Using *Social Network Analysis* in Philosophy

Catherine Herfeld
(with M. C. Doehne – ZU Friedrichshafen)

Summer School on Mathematical Philosophy for Female Students
Munich 2014

This summer school

- At the beginning, you heard about:
  - the mathematician's way;
  - the physicist's way:
    → by modeling and simulation.
- Those who chose Conor's tutorials:
  - know that social networks are used in social epistemology;
  - have heard about social networks in the context of agent-based modeling.

This summer school

- Social/formal epistemology:
  - Optimal network structure for consensus, reliability, etc. (e.g. Zollman 2012);
  - Performance of different scientific communities.
- Moral philosophy:
  - Evolution of moral behavior in society (Alexander 2006).
- Philosophy of (the social) sciences:
  - Reflection about use of network theory in social sciences (e.g. Marchionni 2013).

This summer school

The social scientist's way:

- Economics:
  - How network structure influences consensus and reliability of learning (Golub and Jackson 2012);
- Sociology of science:
  - Diffusion of knowledge;
  - Characteristics of knowledge communities;
  - Co-existence of sub-communities (network clusters) and their relation.
- History of science:
  - Emergence of entire disciplines (e.g. Claveau 2014).

This summer school

How to study social network influences on behavior?

Depends upon question!

- Models and simulations, e.g. agent-based models and network effects;
  - Example: Analyzing the effect of network structure on diffusion.
- **Empirical analysis;**
  - Example: What potential features increase the likelihood of adoption (and diffusion) of an innovation through the social network.
- Often: a combination of both!

Agenda

1. What is Social Network Analysis (SNA)?
2. Science as social enterprise
3. Studying science with SNA
4. Conceptual framework for a SNA
5. An Example: The diffusion of theories
6. Challenges of application and of interpretation
7. Implications
8. Some further readings

The social scientist's way

We live in a connected world:

"If we ever get to the point of charting a whole city or a whole nation, we would have […] a picture of a vast solar system of intangible structures, powerfully influencing conduct, as gravitation does in space. Such an invisible structure underlies society and has its influence in determining the conduct of society as a whole"

Jacob L. Moreno, *New York Times*, April 13, 1933

→ Patterns of connection form a social space, that can be seen in multiple contexts.

Science as social enterprise

CERN:

- Operates world's largest lab for particle physics;
- More than 2300 full-time employees;
- 1,500 part-time employees;
- 10,000 visitors and engineers;
- 21 member states that finance and otherwise support enterprise;
- Numerous non-European countries involved.

Science as social enterprise

Science (2014) Vol. 344, Issue 6191

Formation temperatures of thermogenic and biogenic methane
D. A. Stolper, M. Lawson, C. L. Davis, A. A. Ferreira, E. V. Santos Neto, G. S. Ellis, M. D. Lewan, A. M. Martini, Y. Tang, M. Schoell, A. L. Sessions, and J. M. Eiler

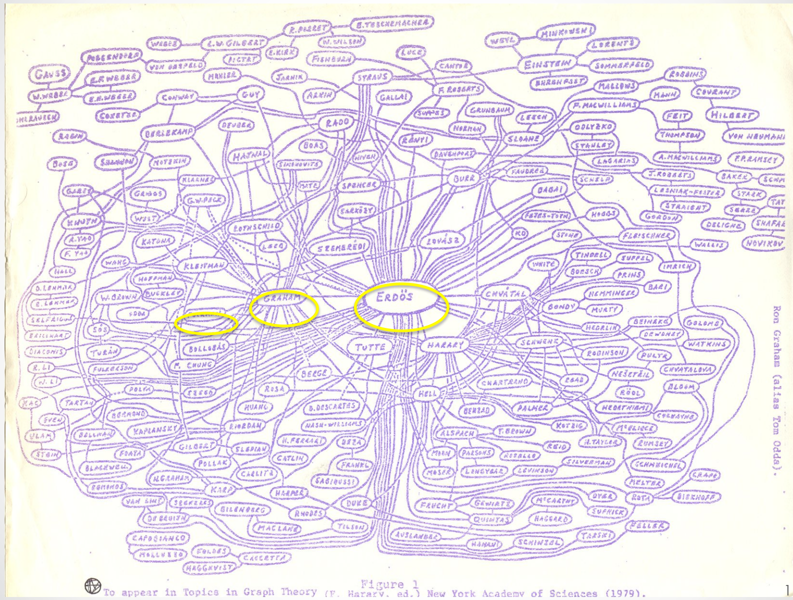
Ediacaran metazoan reefs from the Nama Group, Namibia
M. Penny, R. Wood, A. Curtis, F. Bowyer, R. Tostevin, and K.-H. Hoffman

Lassa virus entry requires a trigger-induced receptor switch
Lucas T. Jae, Matthijs Raaben, Andrew S. Herbert, Ana I. Kuehne, Ariel S. Wirchnianski, Timothy K. Soh, Sarah H. Stubbs, Hans Janssen, Markus Damme, Paul Saftig, Sean P. Whelan, John M. Dye, and Thijn R. Brummelkamp

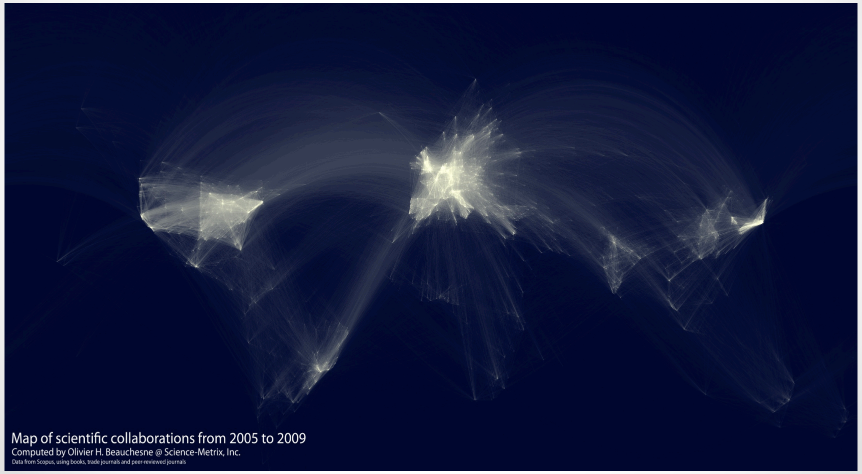
Nucleoside diphosphate kinases fuel dynamin superfamily proteins with GTP for membrane remodeling
Mathieu Boissan, Guillaume Montagnac, Qinfang Shen, Lorena Griparic, Jérôme Guitton, Maryse Romao, Nathalie Sauvonnet, Thibault Lagache, Ioan Lascu, Graça Raposo, Céline Desbourdes, Uwe Schlattner, Marie-Lise Lacombe, Simona Polo, Alexander M. van der Bliek, Aurélien Roux, and Philippe Chavrier

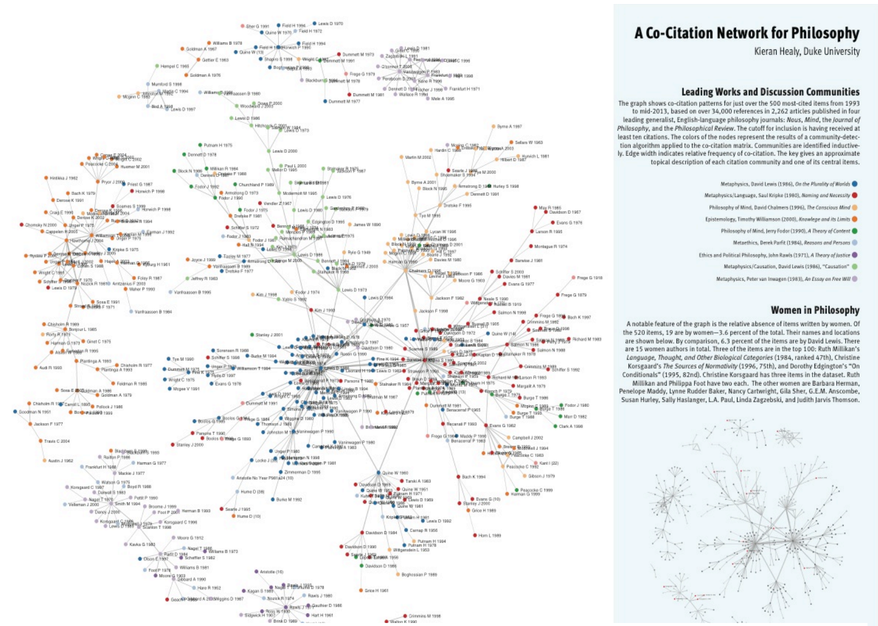
Science as social enterprise

In short:
- Science is socially embedded activity;
- Scientific practices characterized by:
  - Culture;
  - Norms and values;
  - Social relations.
- Social structure shapes scientific endeavor:
  - Position in scientific community;
  - Being (or not being) part of a network.
- Need for methods from social sciences to analyze scientific practices.



Science as social enterprise





http://kieranhealy.org/blog/archives/2013/06/18/a-co-citation-network-for-philosophy/

What is SNA?

- Social Network Analysis (SNA):

  Social network analysis provides a precise way to define important social concepts, a theoretical alternative to the assumption of independent social actors, and a framework for testing theories about structured social relationships. (Wasserman/Faust 1995)

  - Draws on graph theory.
  - Advanced by new computational methods developed from the 1970s and 1980s onwards;
  - allows to quantitatively represent and study (social) relations;
  - grounded in empirical data;
- Basic idea:
  - social structure matters to explain human behavior;
  - social structure can be modeled as social networks;
  - social networks consist of:
    - *positions* of agents;
    - *relations* between agents.

What is SNA?

- General attributes:
  - Comprehensive family of analytical strategies and tools;
  - Systematic analysis of transactions in patterned social relationships;
  - Focuses on relative positions of actors/objects in (social) network.
- Basic challenges of implementation:
  1. Social networks cut across groups, communities, and other entities;
  2. Transactions unfolding within social networks are not necessarily symmetrical in nature;
  3. Social networks evolve in time.

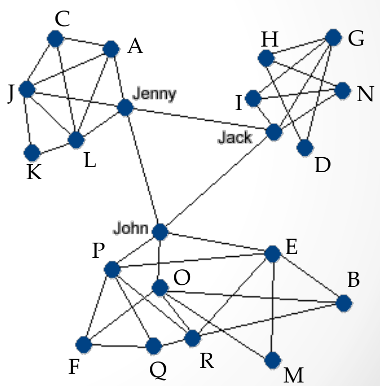
What is SNA?

- Specific focus:
  - studies graphs as representations of symmetric/asymmetric relations;
  - graph is a structure revealed by links between objects;
  - links represent kinds of relationships, e.g.
    - friendships (Ennett and Bauman 1993);
    - scientific collaboration (Newman 2004);
- Specific components:
  - set of nodes (any sort of *object*);
  - set of edges (any sort of *relationship* between objects).
- Specific objectives:
  - to represent and measure structural relations through network topology;
  - to explain occurrence of structural relations;
  - understand consequences of structural relations.

What is SNA?

Illustration of a simple social network:

- Simplified representation that reduces social system to an abstract structure;
- Tools are used to analyze the characteristics and dynamics of this structure;
- Analysis tells us something about social relations and social categories;
- 21 nodes; 40 edges;
- <u>Examples</u>: internet, market interactions, friendship network, etc.;
- <u>Note</u>: networks are not necessarily social.



What is SNA?

Representation of social relations:

- Relational data is stored in matrices:
  - Arrangement of a set of elements (i.e. two-dimensional);
    - Elements arranged in rows and columns;
    - Elements in our context: measurement of relations (data);
  - *nxn Adjacency matrix:* contains who is next to, or adjacent to, whom in "social space".
  - Use of computational algorithms to find, represent, and analyze those social relations.

What is SNA?

<u>Simple (symmetric) matrix (X) presenting a binary network as undirected graph</u>

	Jenny	Jack	John	A	B	C	D	...	R
Jenny	-	1	1	1	0	0	0	0	
Jack	1	-	1	0	0	0	0	0	
John	1	1	-	0	0	0	0	1	
A	1	0	0	-	0	1	0	0	
B	0	0	0	0	-	0	0	1	
C	0	0	0	1	0	-	0	0	
D	0	0	0	0	0	0	-	0	
⋮	0	0	1	0	1	0	0	-	
R	...	...	...	...	...	...	...	...	$x_{m,n}$

- Each node is listed on row (sender) *and* column (receiver).
- The ith row and the jth column (X_{ij}) record value of edge from node i to node j.
- It is *binary* because values are only 1 (related) and 0 (not related).
- Matrix of *undirected* graph: edges 'sent' are same as edges 'received', so that every entry above the diagonal equals entries below diagonal (i.e. symmetric matrix).
- E.g. *who knows whom.*

What is SNA?

<u>Simple (asymmetric) matrix (X) presenting a valued network as nondirected graph</u>

	Jenny	Jack	John	A	B	C	D	...	R
Jenny	-	1	2	1	0	0	0	0	
Jack	1	-	1	0	0	0	0	0	
John	1	1	-	0	0	0	0	1	
A	3	0	0	-	0	1	0	0	
B	0	0	0	0	-	0	0	1	
C	0	0	0	3	0	-	0	0	
D	0	0	0	0	0	0	-	0	
⋮	0	0	1	0	1	0	0	-	
R	...	...	...	...	...	...	...	...	$x_{m,n}$

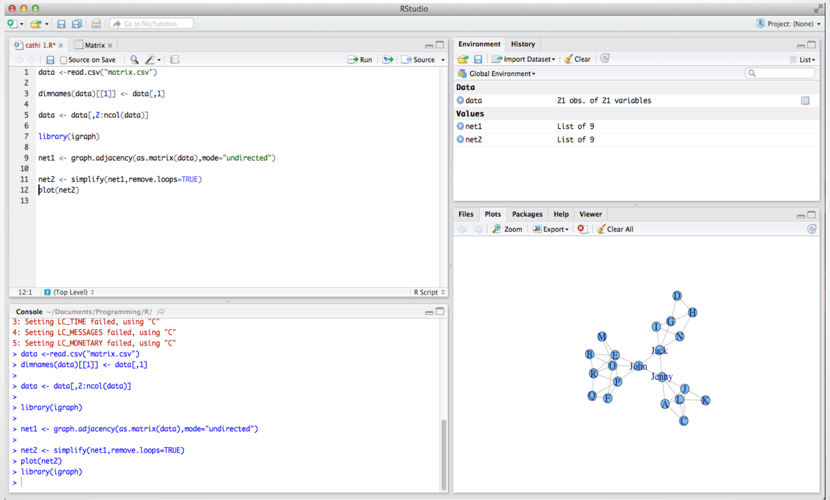
- You can also represent strength of a tie;
- Matrix is *valued*, i.e. specific value is attributed to a relation;
- Matrix of *directed* graph: edges 'sent' are <u>not</u> same as edges 'received' (i.e. asymmetric matrix).
- E.g. *who contacts whom.*

What is SNA?

- Computer software, e.g.:
  - Gephi
  - Pajek
  - UCINET
  - **R**

  → All used for different purposes!

What is SNA?



Conceptual framework

- Steps of empirical SNA:
  1. Specify general research question and general hypothesis;
  2. Formulate empirical hypothesis;
  3. Define a 'case study'
  4. Identify conceptual variables
  5. Identify measurable indicators for conceptual variables;
  6. Collect your data;
  7. Construct the network;
  8. Draw (theoretically informed) inferences with regard to research question.

Example: diffusion of theories

Step 1: Specify general research question and general hypothesis

- General observation:
  - (Social) network structures affect the diffusion of innovations
    - (e.g. for many: Rogers 2003).
  - Does this hold for the diffusion of theories as well?
- General question:
  - What are the factors that lead members of a population to adopt one idea (innovation) and not another?
- General hypothesis:
  - Adoption of an innovation is (partly) explained by *network exposure*.
- Premise:
  - Innovations spread through interpersonal contacts that largely consist of interpersonal communication

Example: diffusion of theories

Step 2: Formulate empirical hypotheses

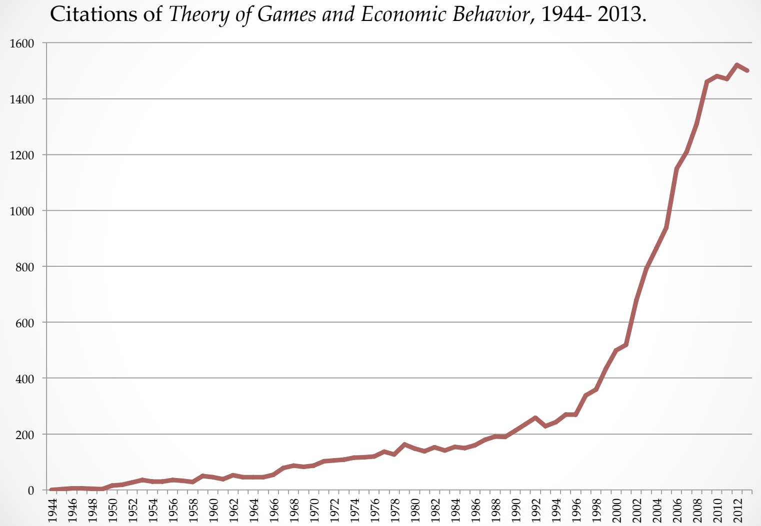
- Specific idea:
  - Social network structures influence the adoption of new scientific theories within and across scientific communities.
- Research question:
  - What leads scientists to adopt a new scientific theory and to not adopt another?
- Empirical hypothesis:
  - Adoption of a new scientific theory is (partly) explained by exposure of scholars to the theory.

Example: diffusion of theories

Step 3: Define a 'case study'

- Case study:
  - Axiomatic theories of rational decision-making.
- Pioneers:
  - John von Neumann and Oskar Morgenstern's *Theory of Games and Economic Behavior* (1944);
- Contribution:
  - Introduced new mathematical tools (i.e. theory of convex sets, mathematical logic, topology) into the social sciences.

Oskar Morgenstern and John von Neumann at Spring Lake, 1949.
Courtesy of the Institute for Advanced Study, Princeton



Source: Google scholar

Example: diffusion of theories

Step 4: Identify conceptual variables

- Narrow focus here:
  - How to conceptualize the diffusion of a *theory* (conceptualized as innovation) ?
  - How do we link the content topology to actors who disseminate theories?
- Main concepts:
  - Adopters and non-adopters
  - Innovation
  - Adoption
  - Exposure

Example: diffusion of theories

Step 5: Identify measurable indicators for conceptual variables

- Measurable indicators for concepts:
  - Adopters and non-adopters:
    - Social scientists after 1944 until 1980s
  - Innovation:
    - John von Neumann and O. Morgenstern's *Theory of Games*
  - **Exposure:**
    - Institutional affiliation in key institutions of the Cold War between 1944 and 1980s;
    - Acknowledgements in published and working papers between 1944 and 1980s.
  - **Adoption:**
    - Citations of von Neumann and Morgenstern's *Theory of Games*.

Example: diffusion of theories

Step 5: Identify measurable indicators for conceptual variables

- Relational data, e.g. for scientific community:
  - Nodes, e.g.
    - individual scientists;
    - journals;
    - publications, working papers;
    - conferences, other science-related events;
    - research institutions.
    - ...
  - Edges, e.g.
    - co-authorship;
    - co-presence at conferences;
    - co-citations;
    - Overlap in references cited;
    - Overlap in classification codes (e.g. JEL classification system in economics, PsycINFO CCCS in psychology).
    - ...

Example: diffusion of theories

Step 6: Collect data

- No way around collecting what you can get!
  - Go into archives (e.g. letters, reports, meeting minutes, member lists, etc.);
  - Get official data about members from institutions.

- HERE: Identification of contributions that have in time proven 'seminal'
  - Co-citation network
  - All journal articles citing the *Theory of Games and Economic Behavior* (ToG) in Scopus 1981-2010;
  - All articles of 1944-1970 that are cited by articles in this group.

Example: diffusion of theories

Step 6: Collect data (*work in progress*)

Database (after preprocessing of data):
- 3,088 published articles that cite *ToG* in 1981-2010;
- 127,269 citations in total (41.2 citations on average per text);
- of which 12,050 citations of 3,161 texts published in 1944-1969 (3.8 citations on average; including 3088 citations of ToG).

- After subtracting *ToG*:
  - 8,962 citations of 3,160 unique texts published in 1944-1969 (2.8 citations on average).

- Source: Scopus

Example: diffusion of theories

- Most frequently cited works in this group of articles (1981-2010) citing *the ToG*.

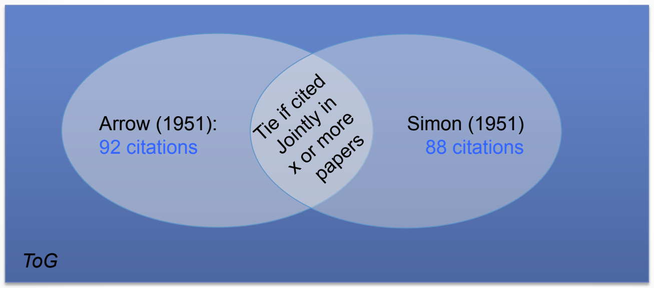
Citation	Times cited
Neumann, J. v., Morgenstern, O. (1944) Theory of Games and Economic Behavior. Princeton, NJ: University Press	3088
Savage, L.J., (1954) The Foundations of Statistics, John Wiley & Sons, New York	337
Nash, J.F., Non-cooperative games (1951) The Annals Mathematics, 54 (2), pp. 286-295	190
Luce, R.D., Raifa, H., (1957) Games and Decision, New York: McGraw-Hill	173
Allais, M., Le comportement de l'homme rationnel devant le risque, Critiques des postulats et axiomes de l'école américaine (1953) Econometrica, 21, pp. 503-546.	167
Ellsberg, D., Risk, Ambiguity, and the Savage Axiom (1961) Quarterly Journal of Economics, 75, pp. 643-649	161
Nash, J.F., Equilibrium points in n-player games (1950) Proceedings of the National Academy of Sciences, 36 (1), pp. 48-49	146
Pratt, J., Risk Aversion in the Large and Small (1964) Econometrica, 32, pp. 122-136	137
Nash, J.F., The Bargaining Problem (1950) Econometrica, 18, pp. 155-162. , April	125
Markowitz, H.M., Portfolio selection (1952) J. Finance, 12, pp. 77-91	112
Arrow, K.J., (1951) Social Choice and Individual Values, 1st ed. New Haven: Yale University Press.	92
Simon, H.A., A behavioral model of rational choice (1955) Quarterly Journal of Economics, 69 (1), pp. 99-118	88
Shapley, L., (1953) A Value for N-person Games, p. 28. , Princeton University Press	80
Friedman, M., Savage, L.J., Utility Analysis of Choices Involving Risk (1948) J. Political Economy, 56, pp. 279-304	73
Anscombe, F., Aumann, R., A definition of subjective probability (1963) Ann. Math. Statist., 34, pp. 199-205	71
Raiffa, H., (1968) Decision Analysis: Introductory Lectures on Choices under Uncertainty, Oxford, England: Addison Wesley	68
Schelling, T.C., (1960) The Strategy of Conflict, Cambridge, MA: Harvard University Press	64
Markowitz, H.M., (1959) Portfolio Selection: Efficient Diversification of Investments, Wiley, New York, NY	54
Harsanyi, J.C., Games with incomplete information played by 'Bayesian' players, Parts I, II and III (1967) Management Science, 14, pp. 159-182	52
Luce, R.D., (1959) Individual Choice Behavior: A Theoretical Analysis, New York: Wiley	46

Basic model

- Think of a theory as a set of (more or less) interdependent ideas;
- Scientific publications encapsulate these ideas;
- Scientific practice involves the citation of prior ideas that influence the idea developed in the present paper;
- If two publications are cited in the same publication, they have both had an influence on the citing publication.

Example: diffusion of theories

- Frequent co-citations suggest topical overlap:



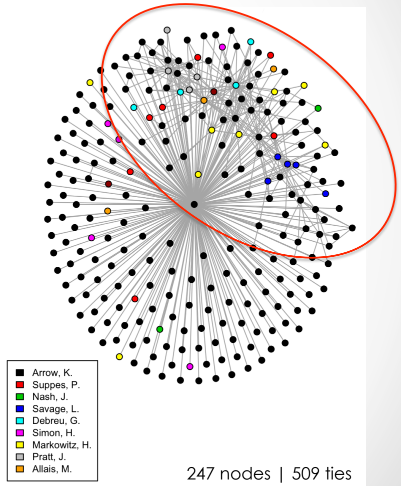
- If overlap is above a set threshold x (here: 5 or more co-citations) then two publications (here: Arrow 1951 and Simon 1951) were contributing to similar topics.

Example: diffusion of theories

Step 7: Construct the network

Including *ToG:*
- Co-citation network;
- *ToG* at center;
- A tie between nodes A and B is established if A and B are both cited together in 5 or more papers;

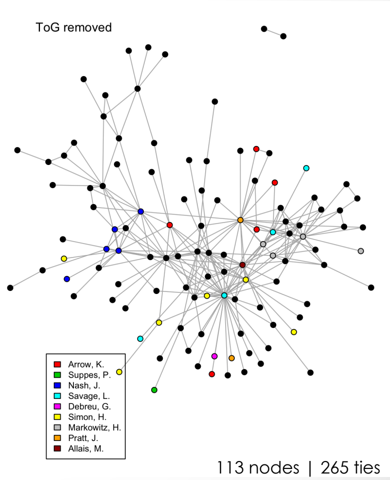
- Reminder: 3088 papers; a tie is established if there are 5 or more co-citations.



Example: diffusion of theories

Excluding *ToG*:

- after removing the isolates, i.e. nodes that are not linked to any other node, the network on the right obtains
- A link exists between any two papers that are cited together in five or more publications in the dataset.

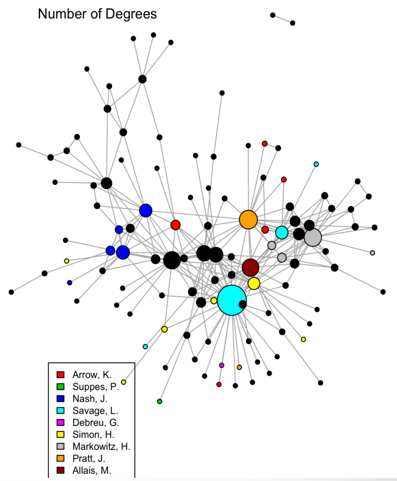


Example: diffusion of theories

Introduce measures of relative importance of contributions for theory:

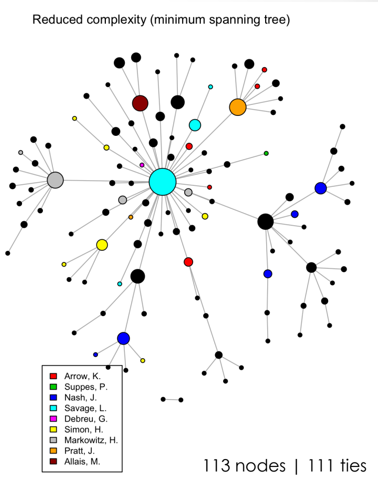
- Degree centrality: the number of edges that link to a node.

$$C_D(v) = \deg(v)$$



Example: diffusion of theories

- Many ways of reducing the complexity of the network to improve interpretability;
- *E.g. Minimum spanning tree* eliminates redundant paths between nodes such that the number of edges in the reduced graph is minimized;
- Note: contributions by actors known to have had an important role in the dissemination of the *ToG* are spread across the reduced network: they are similarly distant from the center of the network and they tend to bridge to further publications.

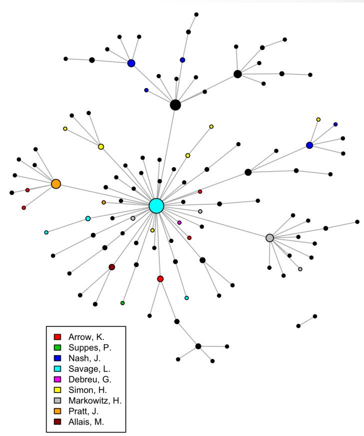


Basic centrality measures (1)

- Betweenness centrality:
- Quantifies how often a node acts as a bridge along the shortest path between two other nodes.

$$G := (V, E)$$

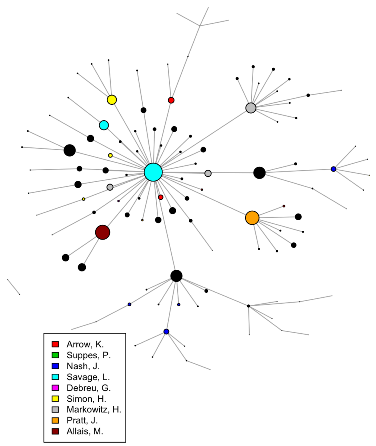
$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$



(mean-normalized betweenness centrality)

Basic centrality measures (2)

- Eigenvector centrality (Bonacich 1972)
  - Problem: high degree positions may be connected to many low degree others, while some low degree positions are connected to a few high degree others
  - Weighting centrality scores by the relative centrality of connected ties

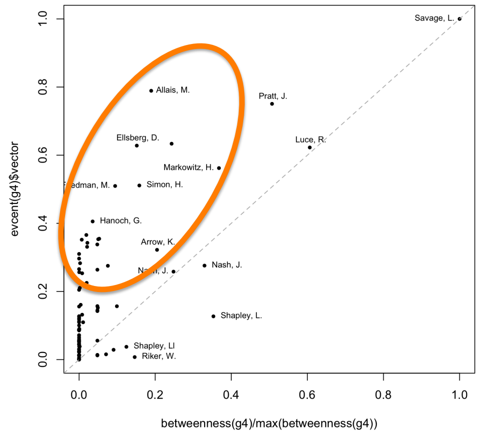


Example: diffusion of theories

Step 7: Interpretation of data

- …high eigenvector centrality indicates being cited with important others
- …low betweenness centrality implies being on the periphery of the network

- (Tentative) claim: contributions with a high eigenvector centrality but low betweenness centrality transport the theory into novel domains of application



Challenges

1. Problem of boundary specification;
2. Characterizing entities 'inside' the boundaries;
3. Conceptualizing causality:
   - How are network dynamics to be accounted for?
   - How are shifts in content and direction of relations to be explained?

Implications for social science

- Macro-level:
  - Replaces 'Society' by an analytically tractable construct ('relational setting');
  - Prescribes caution in assuming primacy of 'nations', 'countries', etc. as units of analysis.
- Meso-level:
  - Local pattern regularities in interactions:
    - locating regularities across transactional processes;
    - specifying recurrent mechanisms, patterns, and sequences in meso-level 'occasions'.
    - Patterns of invisible relations among actors, i.e. relations that are only visible by their absence, equally matter.
- Micro-level:
  - Reconceptualize individual identities and interests;
  - No pre-established interests, desires, beliefs;
  - Formation of 'entities' through transactional processes of social recognition;
  - Intrapsychic processes → 'relational individualism' (≠ drive theory).

Theoretical Implications

The relational view:

- Object of analysis is relation between entities (i.e. transaction);
- Perspective rejects idea that scientist/philosopher of science can posit discrete, pre-given units as starting point for analysis;
- Entities are inseparable from transactional context within which they are embedded;
- Relations are dynamic;
- Examples: buyer and seller are defined by way of interacting in a market setting.

"The terms or units involved in transaction derive their meaning, significance, and identity from the changing functional roles they play within that transaction" (Emirbayer 1997, p. 287).

Some further reading

- Erickson, P. et al (2013): How Reason Almost Lost its Mind, Chicago University Press.
- Knoke, David/Yang, Song (2008): Social Network Analysis, 2nd ed., Sage Publications.
- Newman, M.E.J. (2010): Networks: An Introduction, Oxford University Press.
- Rogers, Everett M. (2005): Diffusion of Innovations, Free Press.
- Wasserman/Faust (1995): Social Network Analysis – Methods and Applications. Cambridge University Press.

But there is much more …