Social Network Analysis

Prof. Vrijendra Singh
Dept of IT, IIIT Allahabad
vrij@iiita.ac.in
7376581984

Fair Use Disclaimer

The presentation contains a very few information/material/images from the web resources, journal papers, books etc. under fair use policy. The materials have been taken for explaining the relevant issues related to the Social Network Analysis to the scientific community. If you wish to use any material from this presentation for the purpose other than the “fair use”, you must obtain the explicit permission from the copyright owner.

Social Networks

- A **social network** is a social structure of people, related (directly or indirectly) to each other through a common relation or interest
- **Social network analysis (SNA)** is the study of social networks to understand their structure and behavior

(Source: Freeman, 2000)

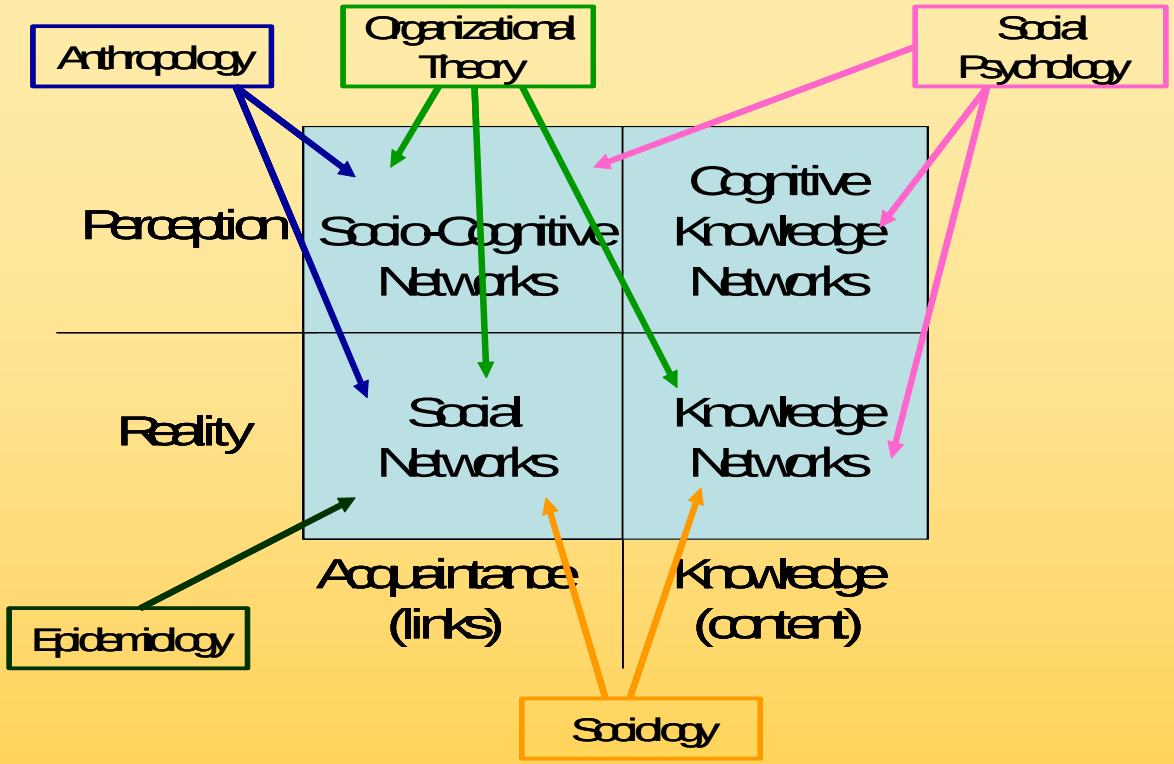
Networks in Social Sciences

- Types of Networks (Contractor, 2006)
  - Social Networks
    - “who knows who”
  - Socio-Cognitive Networks
    - “who thinks who knows who”
  - Knowledge Networks
    - “who knows what”
  - Cognitive Knowledge Networks
    - “who thinks who knows what”

Types of Social Network Analysis

- **Sociocentric (whole) network analysis**
  - Emerged in sociology
  - Involves quantification of interaction among a socially well-defined group of people
  - Focus on identifying global structural patterns
  - Most SNA research in organizations concentrates on sociometric approach
- **Egocentric (personal) network analysis**
  - Emerged in anthropology and psychology
  - Involves quantification of interactions between an individual (called *ego*) and all other persons (called *alters*) related (directly or indirectly) to ego
  - Make generalizations of features found in personal networks
  - Difficult to collect data, so till now studies have been rare

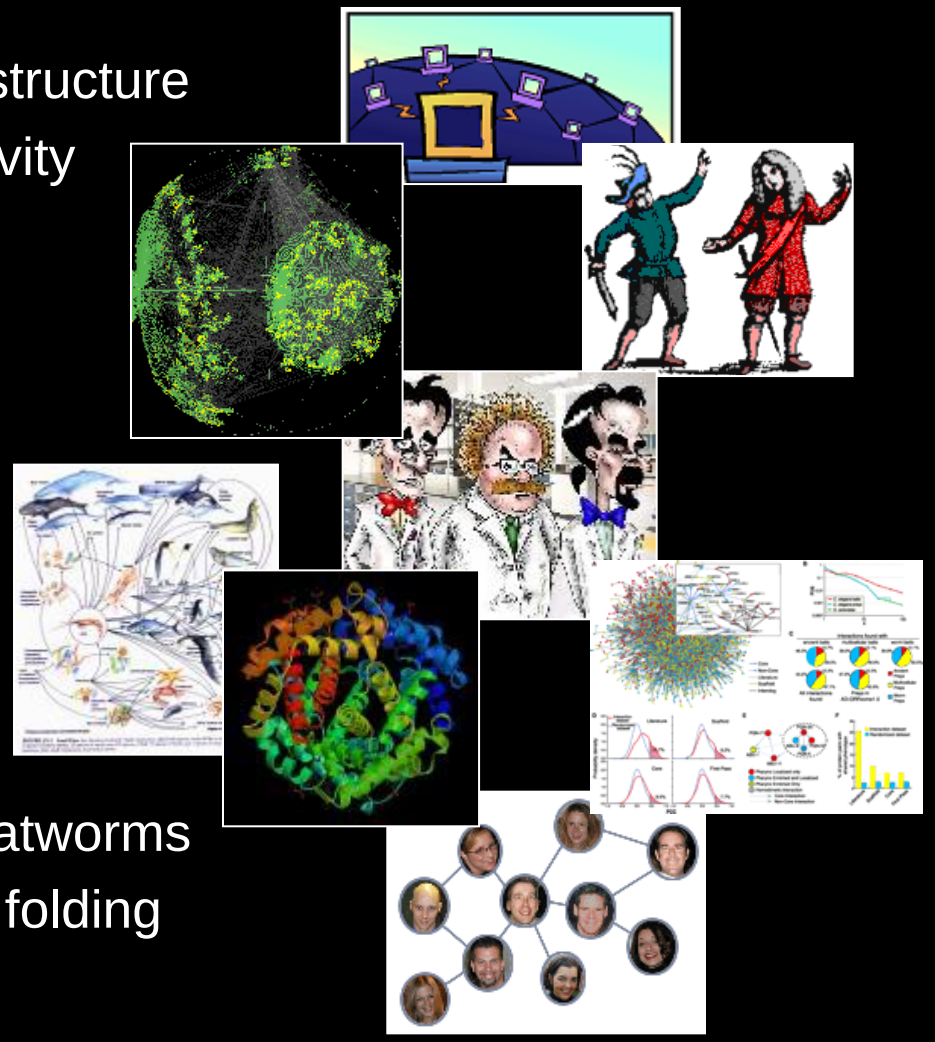
Networks Research in Social Sciences



- Social science networks have widespread application in various fields
- Most of the analyses techniques have come from Sociology, Statistics and Mathematics
- See (Wasserman and Faust, 1994) for a comprehensive introduction to social network analysis

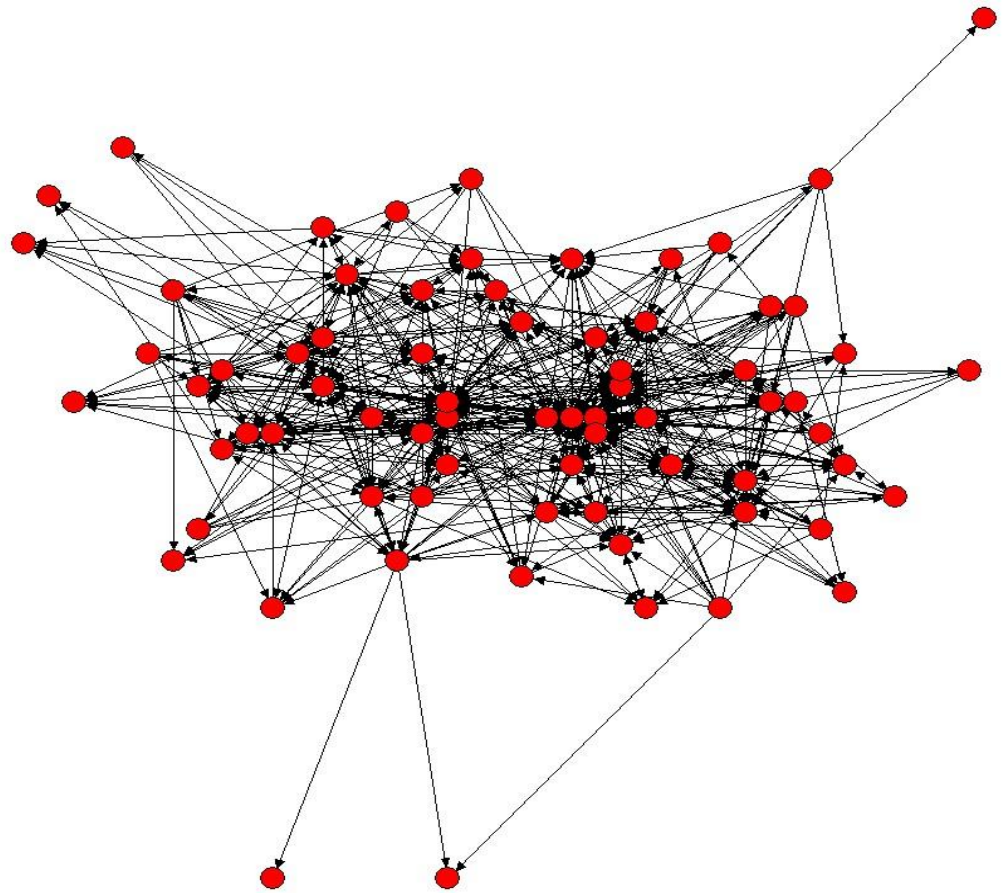
Applications of Network Theory

- World Wide Web and hyperlink structure
- The Internet and router connectivity
- Collaborations among…
  - Movie actors
  - Scientists and mathematicians
- Sexual interaction
- Cellular networks in biology
- Food webs in ecology
- Phone call patterns
- Word co-occurrence in text
- Neural network connectivity of flatworms
- Conformational states in protein folding



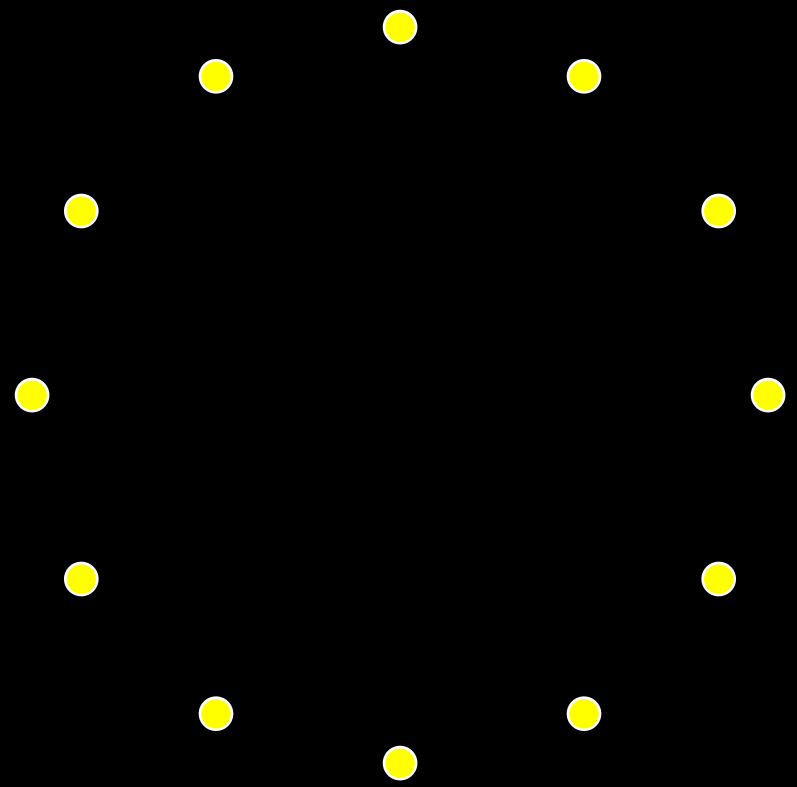
Web Applications of Social Networks

- Analyzing page importance
  - Page Rank
    - Related to recursive in-degree computation
  - Authorities/Hubs
- Discovering Communities
  - Finding near-cliques
- Analyzing Trust
  - Propagating Trust
  - Using propagated trust to fight spam
    - In Email
    - In Web page ranking



Society as a Graph

People are represented as *nodes/actors/vertices/points.*



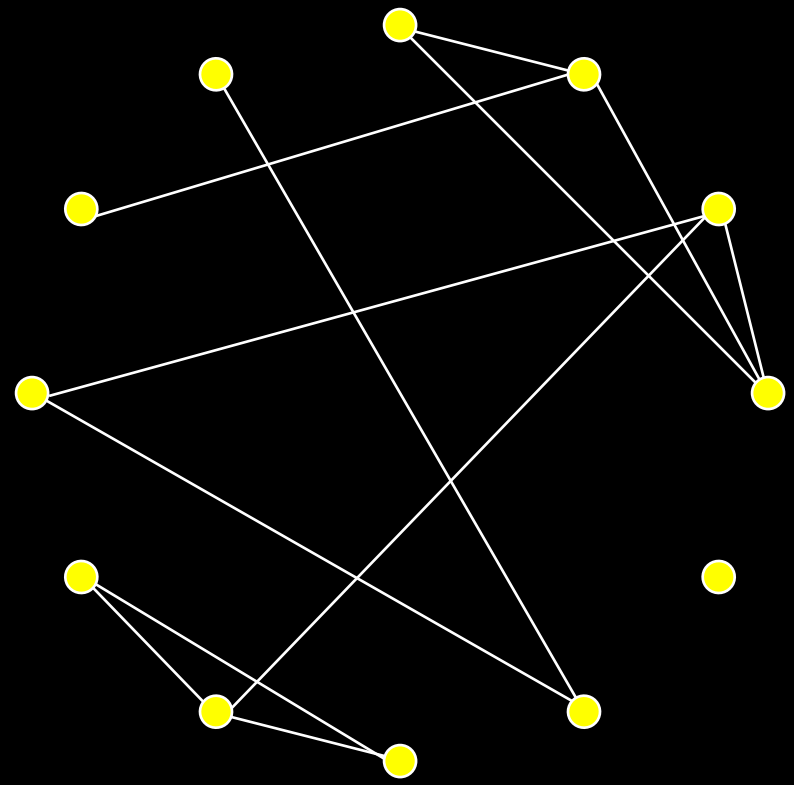
Actors / nodes / vertices / points
Computers / Telephones
Persons / Employees
Companies / Business Units
Articles / Books
Can have properties (attributes)

Society as a Graph

People are represented as *nodes/actors/vertices/points.*

Relationships are represented as *edges/ties/arcs/lines/links.*

(Relationships may be acquaintanceship, friendship, co-authorship, etc.)



Society as a Graph

Actors / nodes / vertices / points

Ties / edges / arcs / lines / links

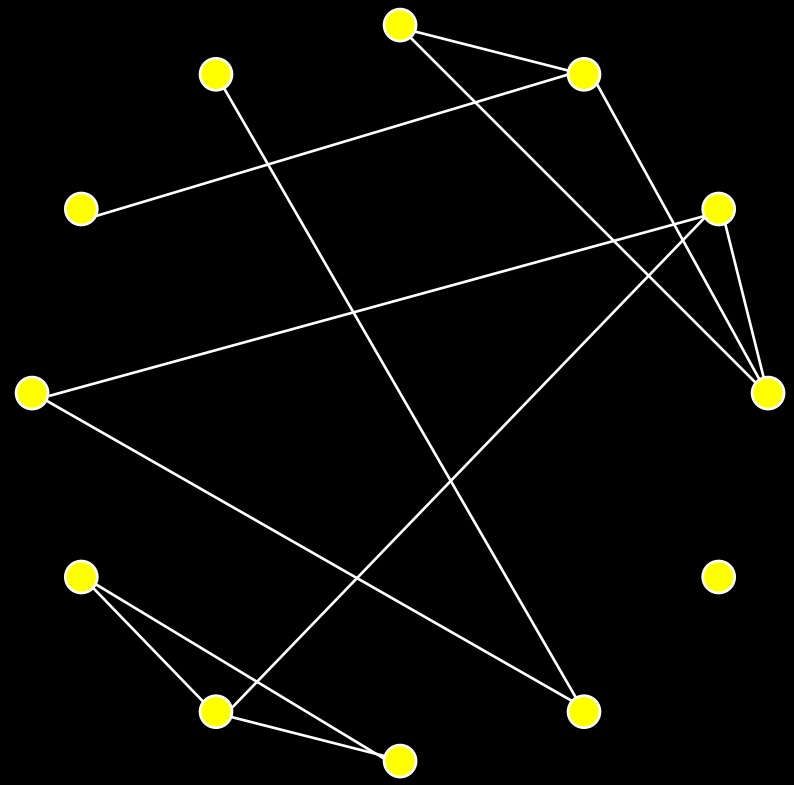
connect pair of actors

types of social relations

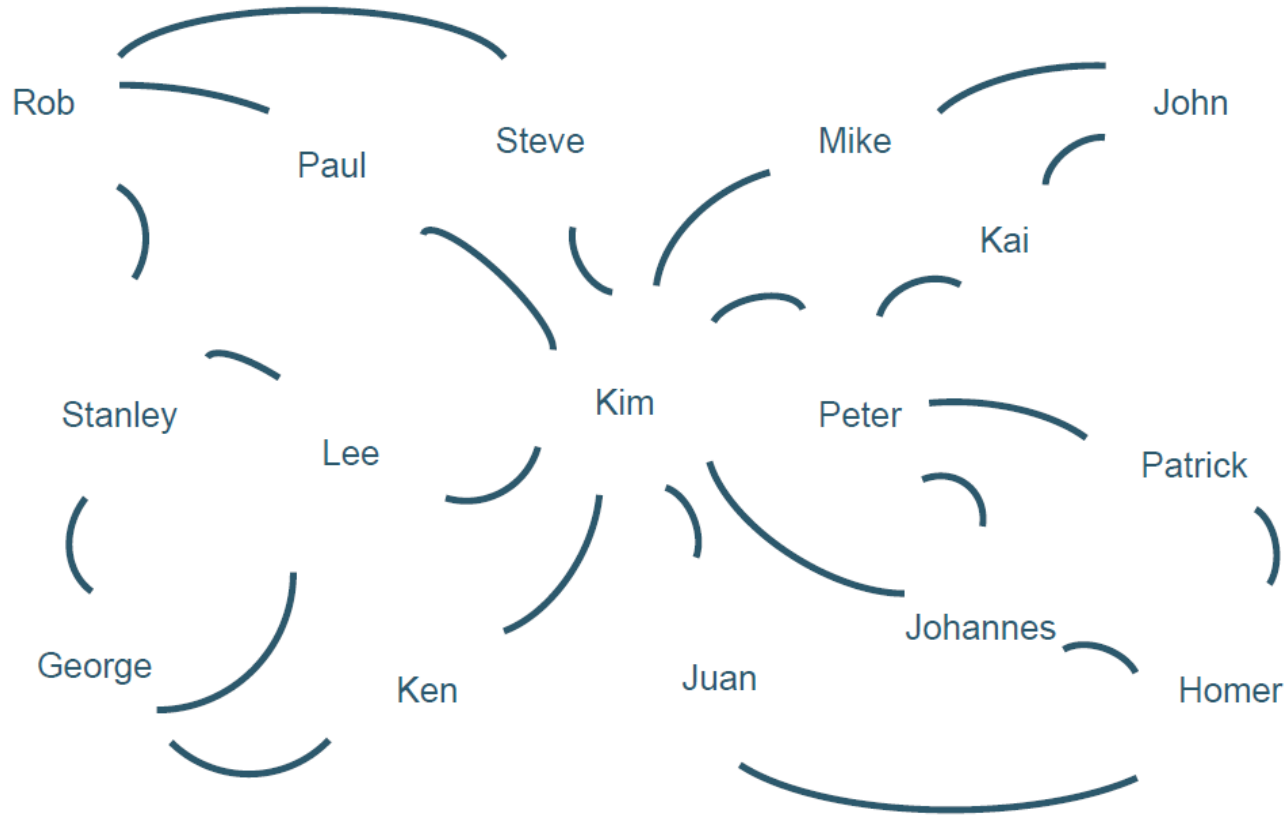
- friendship
- acquaintance
- kinship
- advice
- hindrance
- sex

allow different kind of flows

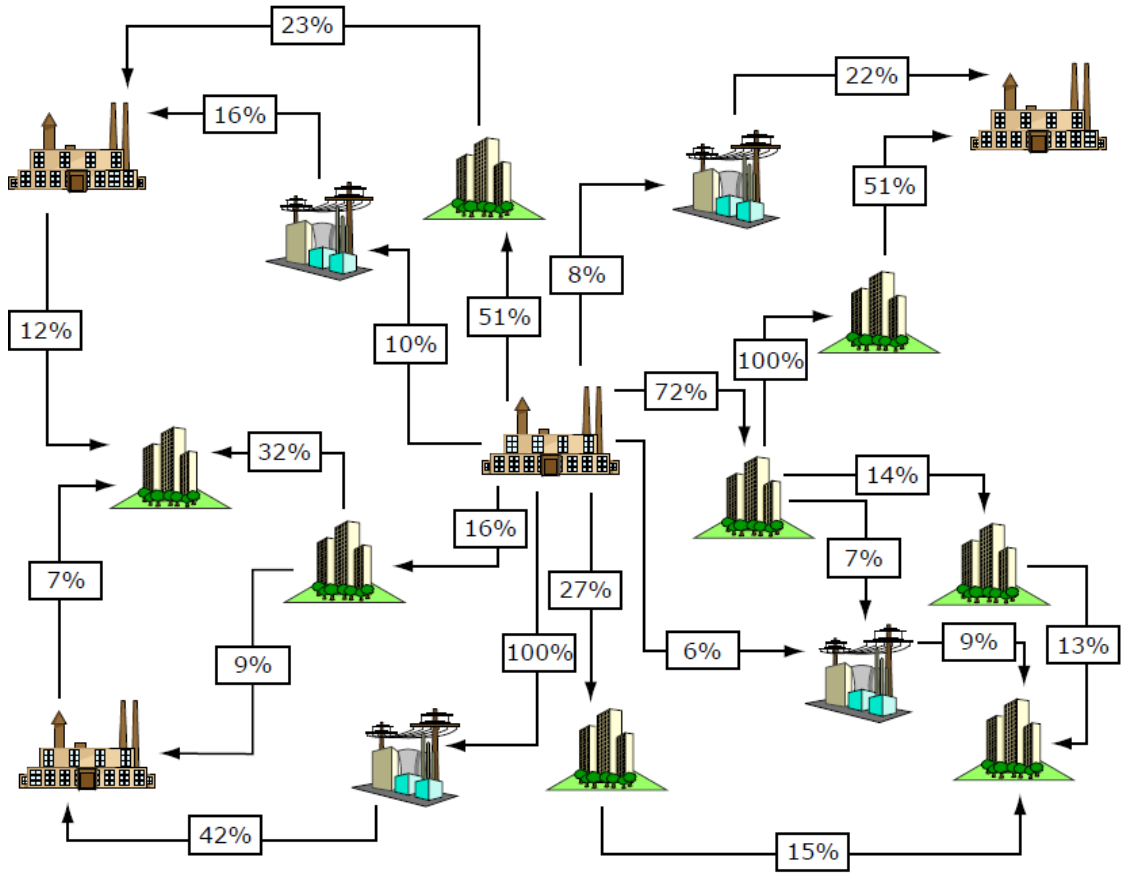
- messages
- money
- diseases



What is a Social Network? - Relations among People



What is a Network? - Relations among Institutions



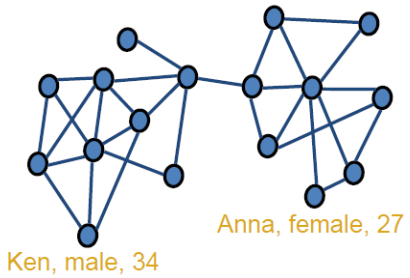
- as institutions
  - owned by, have partnership / joint venture
  - purchases from, sells to
  - competes with, supports
- through stakeholders
  - board interlocks
  - Previously worked for

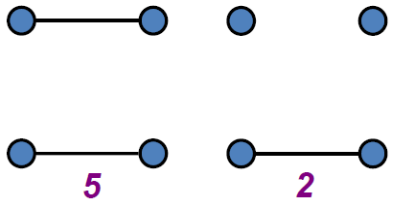
Graphs – Sociograms

(based on Hanneman, 2001)

- Strength of ties:
  - Nominal
  - Signed
  - Ordinal
  - Valued

Strength of a Tie

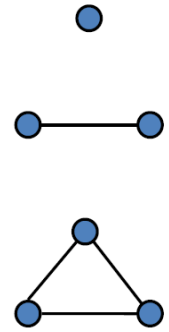




- *Social network*
  - finite set of actors and relation(s) defined on them
  - depicted in *graph/ sociogram*
    - *labeled graph*
- *Strength of a Tie*
  - *dichotomous* vs. *valued*
    - depicted in *valued graph* or *signed graph* (+/-)

Folie: 23

Dyads, Triads and Relations



- *actor*
- *dyad*
- *triad*

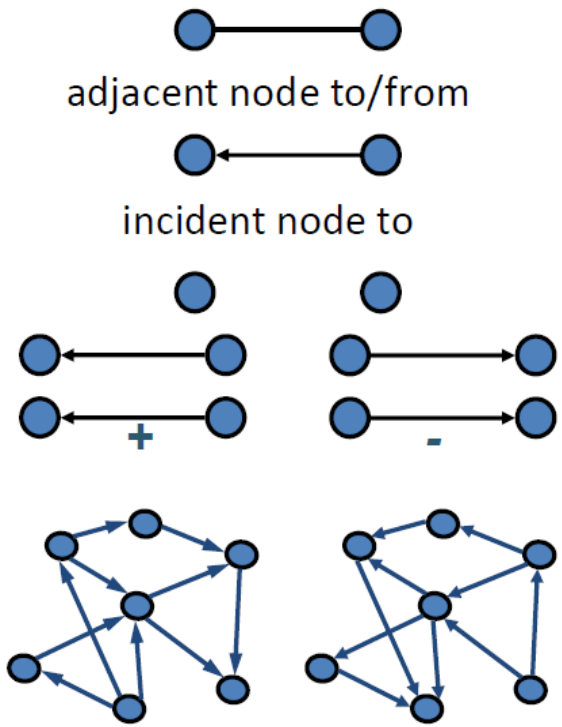
friendship

kinship

- *relation:*
  - collection of specific ties among members of a group

Folie: 22

Strength of a Tie



- *Strength of a Tie*
  - *nondirectional* vs. *directional*
    - depicted in *directed graphs (digraphs)*
    - nodes connected by *arcs*
    - 3 *isomorphism classes*
      - *null dyad*
      - *mutual / reciprocal / symmetrical dyad*
      - *asymmetric / antisymmetric dyad*
    - *converse of a digraph*
      - *reverse direction of all arcs*

Connections

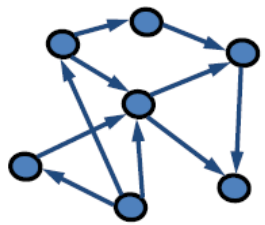
- Size
  - Number of nodes
- Density
  - Number of ties that are present/ the amount of ties that could be present
- Out-degree
  - Sum of connections from an actor to others
- In-degree
  - Sum of connections to an actor

Distance

- Walk
  - A sequence of actors and relations that begins and ends with actors
- Geodesic distance
  - The number of relations in the shortest possible walk from one actor to another
- Maximum flow
  - The amount of different actors in the neighborhood of a source that lead to pathways to a target

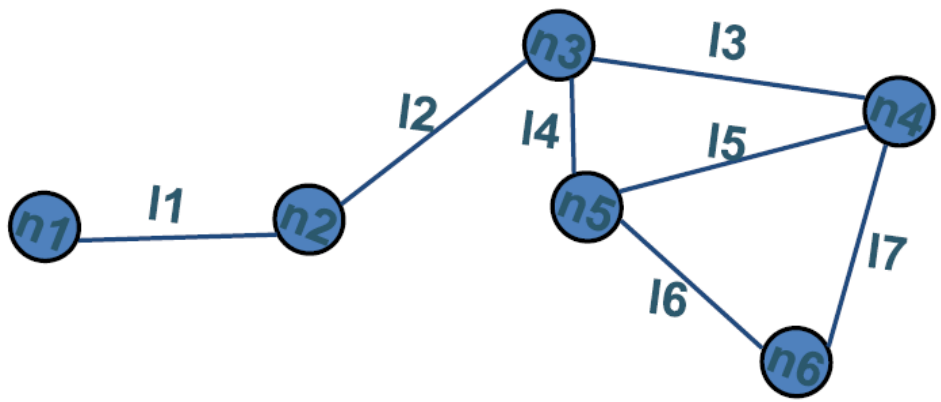
Walks, Trails, Paths

- *(Directed) Walk (W)*
  - sequence of nodes and lines starting and ending with (different) nodes (called *origin* and *terminus*)
  - Nodes and lines can be included more than once
- *Inverse* of a (directed) walk *(W^{-1})*
  - Walk in opposite order
- *Length* of a walk
  - How many lines occur in the walk? (same line counts double, in weighted graphs add line weights)
- *(Directed) Trail*
  - Is a walk in which all lines are distinct
- *(Directed) Path*
  - Walk in which all nodes and all lines are distinct
- Every path is a trail and every trail is a walk



Reachability, Distances and Diameter

- *Reachability*
  - If there is a path between nodes n_i and n_j
- *Geodesic*
  - Shortest path between two nodes
- *(Geodesic) Distance* d(i,j)
  - Length of Geodesic (also called „degrees of separation“)



Some Measures of Power & Prestige

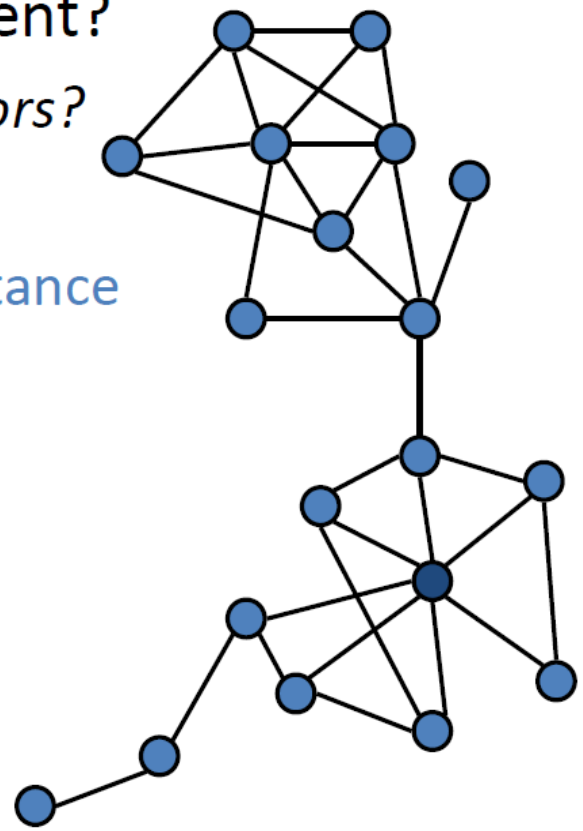
(based on Hanneman, 2001)

- Degree
  - Sum of connections from or to an actor
    - Transitive weighted degree→Authority, hub, pagerank
- Closeness centrality
  - Distance of one actor to all others in the network
- Betweenness centrality
  - Number that represents how frequently an actor is between other actors' geodesic paths

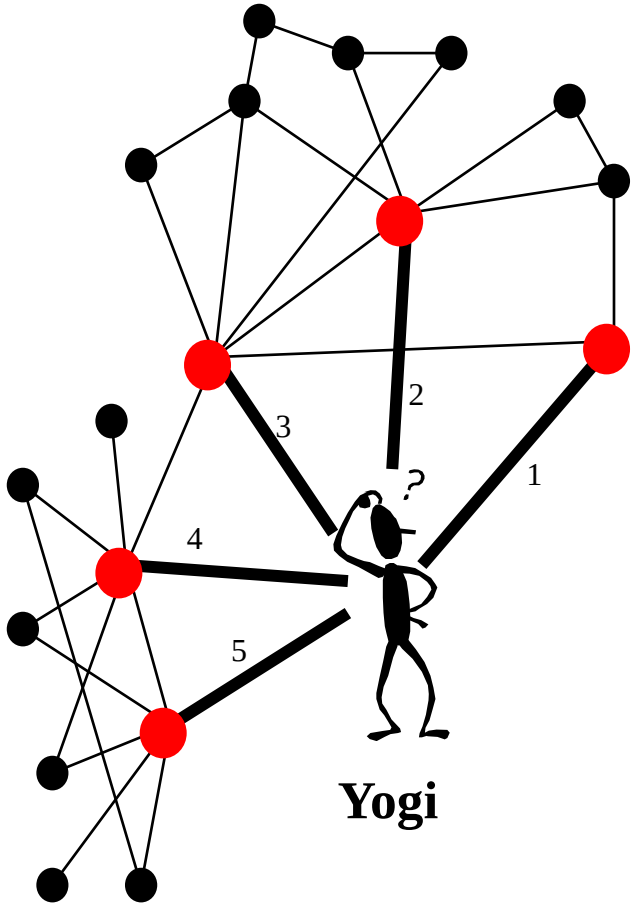
Example: Centrality Measures

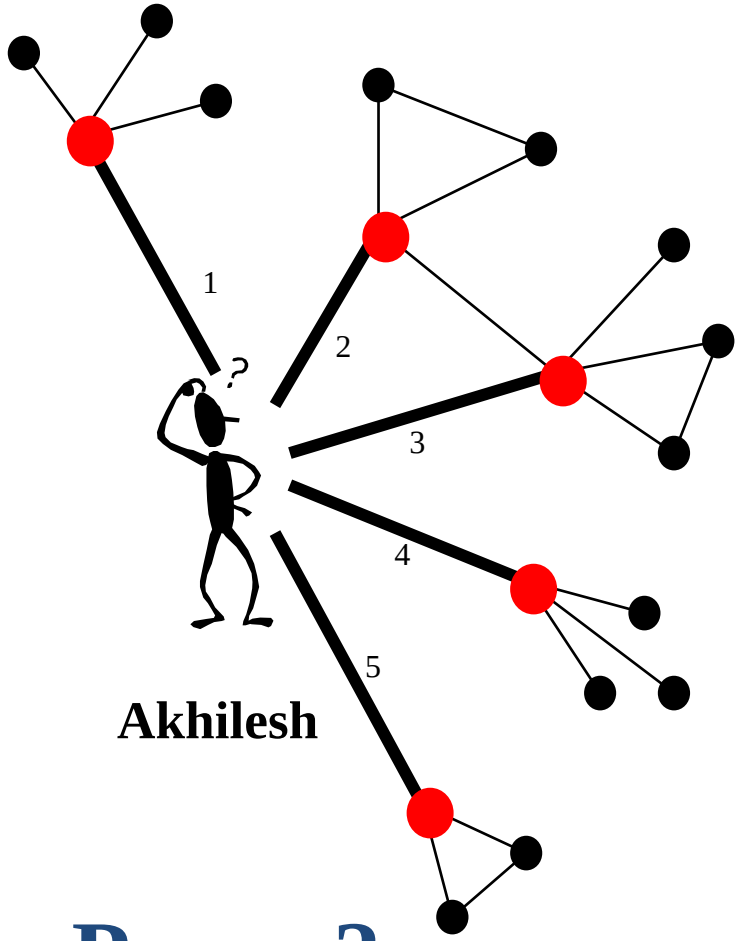
- Who is the most prominent?
  - *Who knows the most actors?*

    (Degree Centrality)
  - Who has the shortest distance to the other actors?
  - Who controls knowledge flows?
  - …



Networks and Power:

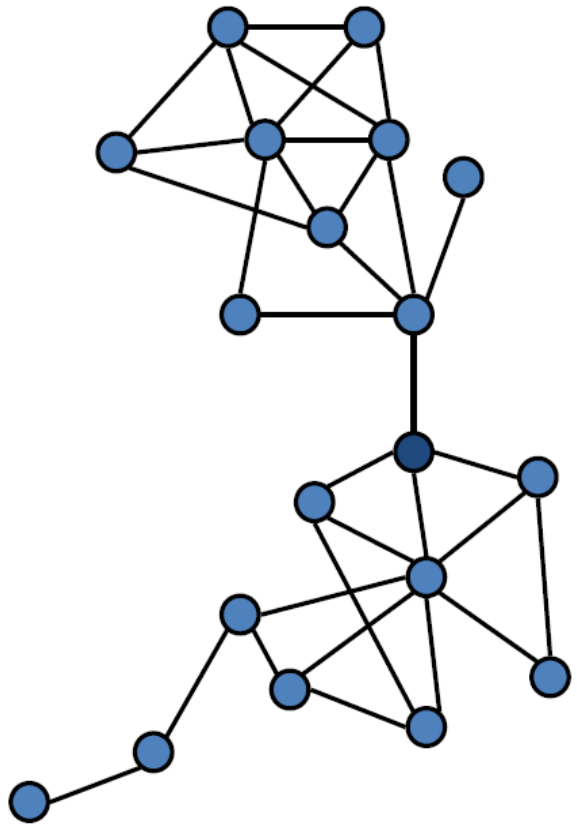




Who has more Power?

Example: Centrality Measures

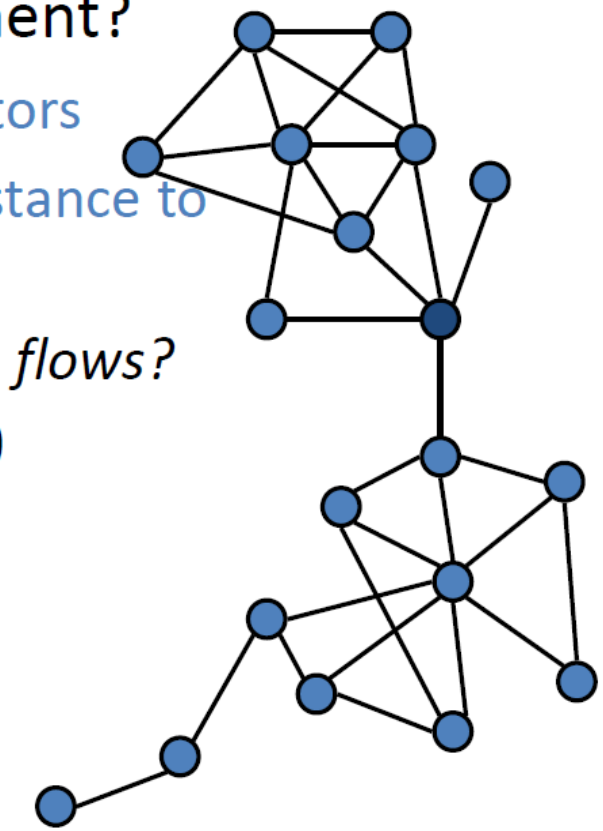
- Who is the most prominent?
  - Who knows the most actors?
  - *Who has the shortest distance to the other actors? (Closesness Centrality)*
  - Who controls knowledge flows?
  - ...



Example: Centrality Measures

- Who is the most prominent?
  - Who knows the most actors
  - Who has the shortest distance to the other actors?
  - *Who controls knowledge flows?*

    (Betweenness Centrality)
  - …



What are 1st, 2nd and 3rd connections in LinkedIn ?

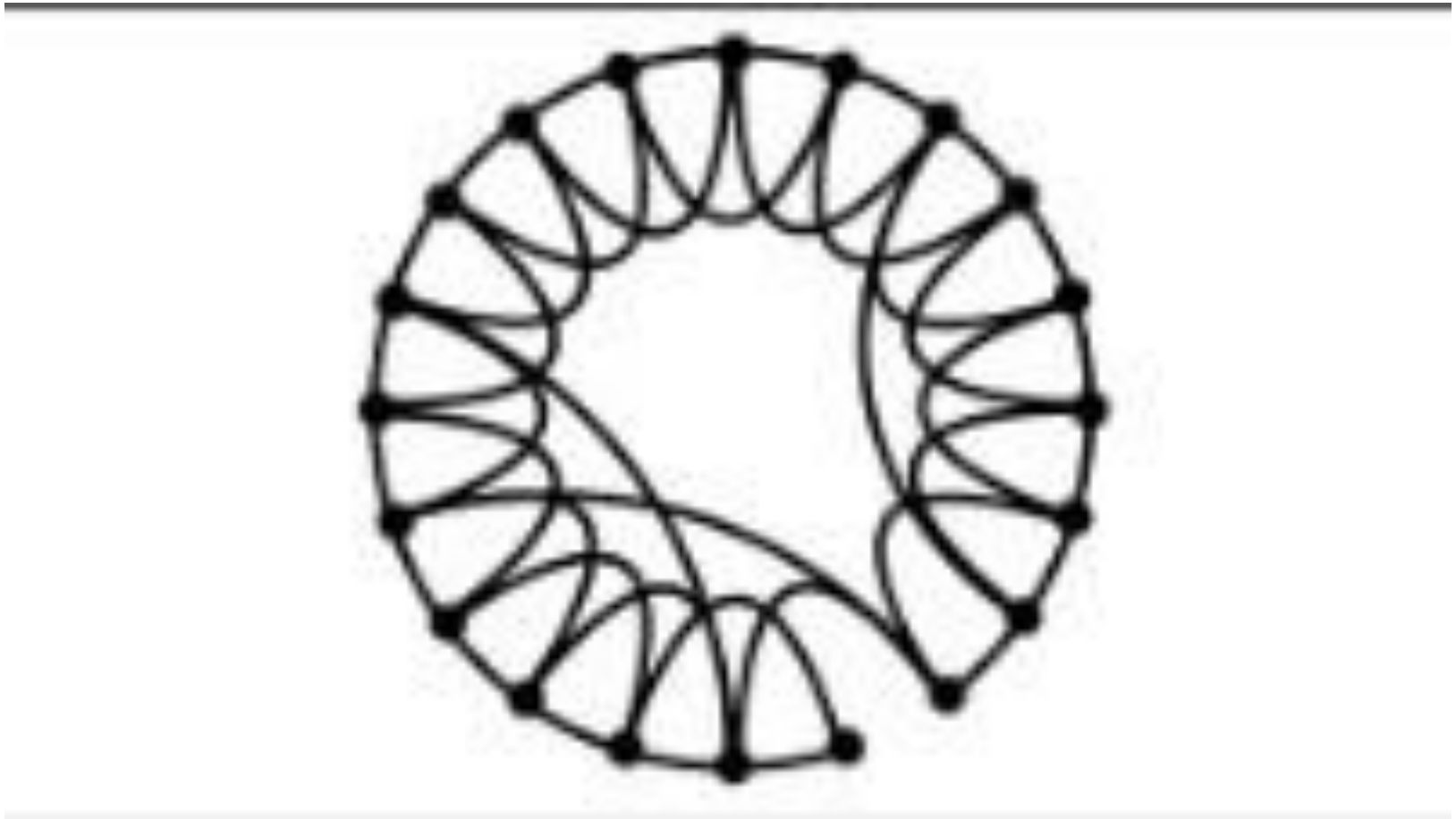
https://www.linkedin.com/feed/

- Many times you people have seen this on LinkedIn that whenever you open your connections, you find 1st, 2nd or 3rd written. Do you know what does that mean and how does LinkedIn does the calculation? I will let you know in detail
- **1st-degree –** People you're directly connected to because you have accepted their invitation to connect, or they have accepted your invitation. You'll see a 1st degree icon next to their name in search results and on their profile. You can contact them by sending a message on LinkedIn.
- **2nd-degree –** People who are connected to your 1st-degree connections. You'll see a 2nd degree icon next to their name in search results and on their profile. You can contact them through an In Mail or an introduction.
- **3rd-degree –** People who are connected to your 2nd-degree connections. You'll see a 3rd degree icon next to their name in search results and on their profile. You can contact them through an In Mail or an introduction.
- **Out of Network –** LinkedIn members who fall outside of the categories listed above. You can contact them through an In Mail.

How the Facebook Algorithm Works in 2021 and How to Make it Work for You

- https://blog.hootsuite.com/facebook-algorithm/

Six Degrees of Separation



Six Degrees of Separation

Milgram (1967)

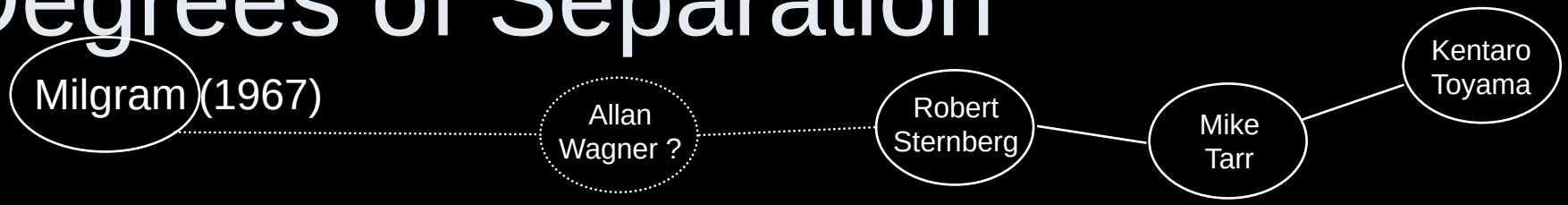
The experiment:

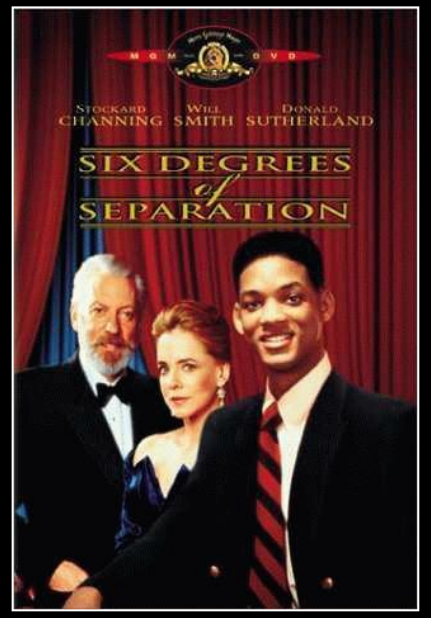
- Random people from Nebraska were to send a letter (via intermediaries) to a stock broker in Boston.
- Could only send to someone with whom they were on a first-name basis.

Among the letters that found the target, the average number of links was six.

Stanley Milgram (1933-1984)

Six Degrees of Separation

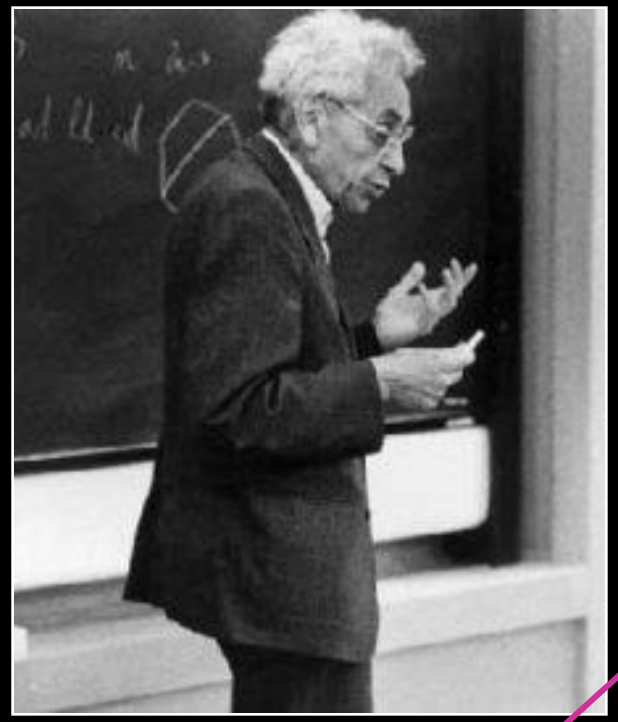




John Guare wrote a play called *Six Degrees of Separation,* based on this concept.

"*Everybody on this planet is separated by only six other people. Six degrees of separation. Between us and everybody else on this planet. The president of the United States. A gondolier in Venice… It's not just the big names. It's anyone. A native in a rain forest. A Tierra del Fuegan. An Eskimo. I am bound to everyone on this planet by a trail of six people…*"

Erdős Number
(Bacon game for Brainiacs ☺)



Paul Erdős (1913-1996)

Unlike Bacon, Erdos has better centrality in his network

Number of links required to connect scholars to Erdős, via co-authorship of papers

Erdős wrote 1500+ papers with 507 co-authors.

Jerry Grossman's (Oakland Univ.) website allows mathematicians to compute their Erdos numbers:

http://www.oakland.edu/enp/

Connecting path lengths, among mathematicians only:

- average is *4.65*
- maximum is *13*

Erdős Number

An Example

Paul Erdős

Alon, N., P. Erdos, D. Gunderson and M. Molloy (2002). On a Ramsey-type Problem. *J. Graph Th.* 40, 120-129.

Mike Molloy

Achlioptas, D. and M. Molloy (1999). Almost All Graphs with 2.522 n Edges are not 3-Colourable. *Electronic J. Comb.* (6), R29.

Dimitris Achlioptas

Achlioptas, D., F. McSherry and B. Schoelkopf. Sampling Techniques for Kernel Methods. *NIPS 2001*, pages 335-342.

Bernard Schoelkopf

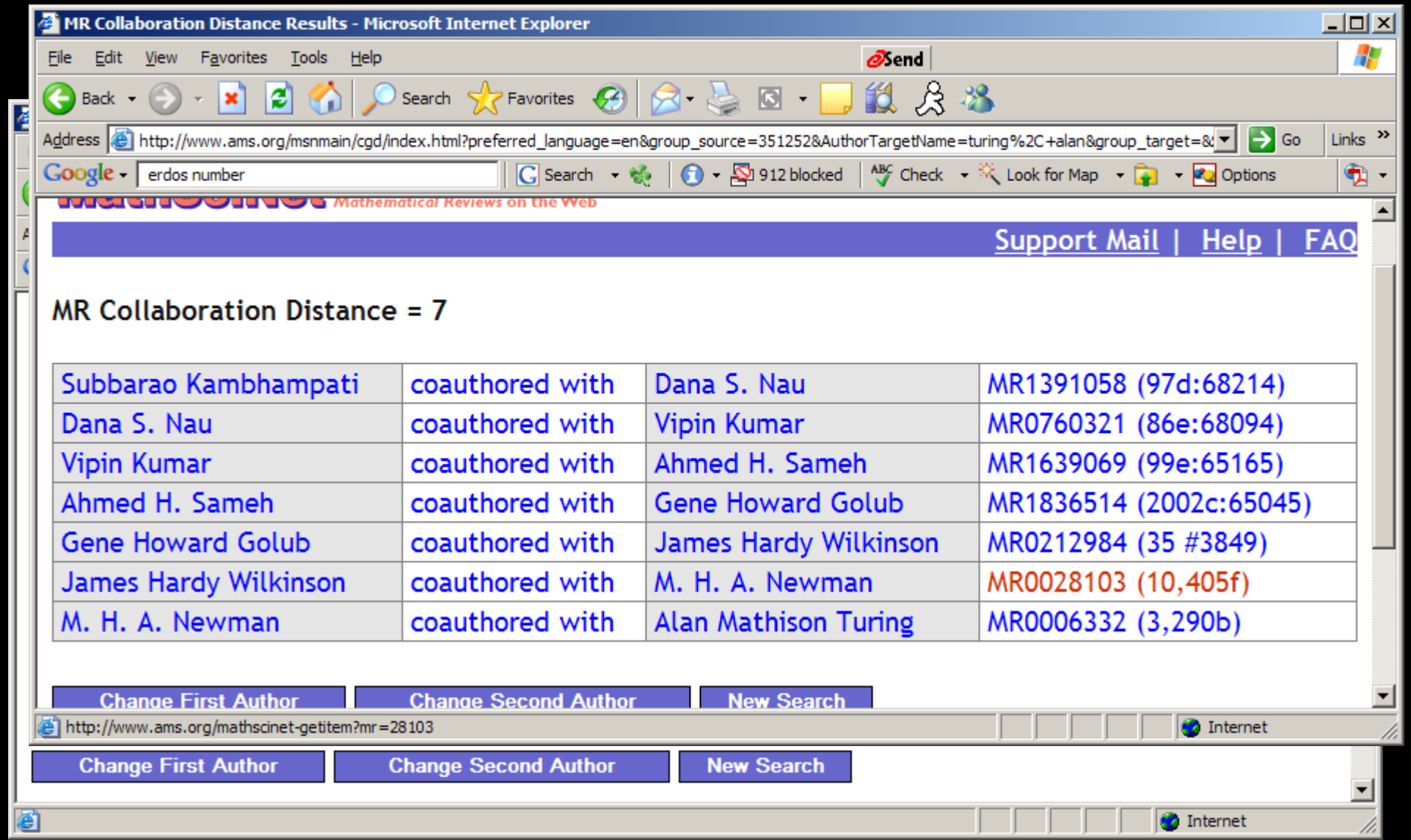
Romdhani, S., P. Torr, B. Schoelkopf, and A. Blake (2001). Computationally efficient face detection. In *Proc. Int'l. Conf. Computer Vision*, pp. 695-700.

Andrew Blake

Toyama, K. and A. Blake (2002). Probabilistic tracking with exemplars in a metric space. *International Journal of Computer Vision*. 48(1):9-19.

Kentaro Toyama

..and Rao has even shorter distance ☺



The Kevin Bacon Game

Boxed version of the Kevin Bacon Game

Invented by Albright College students in 1994:

- Craig Fass, Brian Turtle, Mike Ginelly

Goal: Connect any actor to Kevin Bacon, by linking actors who have acted in the same movie.

Oracle of Bacon website uses Internet Movie Database (IMDB.com) to find shortest link between any two actors:

http://oracleofbacon.org/

The Kevin Bacon Game

An Example

Kevin Bacon

Mystic River (2003)

Tim Robbins

Code 46 (2003)

Om Puri

Yuva (2004)

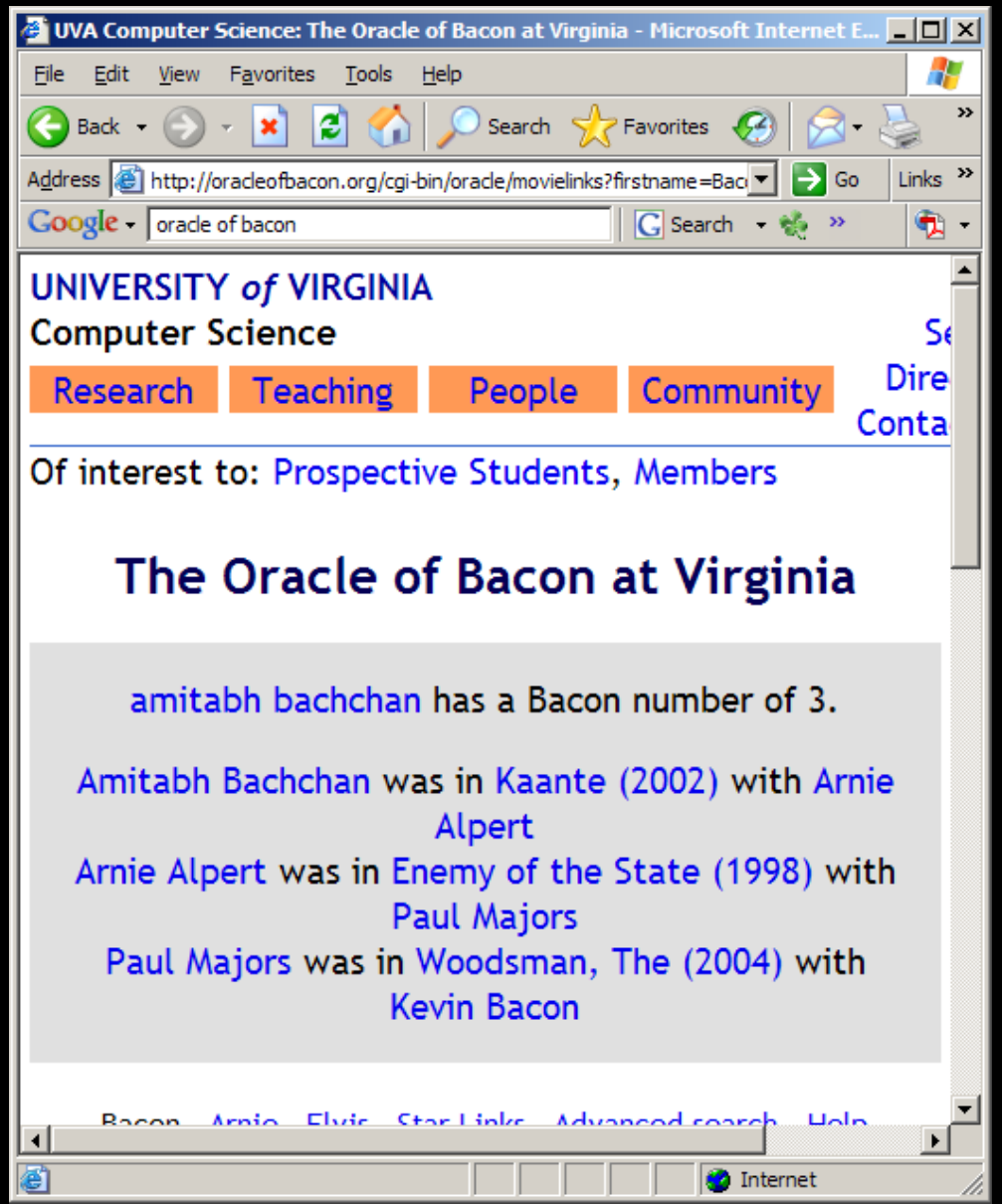
Rani Mukherjee

Black (2005)

Amitabh Bachchan

…actually Bachchan has a Bacon number 3

- Perhaps the other path is deemed more diverse/ colorful…



The Kevin Bacon Game

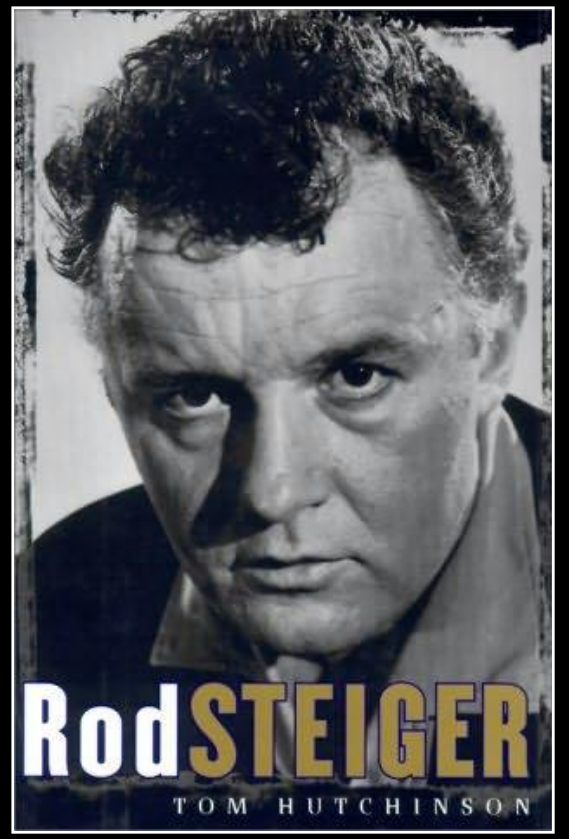
Total # of actors in database: *~550,000*

Average path length to Kevin: *2.79*

Actor closest to “center”: *Rod Steiger (2.53)*

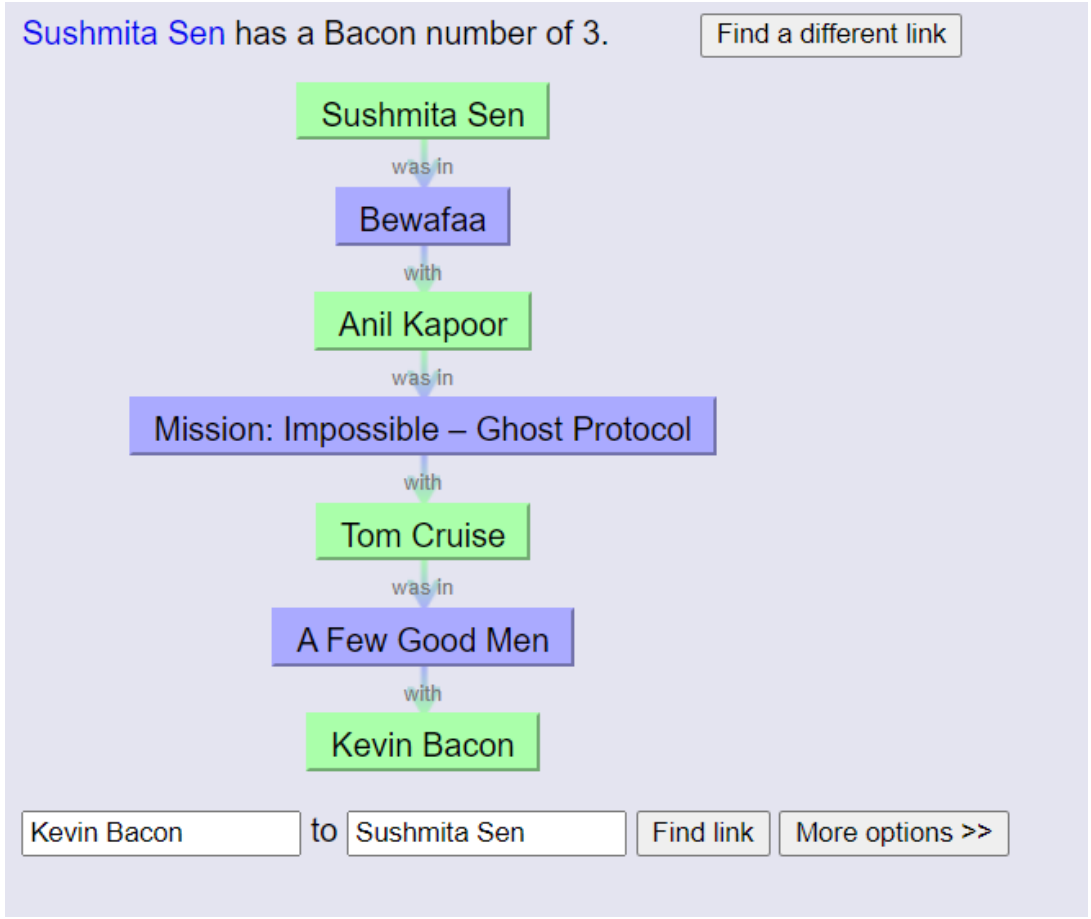
Rank of Kevin, in closeness to center: *876th*

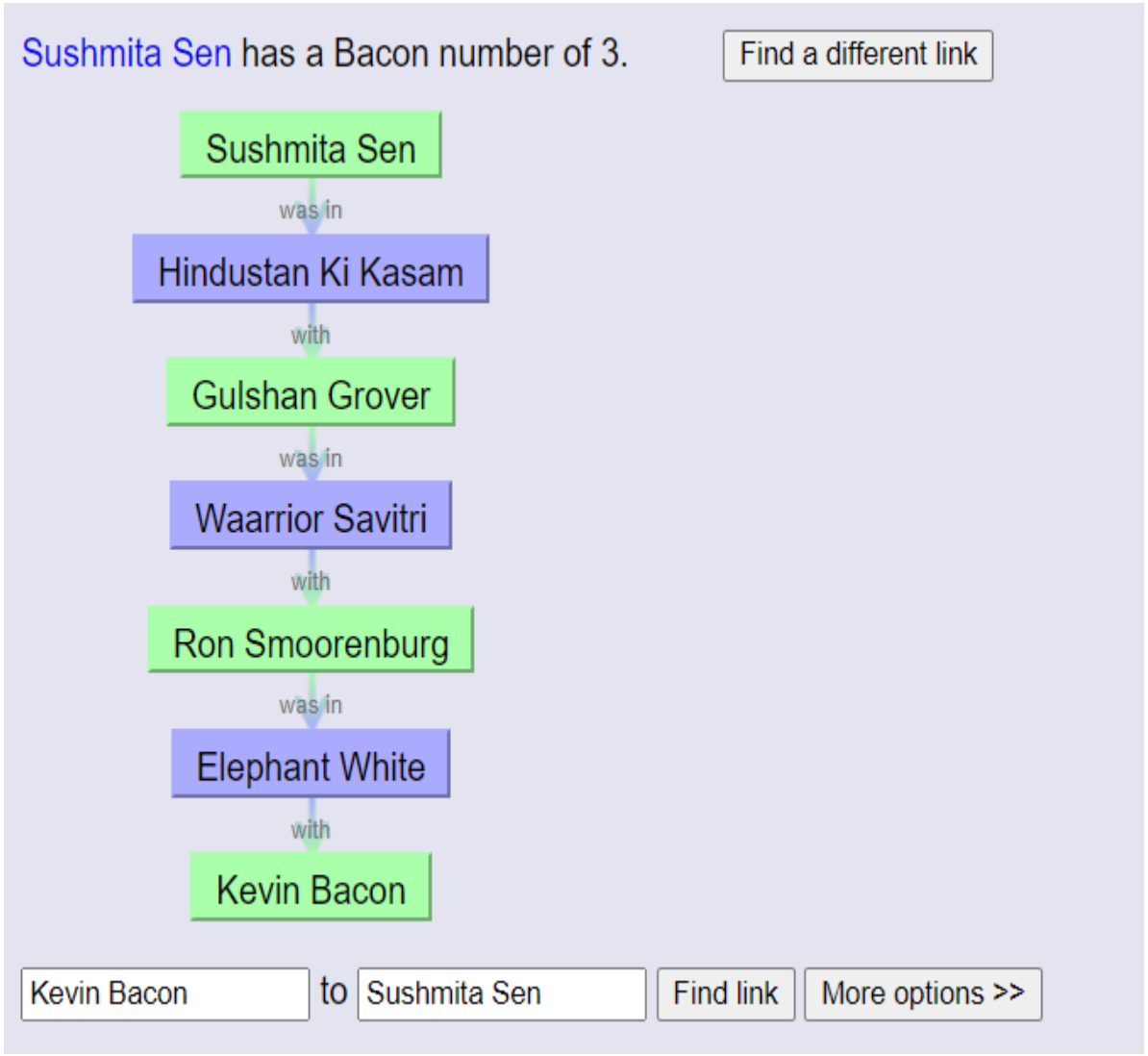
Most actors are within three links of each other!



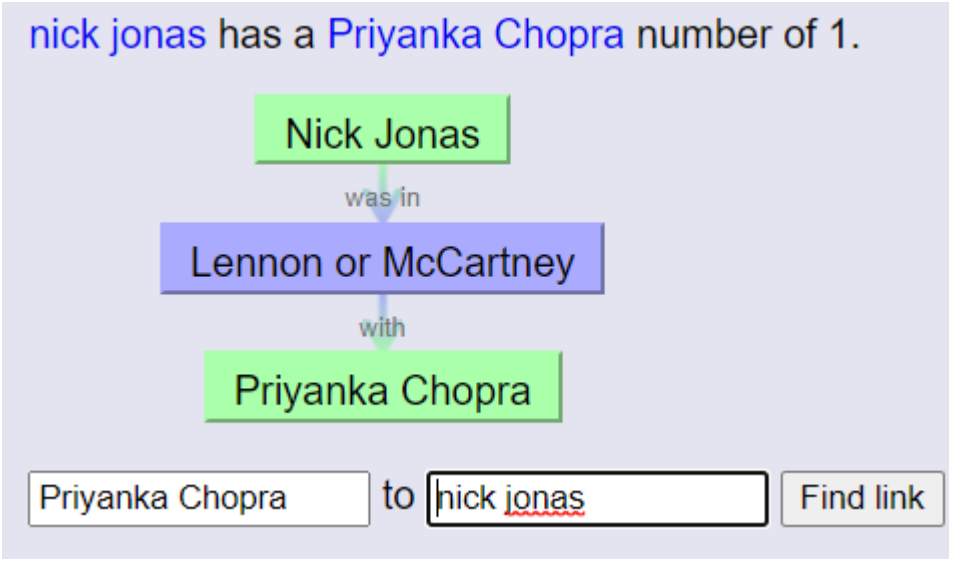
Center of Hollywood?

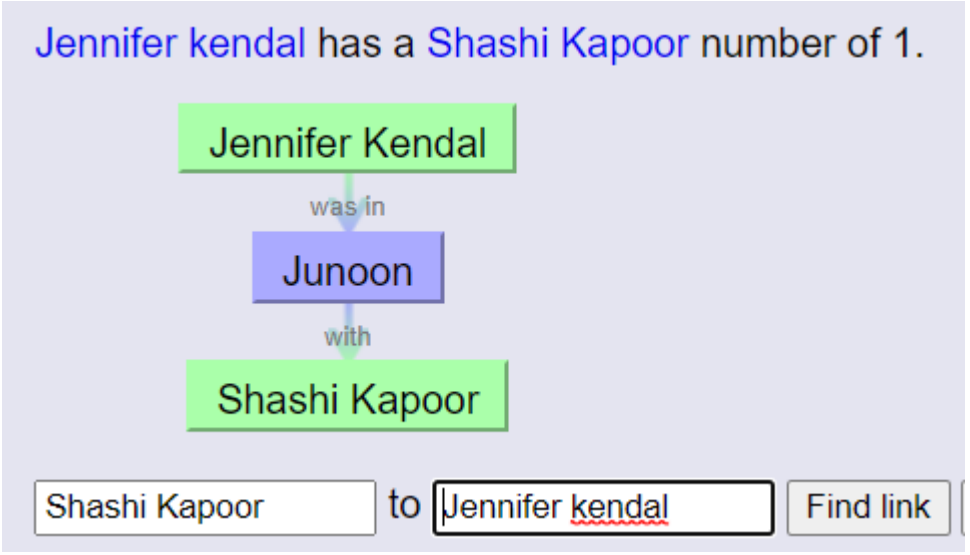
The Oracle of Bacon





The Oracle of Bacon

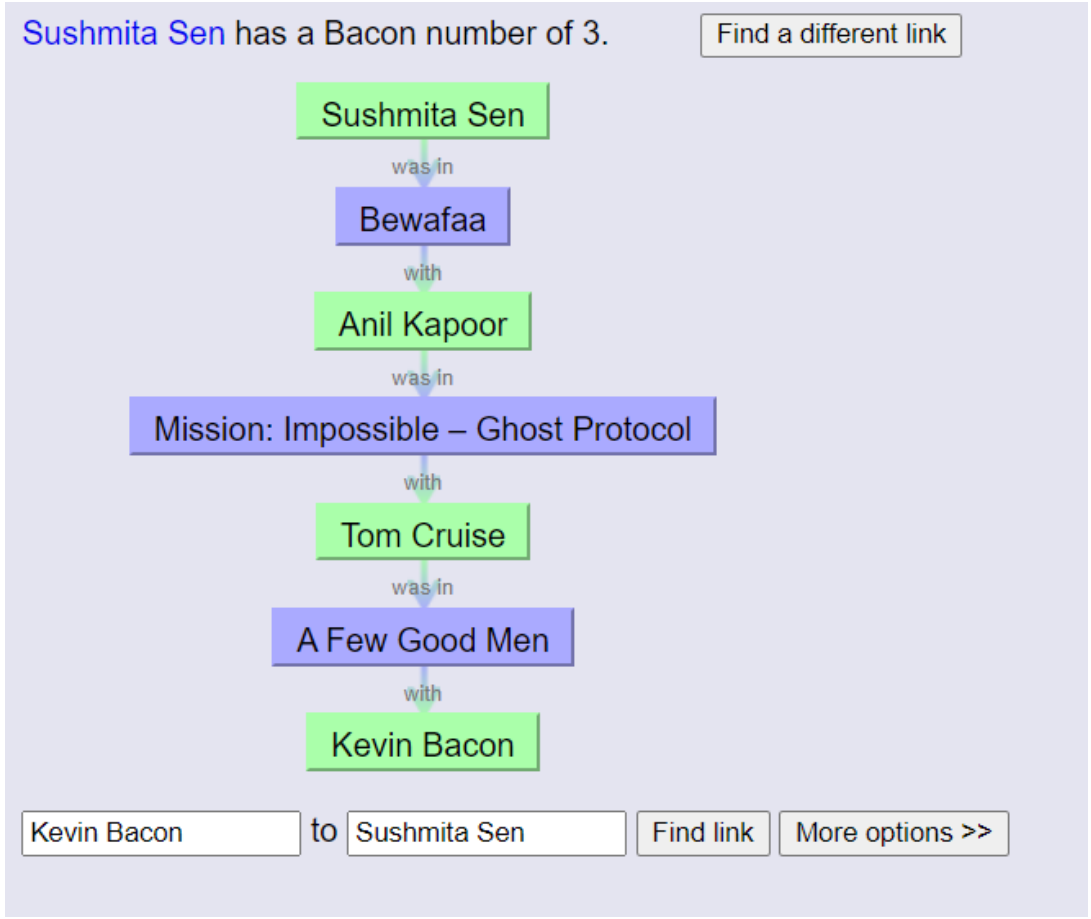


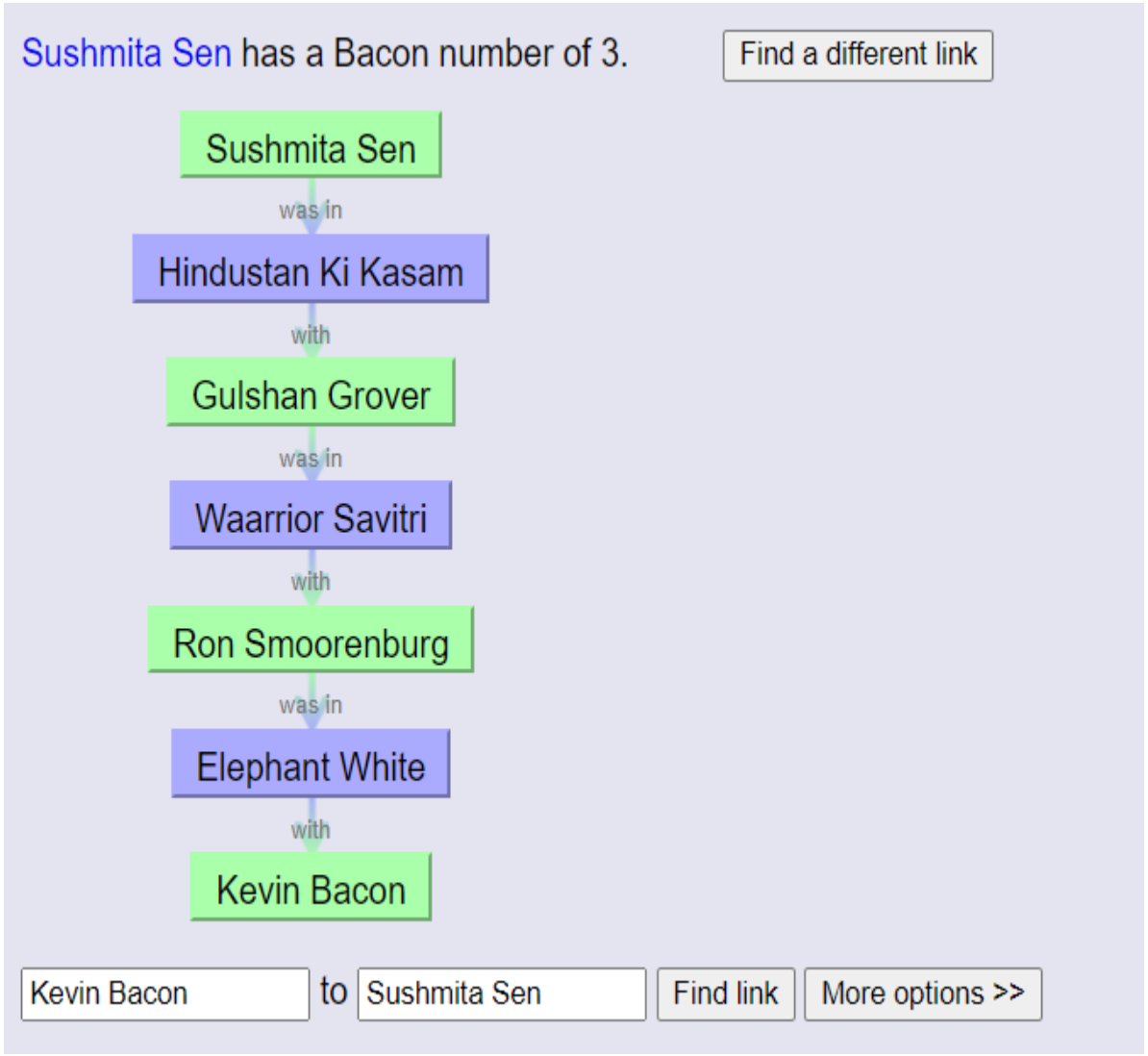


The Oracle cannot find *gene goodenough*. You have probably misspelled your entry, but it's also possible that the actor or actress you seek is not in our database. The Oracle didn't find any close matches, but you may have better luck searching at wikipedia.org.

Preity Zinta to gene goodenough Find link More options >>

The Oracle of Bacon

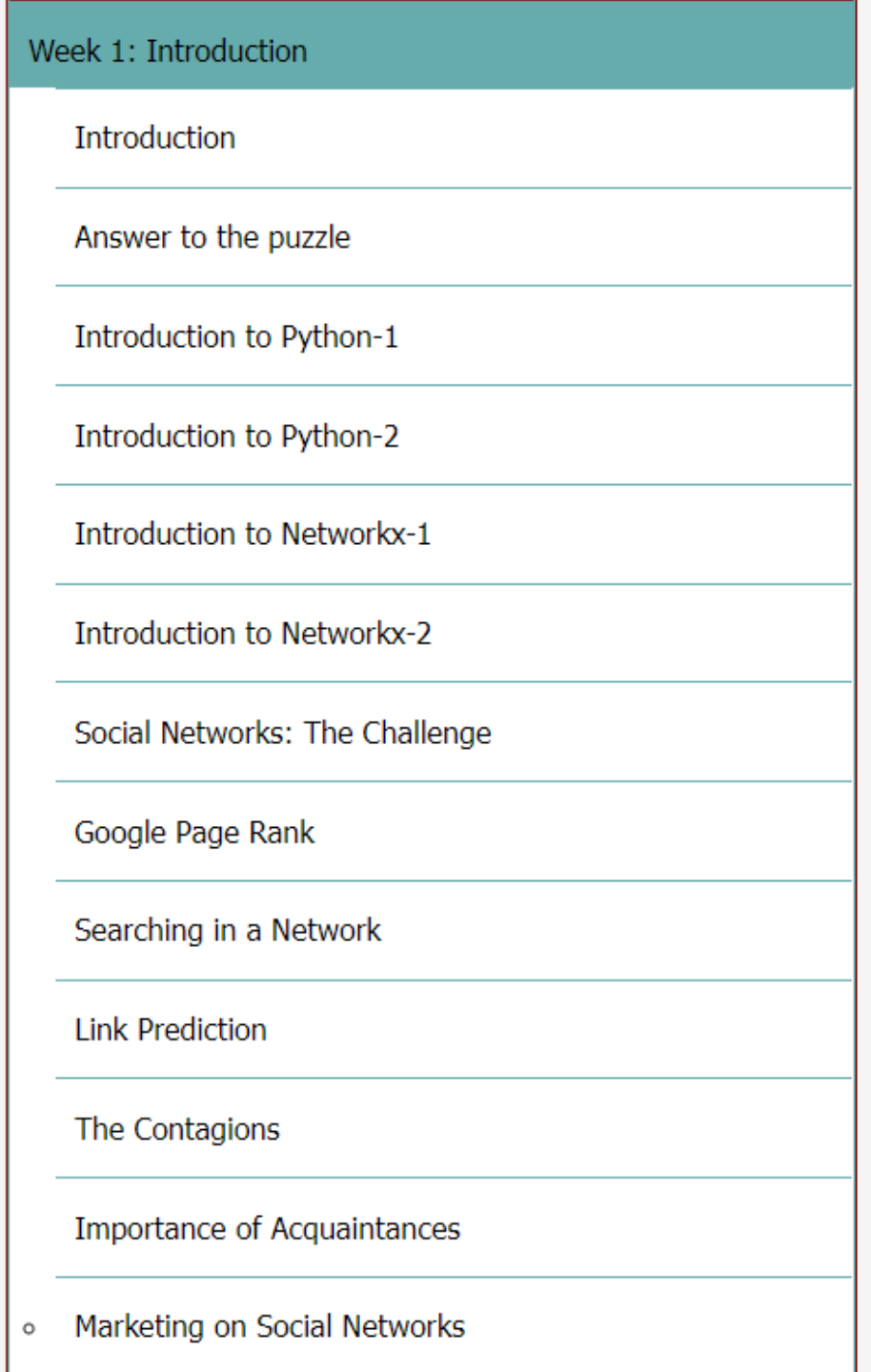




Work to Do?

- https://nptel.ac.in/courses/106/106/106106169/#

https://networkx.org/documentation/stable/tutorial.html



Thank You!