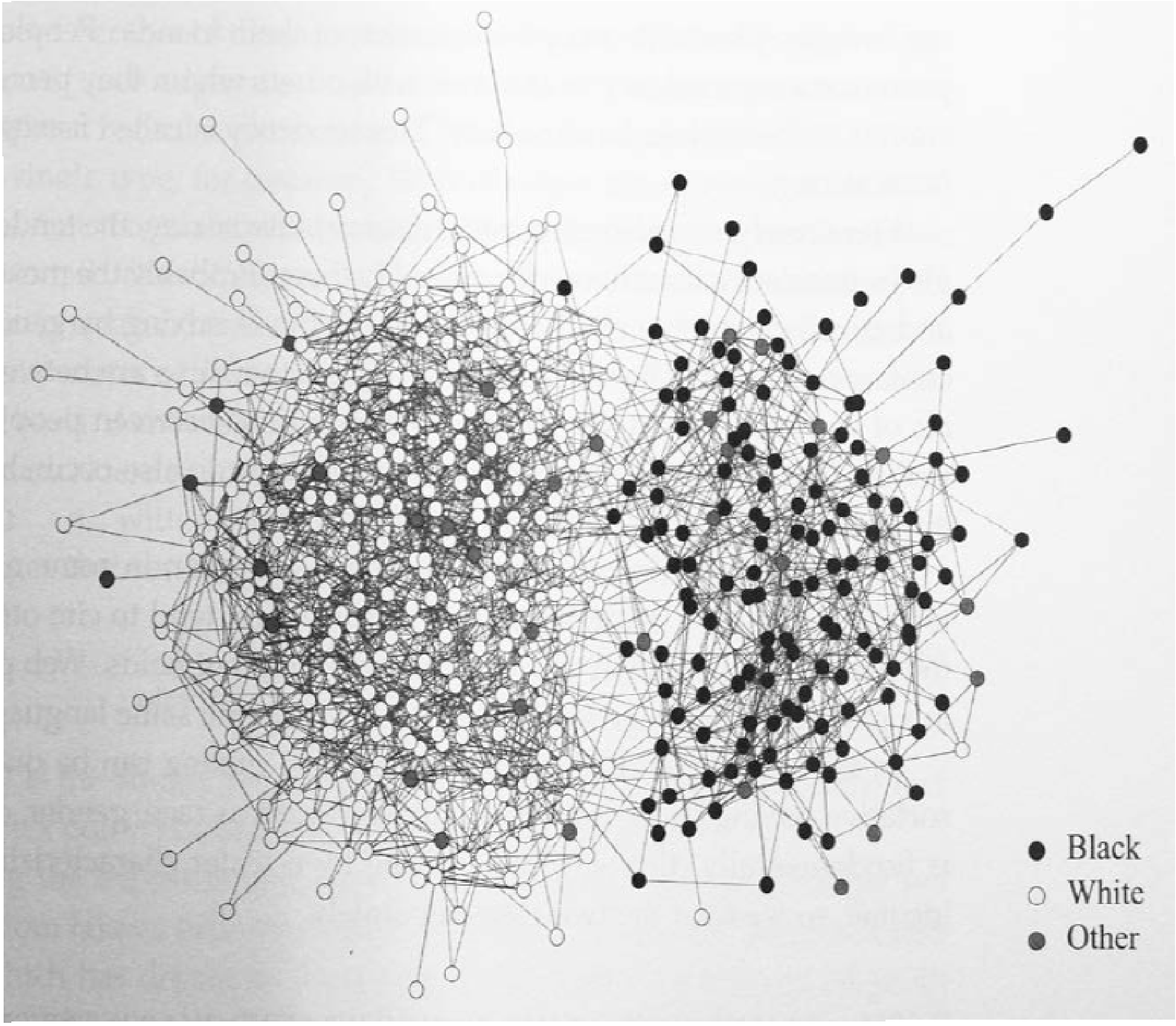
Social Networks

Assortativity (aka homophily)

friendship network at US high school: vertices colored by race



Assortativity

- Assortativity is a preference for a network's nodes to attach to others that are similar (assortative) or different (disassortative) in some way.

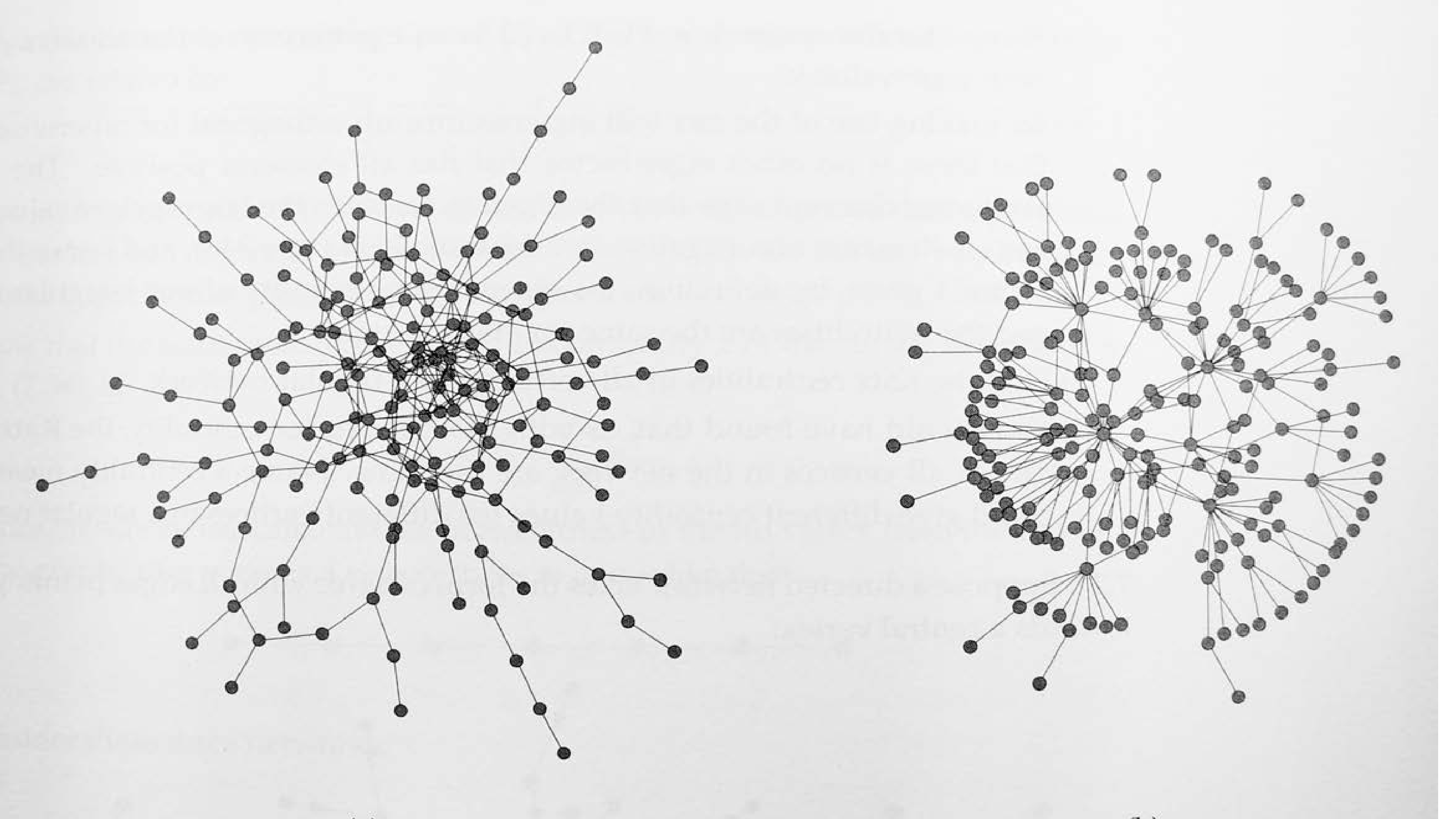
Women

Men

	black	hispanic	white	other
black	506	32	69	26
hispanic	23	308	114	38
white	26	46	599	68
other	10	14	47	32

1958 couples in the city of San Francisco, California --> self-identified their race and their partnership chocies

Assortativity (more examples)



(a) (b)

Assortativity

- Estimate degree correlation (rich goes with rich)
- The average degree of neighbors of a node with degree $k \rightarrow \langle k_{nn} \rangle$
- $\langle k_{nn} \rangle = \sum_{k'} k' P(k'|k)$
- $P(k'|k)$ the conditional probability that an edge of node degree k points to a node with degree k'
- Increasing $\rightarrow$ assortative $\rightarrow$ high degree node go with high degree node
- Decreasing $\rightarrow$ diassortative $\rightarrow$ high degree node go with low degree node

Mixing Patterns

Women

Men

e_{ij}	black	hispanic	white	other	a_i
black	0.258	0.016	0.035	0.013	0.323
hispanic	0.012	0.157	0.058	0.019	0.247
white	0.013	0.023	0.306	0.035	0.377
other	0.005	0.007	0.024	0.016	0.053
b_j	0.289	0.204	0.423	0.084	r = 0.621

$$\sum_{ij} e_{ij} = 1$$

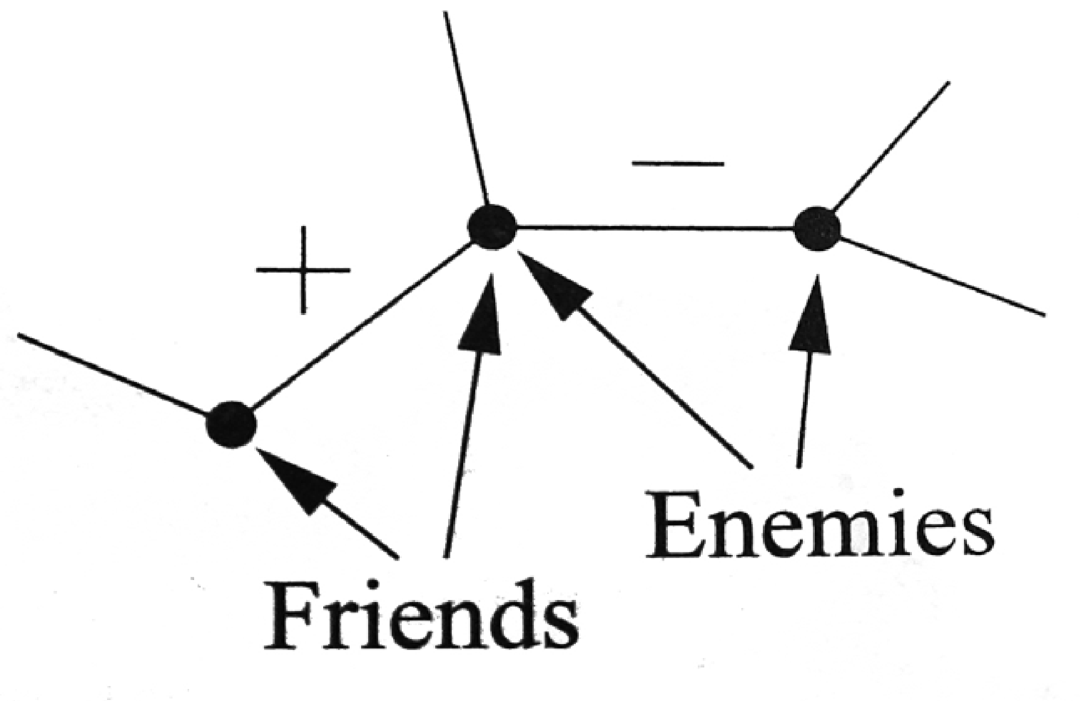
Level of assortative mixing (r):

$$(\sum_i e_{ii} - \sum a_i b_i)/(1 - \sum a_i b_i) = (\mathrm{Tr}\,\boldsymbol{e} - ||\boldsymbol{e}^2||)/(1 - ||\boldsymbol{e}^2||)$$

Perfectly assortative, $r = 1$

Signed Graphs

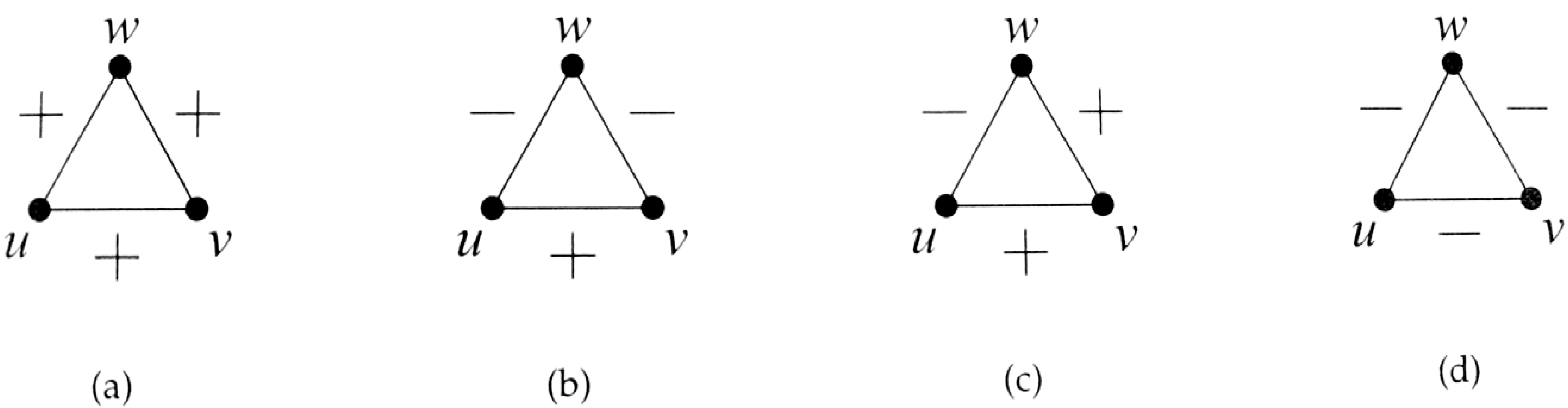
- signed network: network with signed edges
- "+" represents friends
- "-" represents enemies



How will our class look?

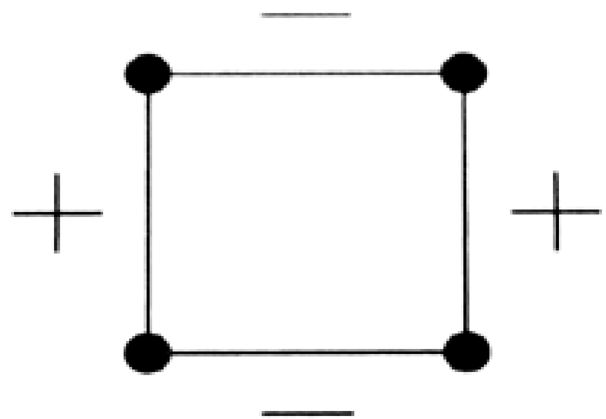
Possible Triads

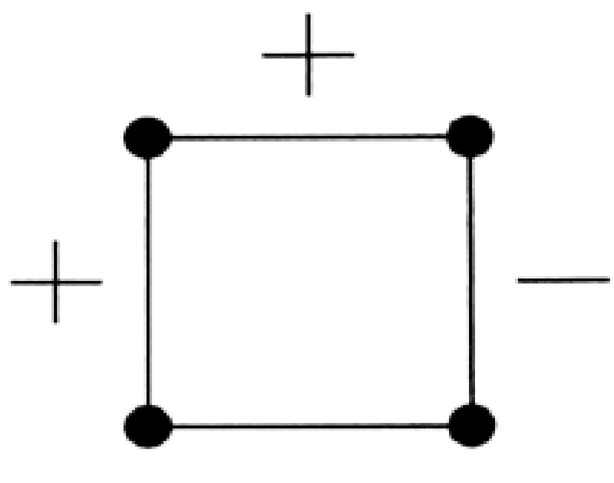
- (a), (b) balanced, relatively stable
- (c), (d) unbalanced, liable to break apart
- triad is stable ← even number of “-” signs around loop



Stable configurations 4 cycles??

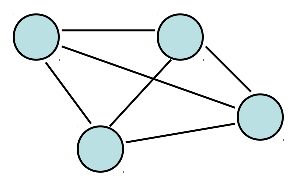
Stable configurations 4 cycles??



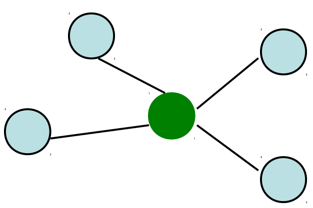


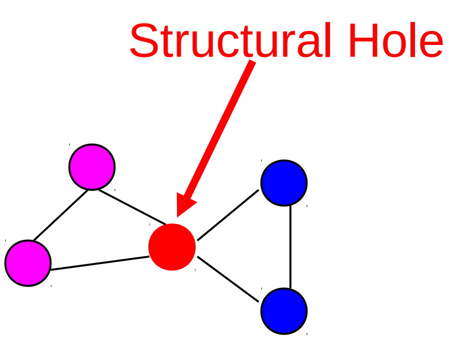
Structural Holes

- Structural holes are nodes (mainly in a social network) that separate non-redundant sources of information, sources that are additive than overlapping
- Redundancy
  - Cohesion – contacts strongly connected to each other are likely to have similar information and therefore provide redundant information (same clor nodes) benefits



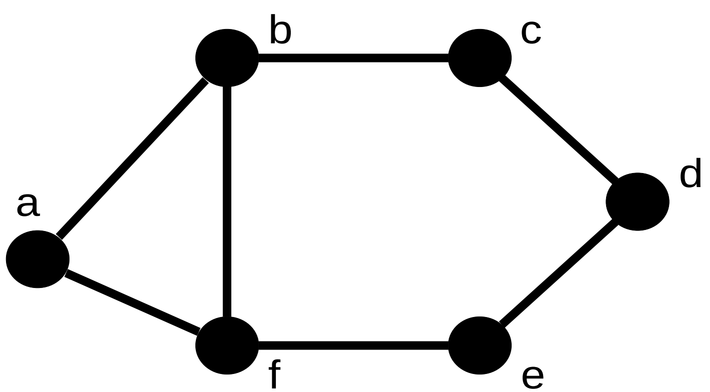
  - Equivalence – contacts that link a manager to the same third parties have same sources of information and therefore provide redundant information benefit





Social Cohesiveness (distance based)

- Refers to the cliquishness
- A complete clique is too strict to be practical
- Most of the real groups have at least a few members who don’t know each other --> relaxing the definition
- *k-cliques* => Any maximal set S of nodes in which the geodesic path between every pair of nodes $\{u, v\} \in S$ is <=k



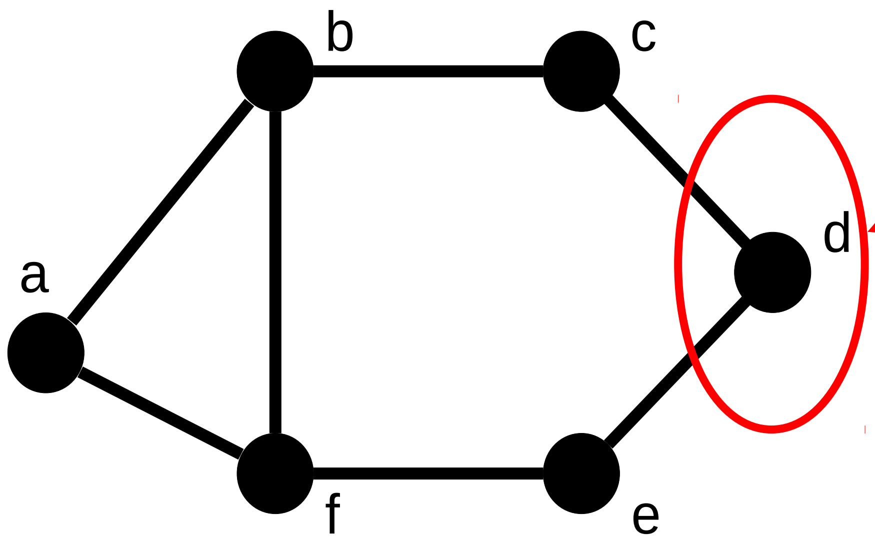
{a, b, c, f, e} is a k-clique
--> What is the value of k?

Social Cohesiveness (distance based)

- Do you see the problem with the previous definition?

Social Cohesiveness (distance based)

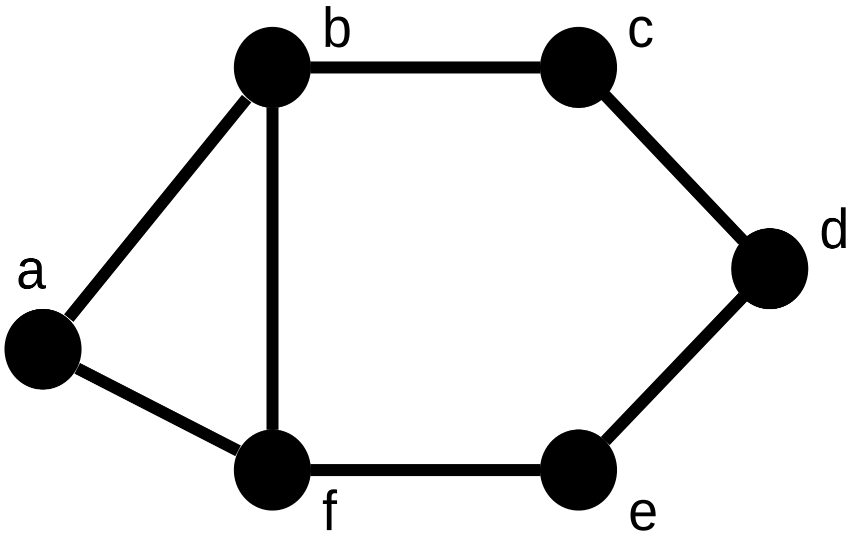
- Do you see the problem with the previous definition?
- k-cliques might not be as cohesive as they look!



Not a member of the clique {a, b, c, f, e} --> causes the distance between c and e to be 2

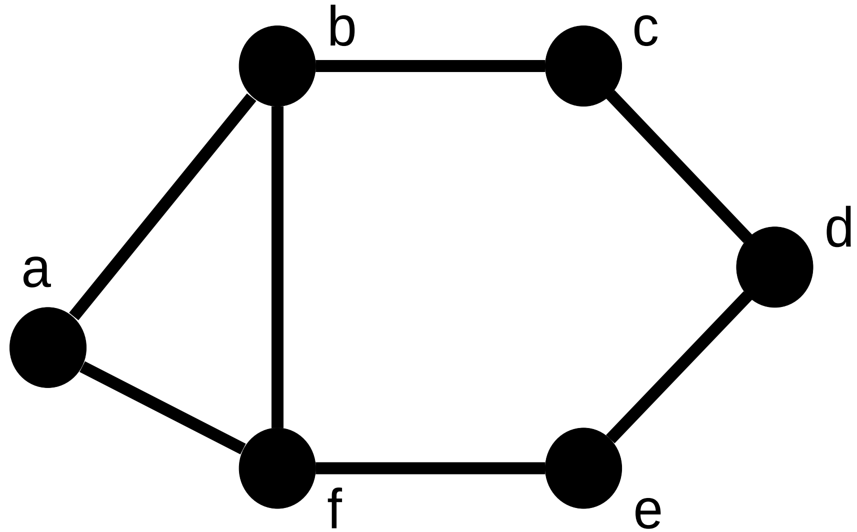
Social Cohesiveness (distance based)

- Soultion: *k-clans*
- A k-clique in which the subgraph induced by S has a diameter <= k
- {b, c, d, e, f} is a k-clan
- {b, e, f} also induces a subgraph that has diameter = 2



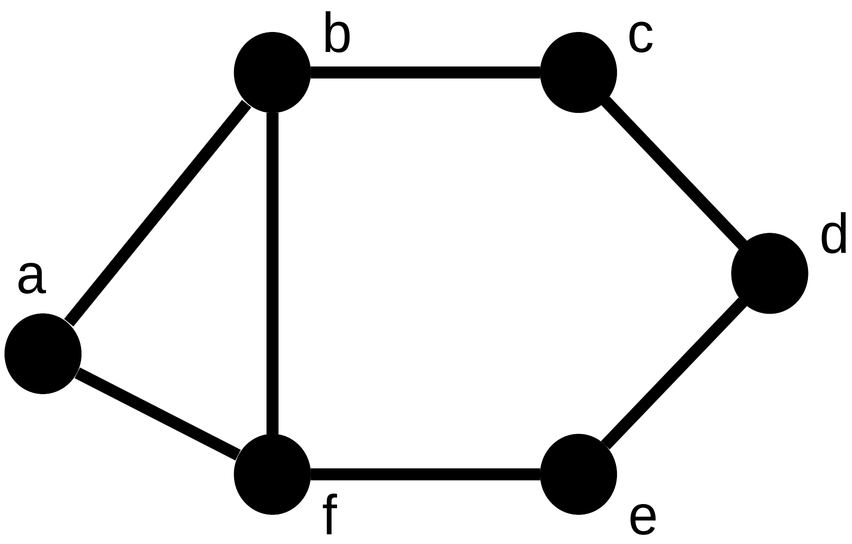
Social Cohesiveness (distance based)

- Soultion: *k-clans*
- A k-clique in which the subgraph induced by S has a diameter <= k
- {b, c, d, e, f} is a k-clan
- {b, e, f} also induces a subgraph that has diameter = 2 <-- NOT a k-clan
  - {a, b, f, e} is not maximal (k-clique criteria)



Social Cohesiveness (distance based)

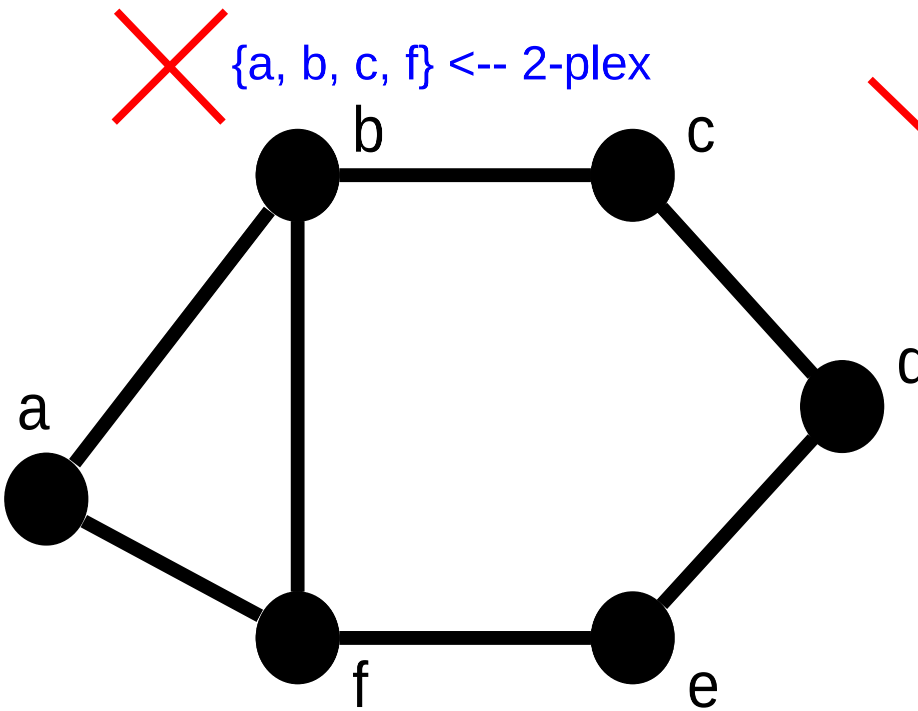
- Relax maximality condition on k-clans: *k-club*
- {a, b, f, e} is a k-club

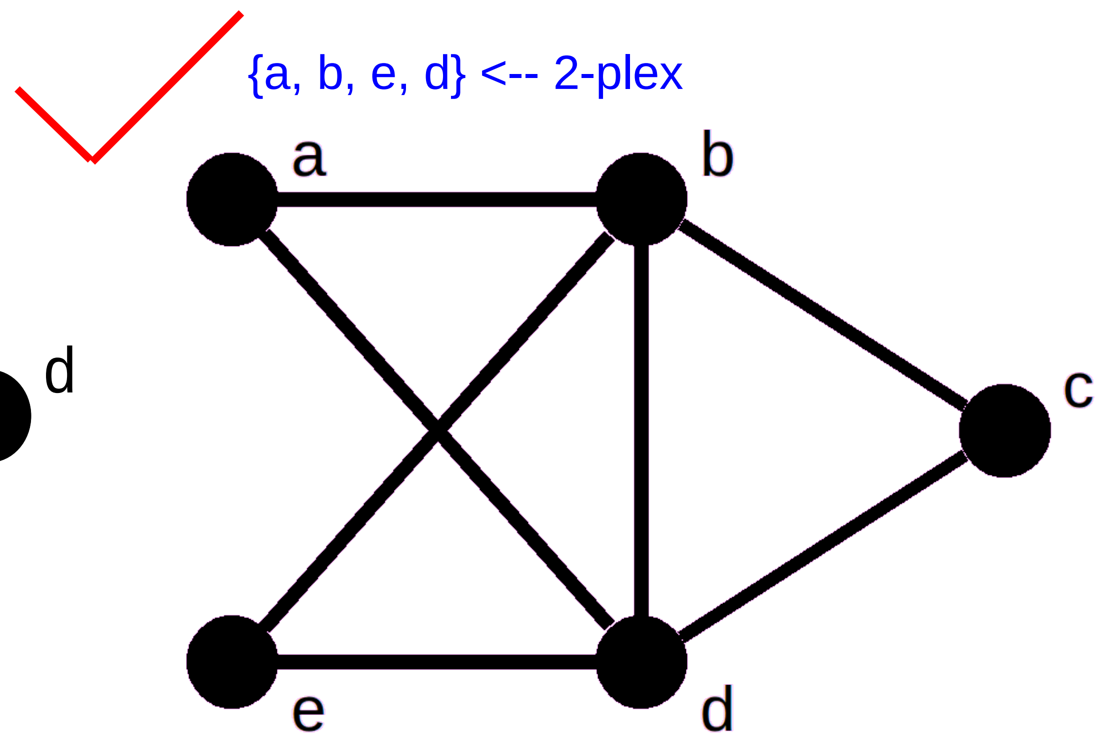


- k-clan is a k-clique and also a k-club

Social Cohesiveness (degree based)

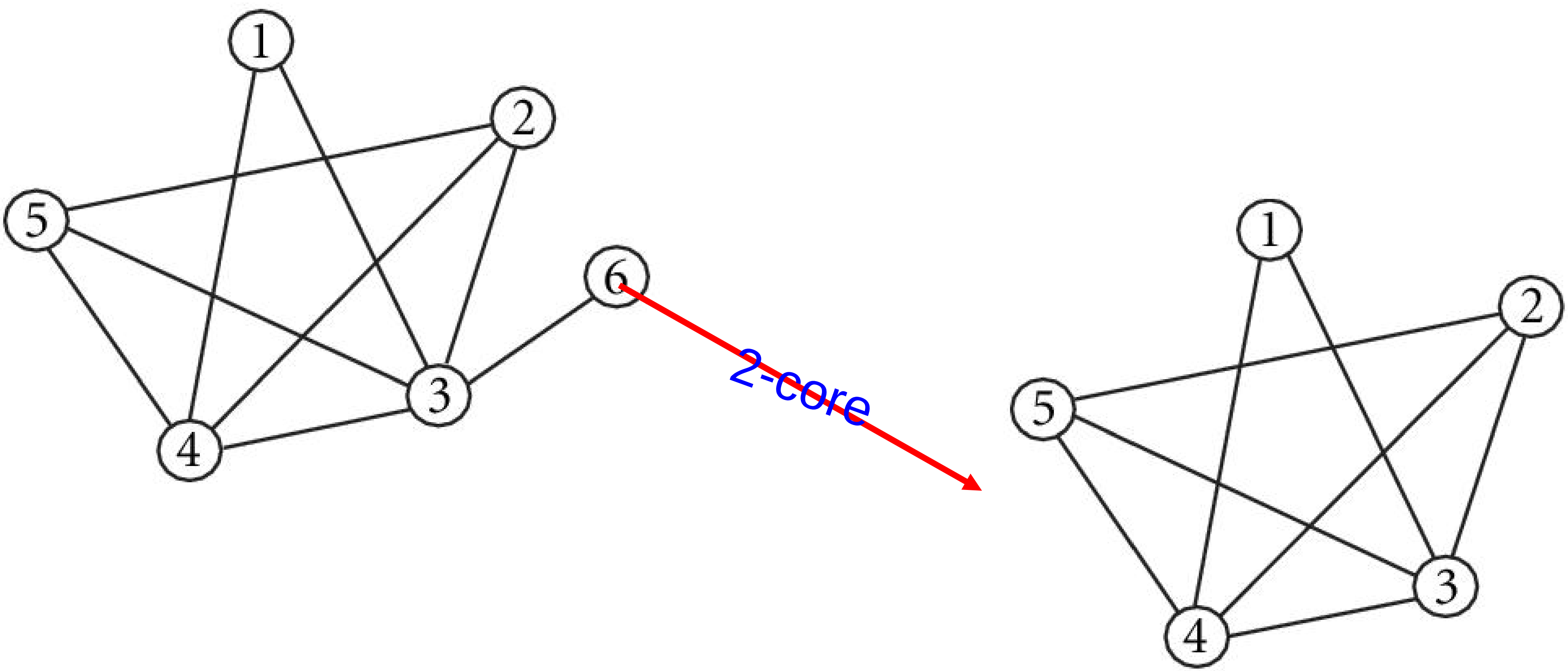
- A *k-plex* is a maximal subset S of nodes such that every member of the set is connected to n-k other members, where n is the size of S





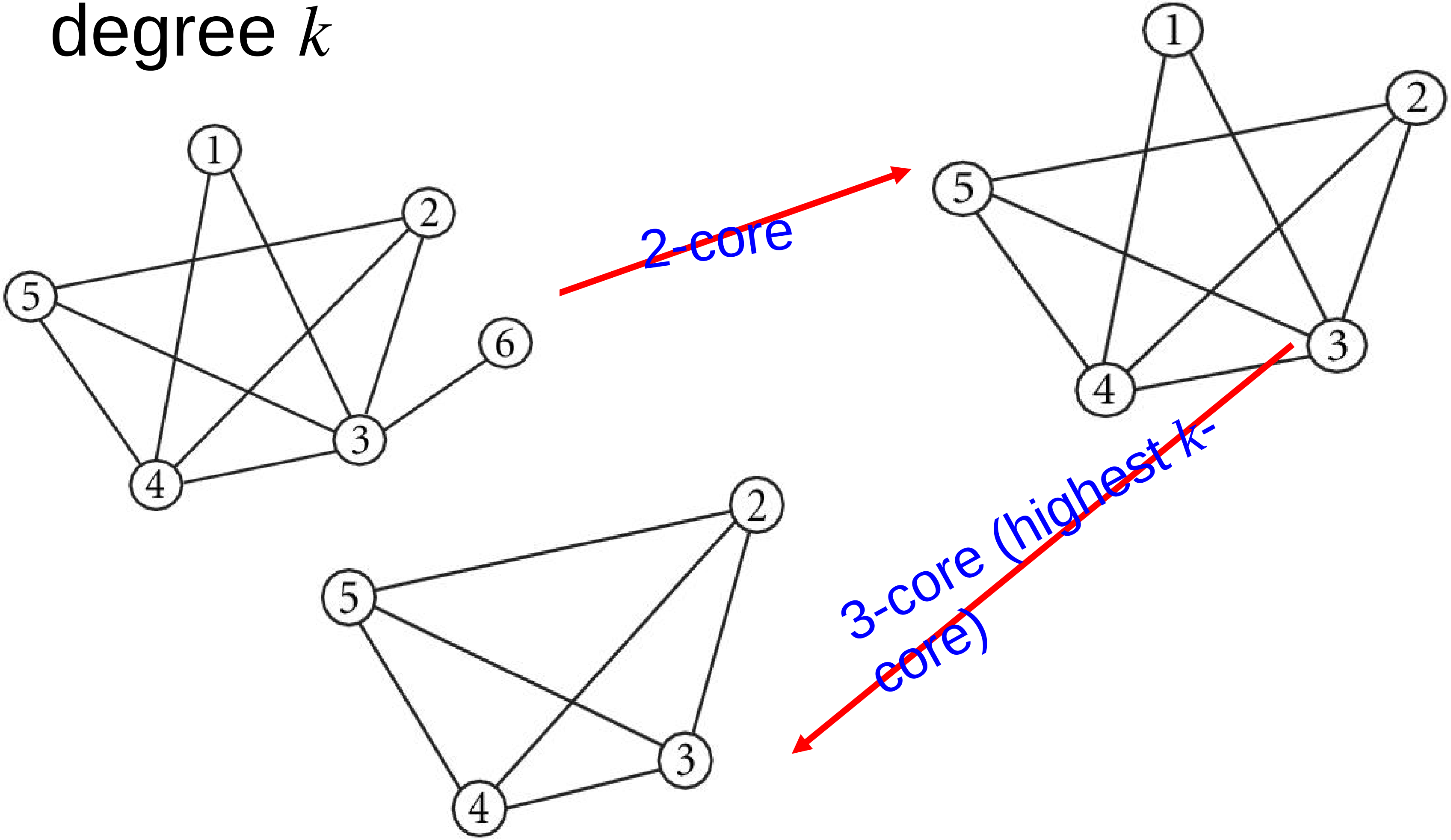
Social Cohesiveness (degree based)

- A *k-core* of a graph is a maximal subgraph such that each node in the subgraph has at least degree k



Social Cohesiveness (degree based)

- A *k-core* of a graph is a maximal subgraph such that each node in the subgraph has at least degree k

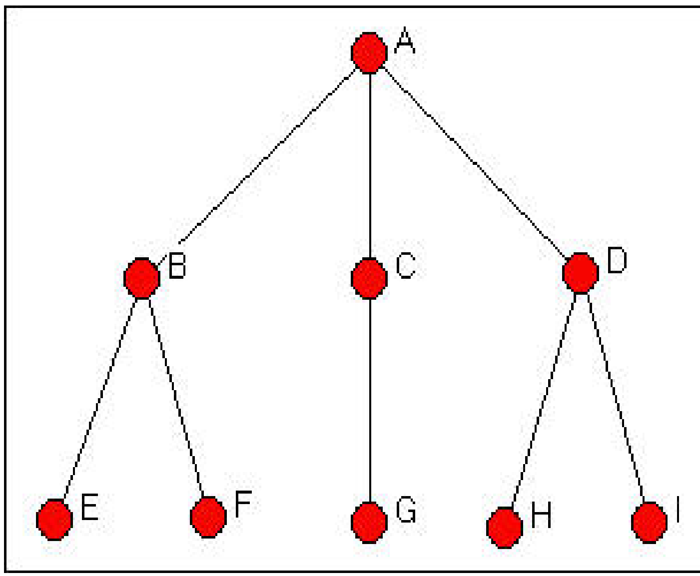


Social Roles

- "Positions" or "roles" or "social categories" are defined by "relations" among actors
- Two actors have the same "position" or "role" to the extent that their pattern of relationships with other actors is the same
- How does one define such a similarity?
- Me (faculty) → IIT ←you (students): In some ways our relationships are same (“being governed”); but there are differences also: you pay them while they pay me
- Which relations should count and which ones not, in trying to describe the roles of "faculty" and "student”?

Equivalence

- Structural Equivalence



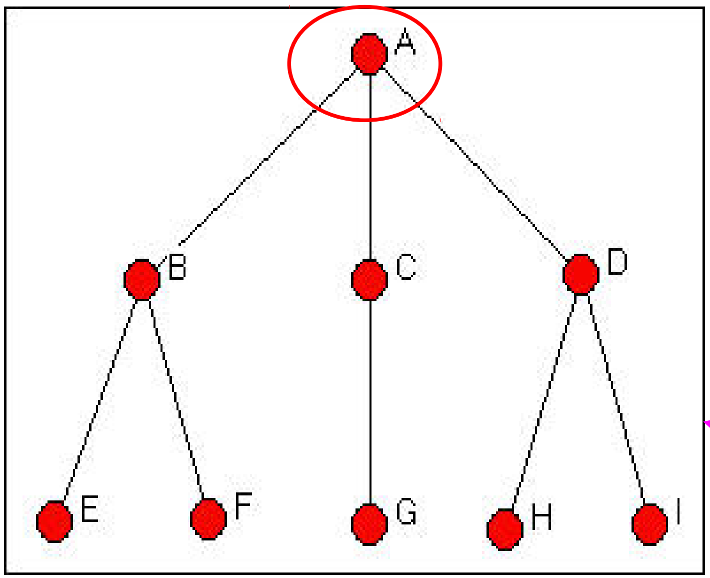
Wasserman-Faust network

Two nodes are said to be exactly structurally equivalent if they have the same relationships to all other nodes --> One should be perfectly substitutable by the other.

What are the equivalence classes?

Equivalence

- Structural Equivalence



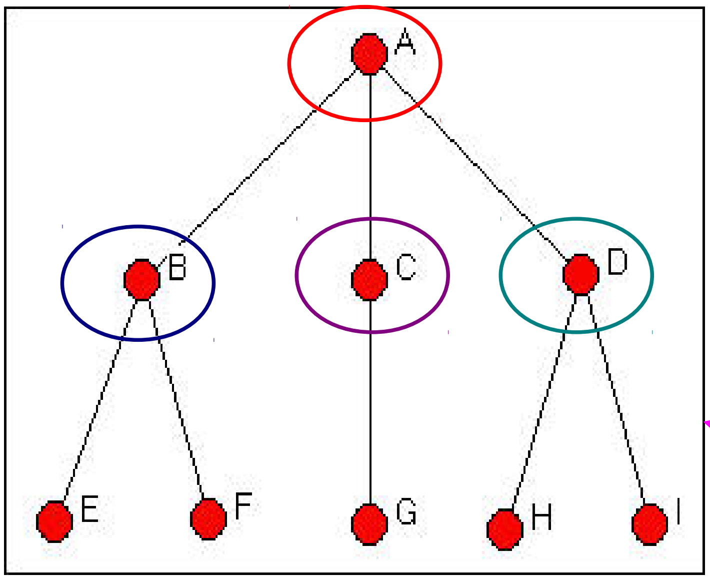
Wasserman-Faust network

Two nodes are said to be exactly structurally equivalent if they have the same relationships to all other nodes --> One should be perfectly substitutable by the other --> connected to exactly the same set of neighbors

What are the equivalence classes?

Equivalence

- Structural Equivalence



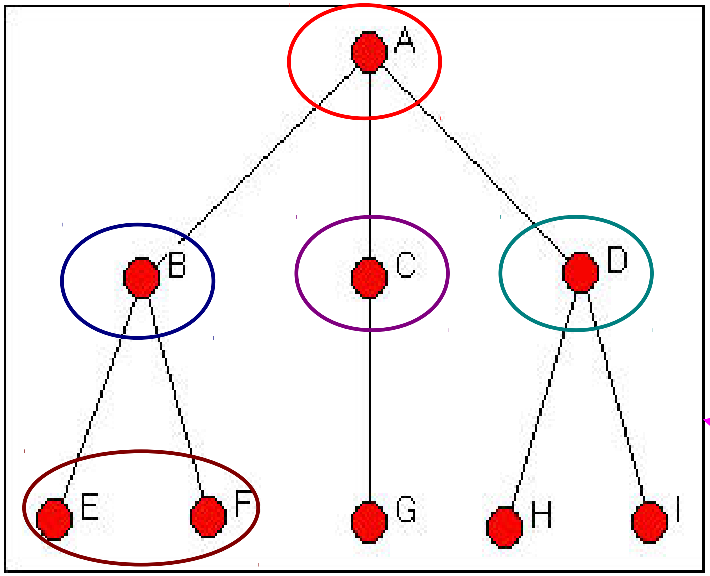
Wasserman-Faust network

Two nodes are said to be exactly structurally equivalent if they have the same relationships to all other nodes --> One should be perfectly substitutable by the other --> connected to exactly the same set of neighbors

What are the equivalence classes?

Equivalence

- ## Structural Equivalence



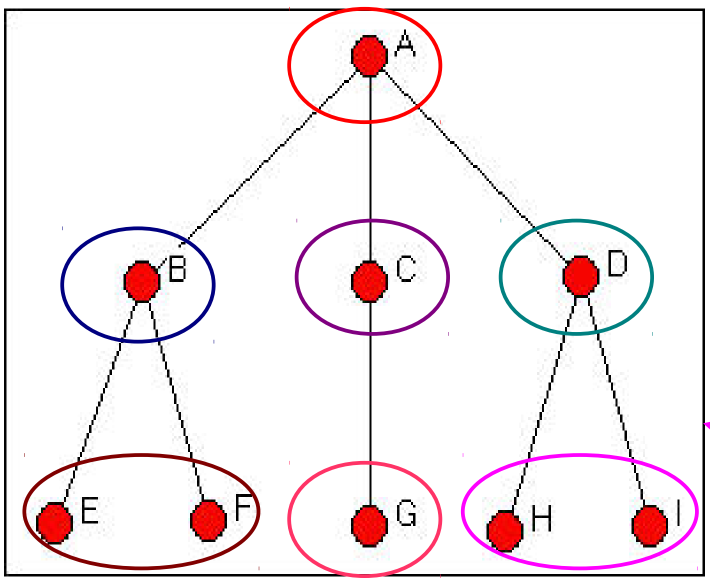
Wasserman-Faust network

Two nodes are said to be exactly structurally equivalent if they have the same relationships to all other nodes --> One should be perfectly substitutable by the other --> connected to exactly the same set of neighbors

What are the equivalence classes?

Equivalence

- Structural Equivalence



Wasserman-Faust network

Two nodes are said to be exactly structurally equivalent if they have the same relationships to all other nodes --> One should be perfectly substitutable by the other --> connected to exactly the same set of neighbors

What are the equivalence classes?

Exact structural equivalence is likely to be rare (particularly in large networks) --> Examine the degree of structural equivalence --> Any idea about how to measure?

Degree of Equivalence

- How to measure?
- Hint: Number of common neighbors

Degree of Equivalence

- How to measure?
- Hint: Number of common neighbors
- $n_{ij} = \sum_k A_{ik} A_{kj} \leftarrow ij^{th}$ element of the matrix A^2
- closely related to the cocitation measure (in directed networks)
- Any problem with this measure $\leftarrow$ remember you are measuring the extent of similarity

Degree of Equivalence

- How to measure?
- Hint: Number of common neighbors
- $n_{ij} = \sum_k A_{ik} A_{kj} \leftarrow ij^{th}$ element of the matrix A^2
- closely related to the cocitation measure (in directed networks)
- Any problem with this measure $\leftarrow$ remember you are measuring the extent of similarity
- Appropriate normalization

Cosine similarity

- Inner product of two vectors

$$similarity(x, y) = cos(\theta) = \frac{x \cdot y}{||x|| * ||y||}$$

- Ө=0 → maximum similarity Ө=90 → no similarity
- Consider i^{th} and the j^{th} row as vectors
- cosine similarity between vertices i and j
- $\sigma_{ij} = (\sum_k A_{ik} A_{kj})/(\sqrt{\sum_k A_{ik}^2} \sqrt{\sum_k A_{kj}^2}) = n_{ij}/\sqrt{(k_i k_j)}$

Pearson Correlation

- Correlation coefficient between rows *i* and *j*

$$r_{ij} = \frac{\mathrm{Cov}(X_i, X_j)}{\sqrt{Var(X_i)Var(X_j)}}$$

$$= \frac{\sum_k A_{ik}A_{jk} - \frac{k_i k_j}{n}}{\sqrt{k_i - \frac{k_i^2}{n}}\sqrt{k_j - \frac{k_j^2}{n}}}$$

Euclidean Distance

$$d_{ij} = \sum_k (A_{ik} - A_{jk})^2$$

- For a binary graph??

Euclidean Distance

$$d_{ij} = \sum_k (A_{ik} - A_{jk})^2$$

- For a binary graph → Hamming distance
- Normalization → What could be the maximum possible distance?

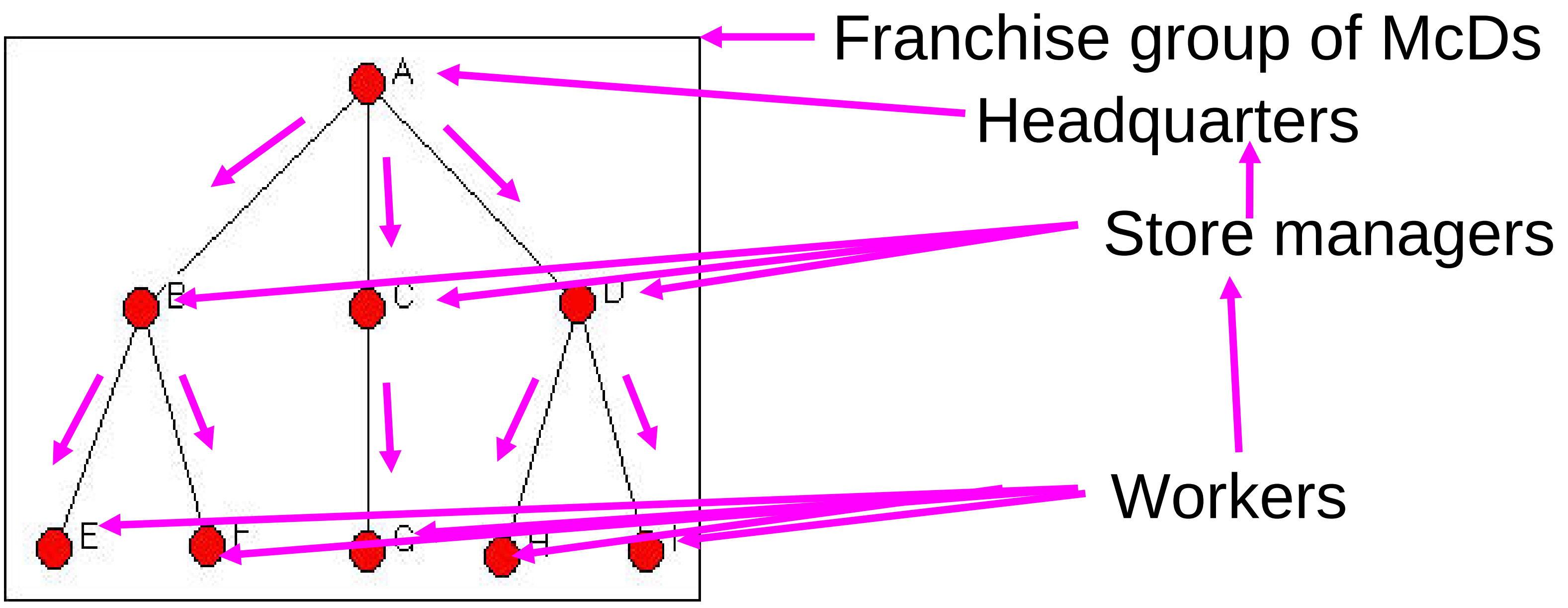
Euclidean Distance

$$d_{ij} = \sum_k (A_{ik} - A_{jk})^2$$

- For a binary graph $\rightarrow$ Hamming distance
- Normalization $\rightarrow$ What could be the maximum possible distance?
- None of i's neighbors (k_i) match with j's neighbors (k_j) $\rightarrow k_i + k_j$
- Similarity = $\frac{d_{ij}}{k_i + k_j}$

Equivalence

- Automorphic Equivalence

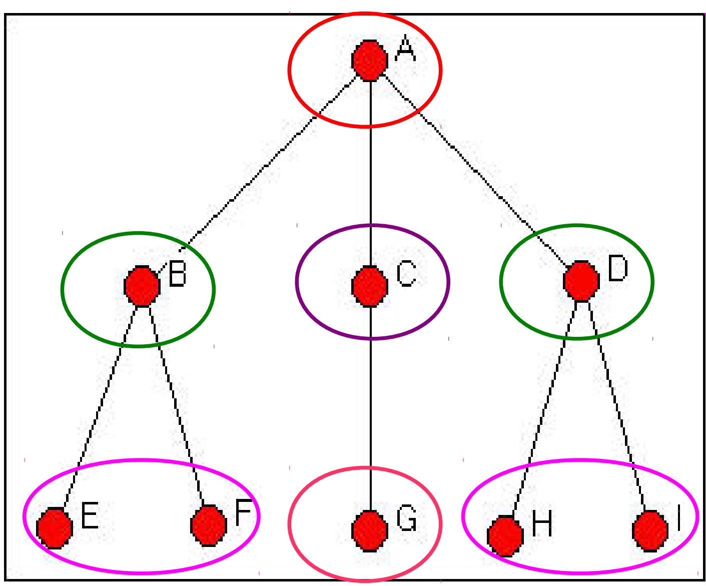


Wasserman-Faust network

B and D are not structurally equivalent --> But do they look equivalent?

Equivalence

- Automorphic Equivalence



Wasserman-Faust network

Franchise group of McDs

Headquarters

Store managers

Workers

B and D are not structurally equivalent --> But do they look equivalent?
Yes --> they have exactly the same boss and the same number of workers --> If we swapped them, and also swapped the four workers, all of the distances among all the actors in the graph would be exactly identical

Equivalence

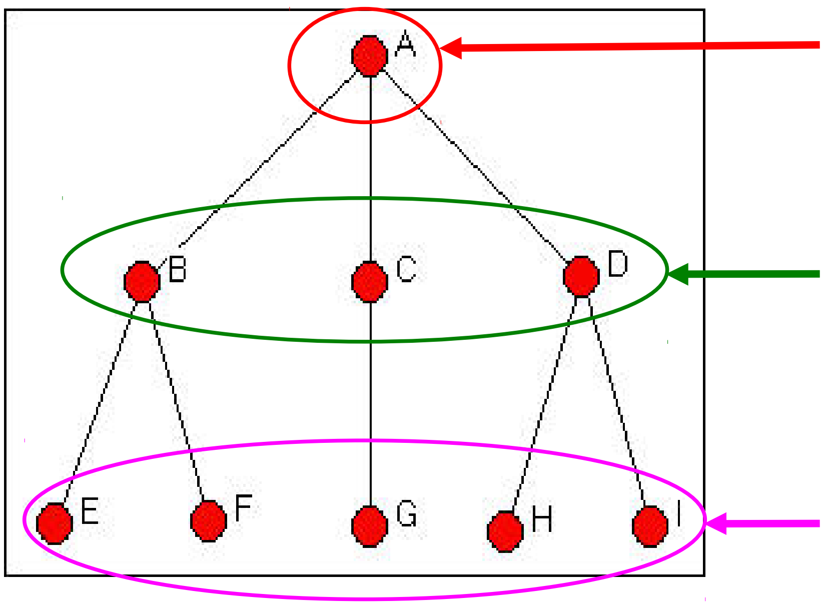
- Regular Equivalence
- nodes i and j are regularly equivalent if their profile of ties is similar to other set of actors that are also regularly equivalent → recursive
- Consider the social role “mother”: what are the social ties?

Equivalence

- Regular Equivalence
- nodes i and j are regularly equivalent if their profile of ties is similar to other set of actors that are also regularly equivalent → recursive
- Consider the social role “mother”: what are the social ties → husband, children, in-laws etc.
- All mothers will have a similar profile of such ties
- Not structurally equivalent ← two mothers usually don't have same husband, children or in-law
- Not automorphically equivalent ← Different mother have different numbers of husbands, children or in-law
- But still (possibly functionally) they share some similarity

Equivalence

- Regular Equivalence



Class 1: at least 1 tie to class 1 + no ties to class 2

Class 2: at least 1 tie to class 1 + at least one tie to class 2

Class 3: no ties to class 1 and at least one tie to class 2

Computing Regular Equivalence

- Regular Equivalence: vertices i, j are similar if i has a neighbor k that is itself similar to j

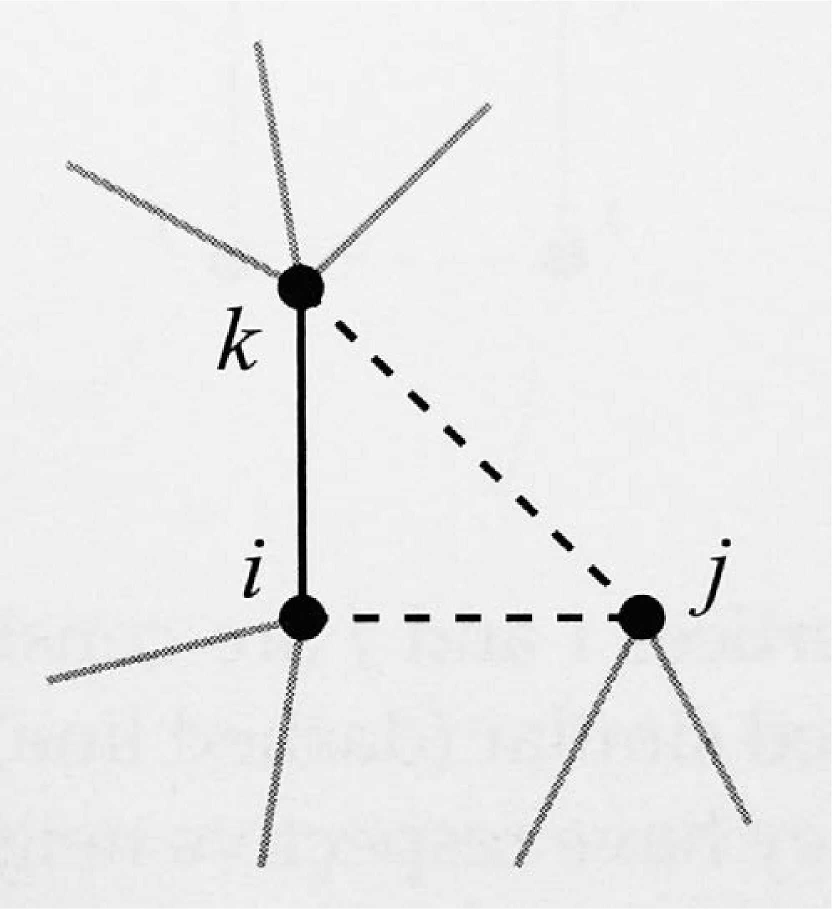
$$\sigma_{ij} = \alpha \sum_k A_{ik}\sigma_{kj} + \delta_{ij}$$

- Matrix form

$$\sigma = \alpha \mathbf{A}\sigma + \mathbf{I}$$

$$\downarrow$$

$$\sigma = (\mathbf{I} - \alpha \mathbf{A})^{-1}$$



Ego-Centric Networks

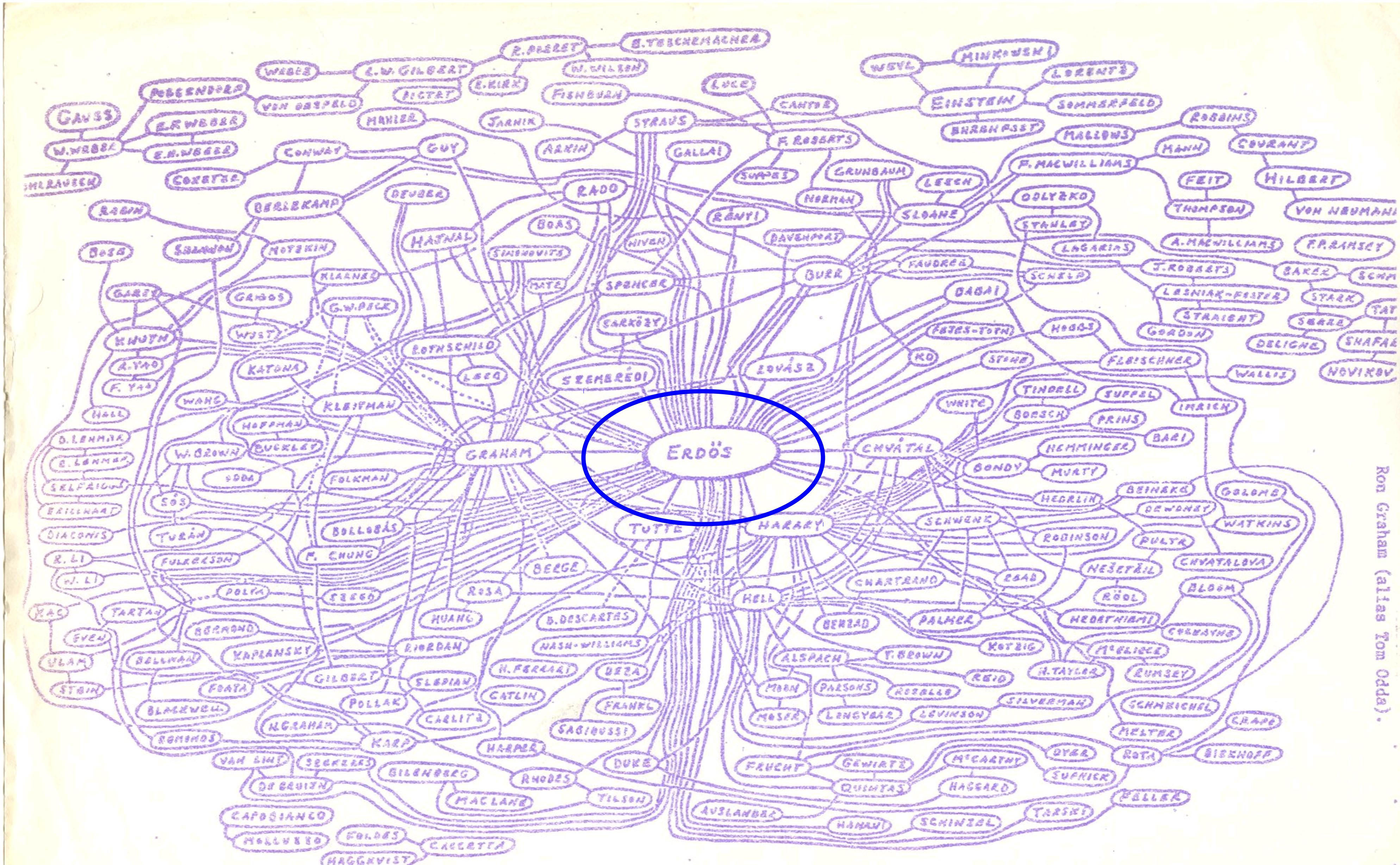


Figure 1
To appear in Topics in Graph Theory (F. Harary, ed.) New York Academy of Sciences (1979).