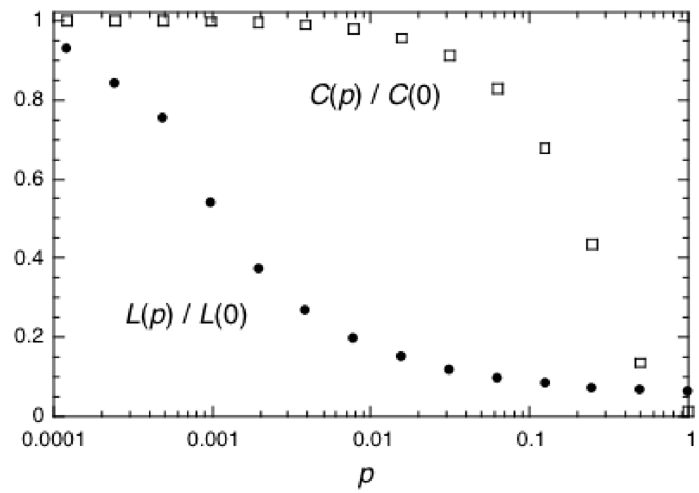


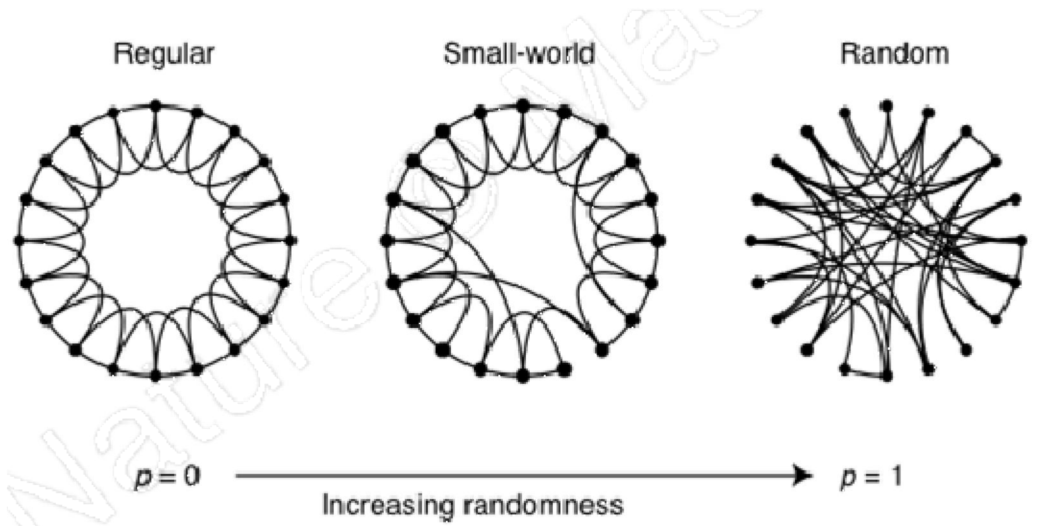
Paths & clustering across networks

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	Reference
WWW, site level, undir.	153,127	35.21	3.1	3.35	Adamic 1999
Internet, domain level	3015 - 6209	3.52 - 4.11	3.7 - 3.76	6.36 - 6.18	Yook *et al.* 2001a, Pastor-Satorras *et al.* 2001
Movie actors	225,226	61	3.65	2.99	Watts, Strogatz 1998
LANL coauthorship	52,909	9.7	5.9	4.79	Newman 2001a,b
MEDLINE coauthorship	1,520,251	18.1	4.6	4.91	Newman 2001a,b
SPIRES coauthorship	56,627	173	4.0	2.12	Newman 2001a,b,c
NCSTRL coauthorship	11,994	3.59	9.7	7.34	Newman 2001a,b
Math coauthorship	70,975	3.9	9.5	8.2	Barabási *et al.* 2001
Neurosci. coauthorship	209,293	11.5	6	5.01	Barabási *et al.* 2001
E. coli, substrate graph	282	7.35	2.9	3.04	Wagner, Fell 2000
E. coli, reaction graph	315	28.3	2.62	1.98	Wagner, Fell 2000
Ythan estuary food web	134	8.7	2.43	2.26	Montoya, Solé 2000
Silwood park food web	154	4.75	3.40	3.23	Montoya, Solé 2000
Words, cooccurence	460.902	70.13	2.67	3.03	Cancho, Solé 2001
Words, synonyms	22,311	13.48	4.5	3.84	Yook *et al.* 2001
Power grid	4,941	2.67	18.7	12.4	Watts, Strogatz 1998
C. Elegans	282	14	2.65	2.25	Watts, Strogatz 1998

Small worlds

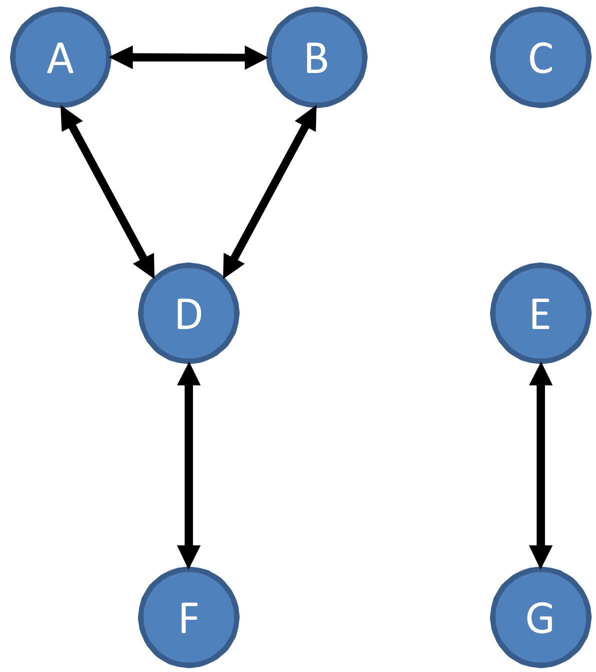
- **Paradox**: Individuals within the network are highly clustered but also have small average geodesics to other members
- Randomly rewiring a fraction of links on a regularly-clustered network drastically shortens average eccentricity
- Random rewiring, however still maintains high clustering over several orders of magnitude





Egos & alters

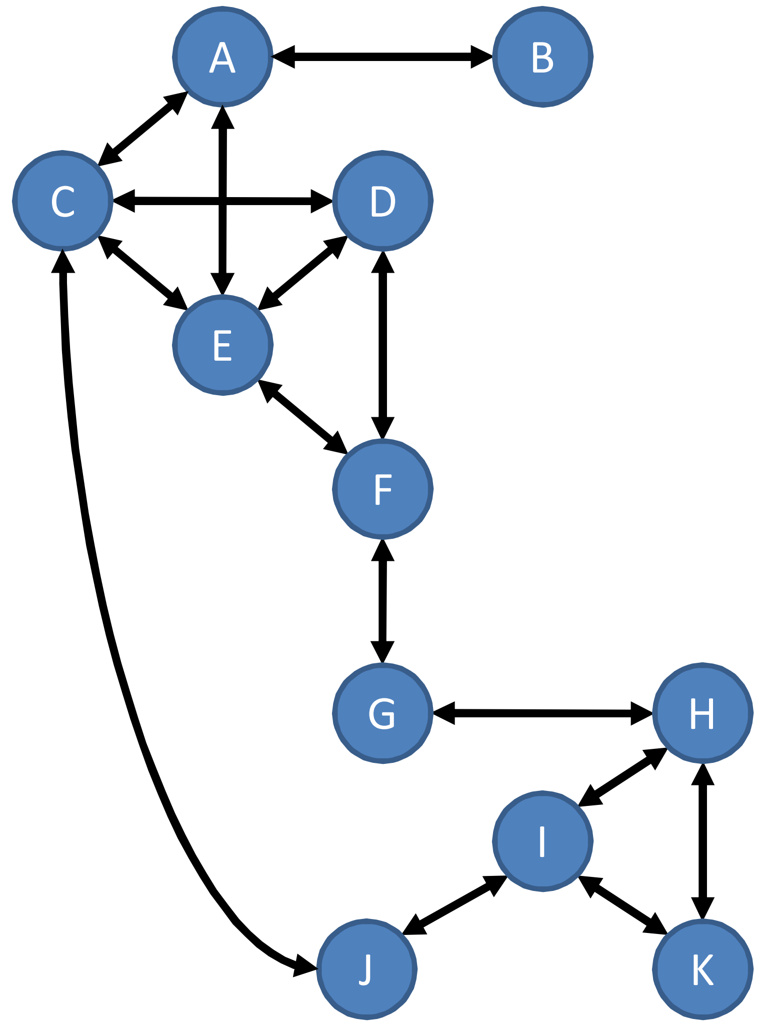
If **A** is "*ego*"



B and **D** are his "*alters*"

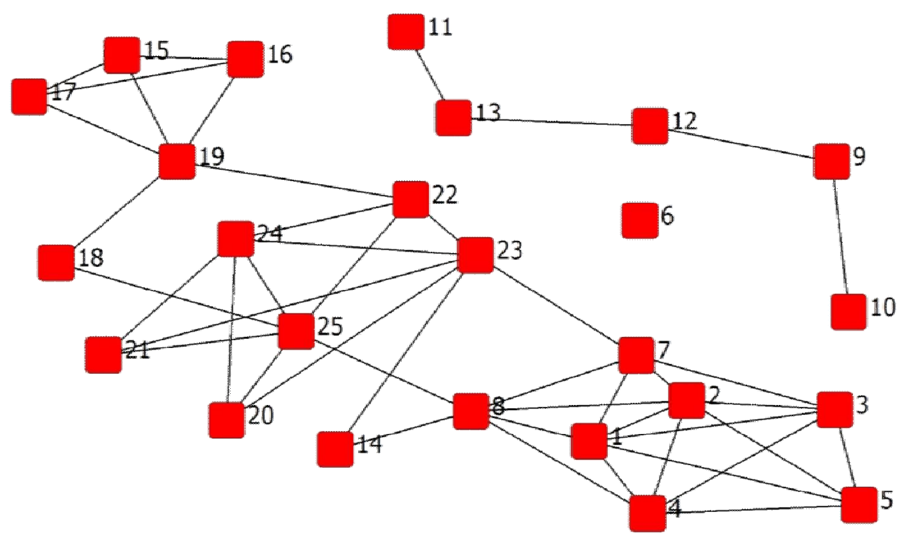
Ego network

- **<u>N-step ego network</u>**: network of all actors and their shared ties, N steps away from ego
  - E's 1-step ego network
  - E's 2-step ego network
  - E's 3-step ego network
  - E's 4-step ego network

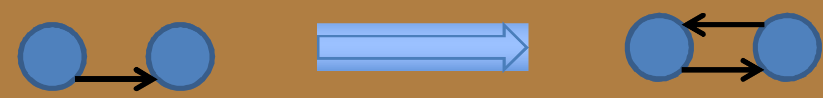
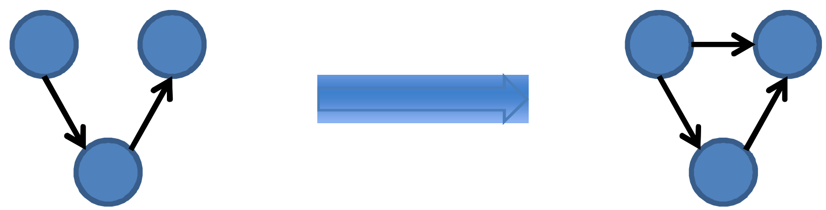
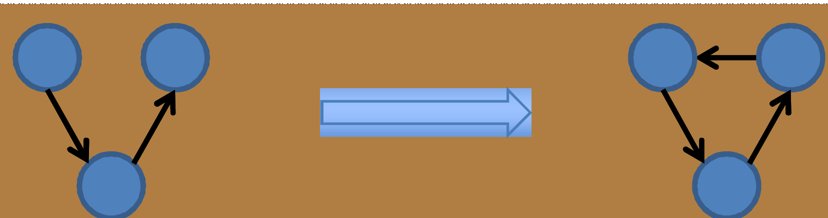
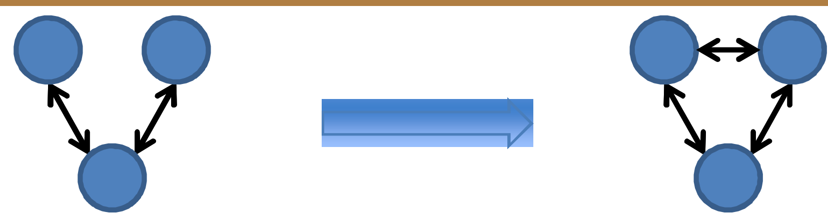


Components and cliques

- Connected component
  - Every actor is reachable from every other actor
- Giant component
  - Largest connected component
- Subgraphs
  - Subsets of actors that are disconnected from each other
- Clique
  - Subset of maximally-connected actors
- Clans, plexes, cores, & more!
  - See week 4!

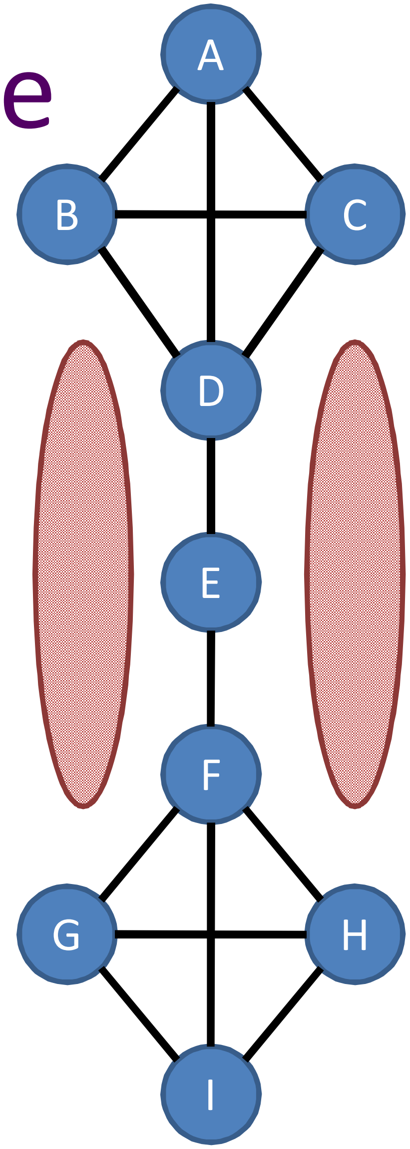


Reciprocity, transitivity, & closure

	T_1 → T_2
Reciprocity	
Transitivity	
Cyclicality	
Closure	

Structural holes & Brokerage

- **Structural holes**
  - Places where people are unconnected in a network
- **Brokers**
  - Actors who exploit structural holes
  - Gain access to information, power to filter, timing for competitive advantage, and ability to refer other actors
  - Difficult entrée, requires accurate maps of relationships in each groups, costly to maintain, high potential to be undercut



Equivalence & Closure

- **Structural equivalence**
  - One actor having the same set of relations as another actor (siblings)
  - {A,B,C}, {G,H,I}
- **Regular equivalence**
  - One actor occupying a similar position as another actor (division managers)
  - {A,B,C,G,H,I}, {D,F}
- **Closure**
  - Process of generating highly equivalent positions
  - Greater trust, high reciprocity and exchange
  - Increasing redundancy, greater constraint