Social and Knowledge Networks

Tina Wakolbinger

11-03-03

Agenda

- Social networks
  - History of social network theory
  - Basics of social network theory
  - Applications
  - Dynamic social networks
- Knowledge networks
  - Definitions
  - Schools of thought
  - Economics of knowledge based goods

Social Networks

History

- Spinoza, 17th century, first model
- Clausewitz, 1832, structuralist vision of warfare
- J.L. Moreno, 1937, introduced sociometry
- A. Bavelas founded the group networks laboratory at MIT in 1948, specified centrality
- R.D. Luce, A.D. Berry, 1949, definition of a clique
- C.L. Strauss, 1947, definition of kinship
- A. Rapaport, 1949, developed a probability based model of information flow

History 2

- 50s and 60s: distinct research by individual researchers
- 70s: field of social network analysis emerged
  - New features in graph theory – more general structural models
  - Better computer power – analysis of complex relational data sets

Social Network Methods

- Introduction to Social Network Methods
  - R.A. Hannemann, Department of Sociology, University of California, Riverside, 2001
  - Based on work by
    - L.C. Freeman, S. Borgatti, and M.G. Everett
    - S. Wasserman and K. Faust
    - P. Bonacich

Software

- UCINET 6 for Windows
  - S. Borgatti, M.G. Everett, and L.C. Freeman, 2002
  - All Data are described as matrices
- Krackplot 3.3
  - D. Krackhardt, J. Blythe, and C. McGrath, 2002
  - Graph layout software for social network analysis

Definition of Social Networks

- “A social network is a set of actors that may have relationships with one another. Networks can have few or many actors (nodes), and one or more kinds of relations (edges) between pairs of actors.” (Hannemann, 2001)

- Social network theory uses matrices and sociogramms

Matrices

- Square array of measurements
- Rows and columns are cases, subjects, or observations (nodes)
- Cells represent relationships (edges)
- Example: Who reports liking whom?

	Ann	Rob	Sue	Nick
Ann	---	1	0	0
Rob	1	---	1	0
Sue	1	1	---	1
Nick	0	0	1	---

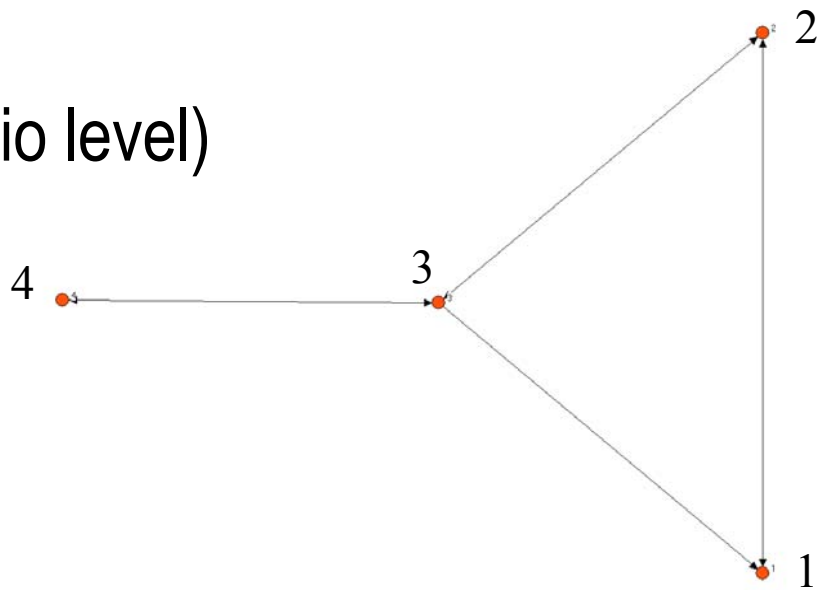
How are actors embedded in the network?
How is the overall density?

Graphs - Sociogramms

- Labeled circle for each actor in the population
- Line segments between pairs of actors represent ties between them
- A graph may represent a single type of relations or more than one kind
- Each tie can be directed or represent co-occurrence
- Arrows represent directed ties

Graphs – Sociogramms 2

- Strength of ties:
  - Nominal (presence or absence)
  - Signed (negative, positive, no tie)
  - Ordinal
  - Valued (interval or ratio level)
- Example



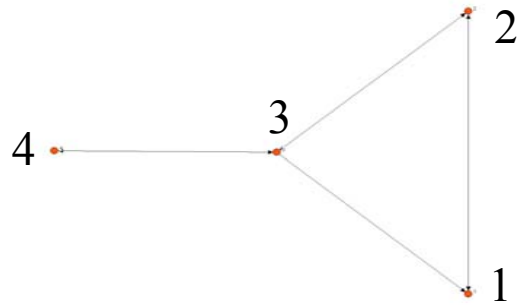
Connections

- Size
  - Number of nodes
- Density
  - Proportion of all ties that are present to the amount that could be present
- Out-degree
  - Sum of connections from an actor to others
  - Measures an actor's influence
- In-degree
  - Sum of each column in the adjacency matrix
  - High amounts: actors may be prestigious, powerful but also suffer from information overflow

	Ann	Rob	Sue	Nick
Ann	---	1	0	0
Rob	1	---	1	0
Sue	1	1	---	1
Nick	0	0	1	---

Connections 2

- Reachability
  - An actor is reachable by another if there exists any set of connections by which we can trace from the source to the target actor
- Reciprocity
  - Balance theory: If A is tied to B then B should be tied to A
- Transitivity
  - If A is tied to B, and B is tied to C, then A should be tied to C
- Neighborhood size
  - Number of other actors to whom the actor is adjacent



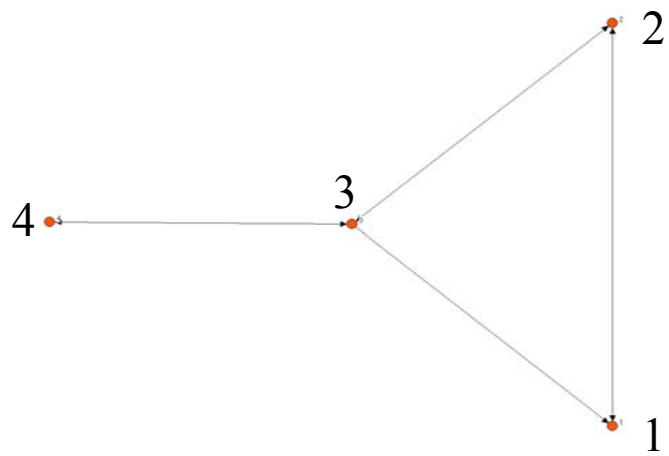
Distance

- Walk
  - A sequence of actors and relations that begins and ends with actors; a *closed walk* is one where the beginning and end point of the walk are the same actor
- 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

Example: Geodesic Distance

Geodesic Distances: The number of relations in the shortest possible walk from one actor to another

	1	2	3	4
1	0	1	2	3
2	1	0	1	2
3	1	1	0	1
4	2	2	1	0

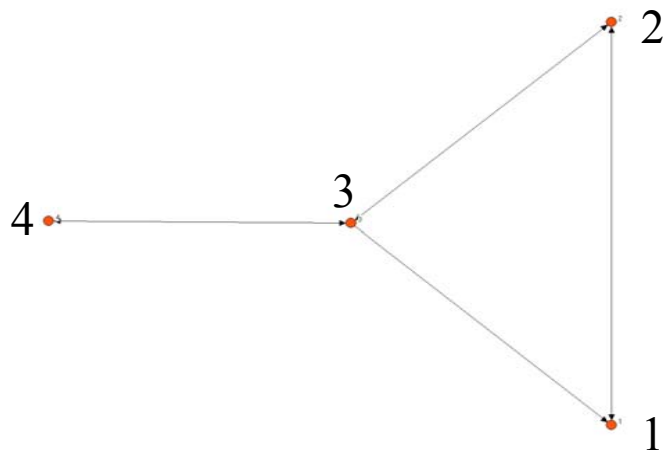


Borgatti, S.P., Everett, M.G. and Freeman, L.C. 2002.
Ucinet for Windows: Software for Social Network Analysis. Harvard: Analytic Technologies.

Example: Maximum Flow

Maximum Flow: How many different actors in the neighborhood of a source lead to pathways to a target

	1	2	3	4
1	0	1	1	1
2	2	0	1	1
3	2	2	0	1
4	1	1	1	0



Borgatti, S.P., Everett, M.G. and Freeman, L.C. 2002.
Ucinet for Windows: Software for Social Network Analysis. Harvard: Analytic Technologies.

Distance 2

- The Hubbell and Katz approaches
  - Count the total connections between actors. Each connection, however, is given a weight, according to it's length.
- The Taylor measure
  - The column marginals for each actor are subtracted from the row marginals, and the result is then normed
    - Shows senders and receivers of information

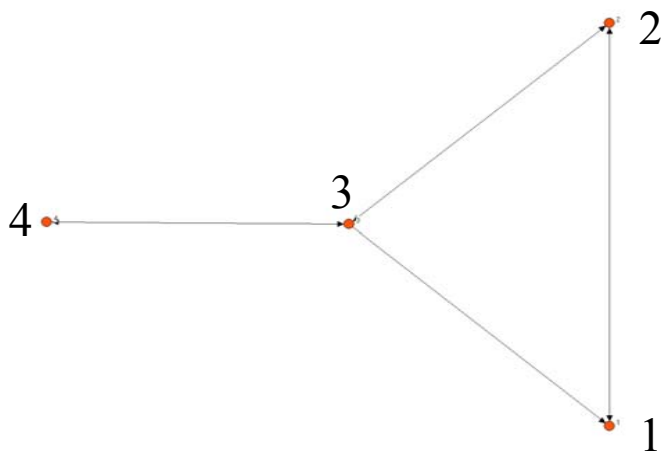
Measures of Power

- Degree
  - Sum of connections from or to an actor
- Closeness centrality
  - Distance of one actor to all others in the network
  - Sum of geodesic distance is farness
- Betweenness centrality
  - Number how frequently an actor is between other actors' geodesic paths
- Flow centrality
  - Measure how often the actor is in all the flows between all other pairs of actors

Example: Betweenness Centrality

Betweenness Centrality: Number how frequently an actor is between other actors' geodesic paths

3	4.000
2	2.000
1	0.000
4	0.000



Borgatti, S.P., Everett, M.G. and Freeman, L.C. 2002.
Ucinet for Windows: Software for Social Network Analysis. Harvard: Analytic Technologies.

Measures of Power

- Bonacich power index
  - The more connections the actors in the neighborhood have the more central and the less powerful an actor is
  - Iterative estimation

Cliques and Sub-Groups

- A Clique
  - Is a sub-set of actors who are more closely tied to each other than they are to actors who are not part of the group
- Bottom-up approaches
  - N-cliques
    - Member of a clique if they are connected to every other member of the group at distance n
  - K-plex
    - Actors may be member of a clique if they have ties to all but k other members

Cliques and Sub-groups 2

- Top-down approaches
  - Components
    - Parts that are connected within, but disconnected between sub-graphs
  - Cutpoints
    - Nodes that if they are removed the structure becomes divided into un-connected systems
  - Blocks
    - Divisions into which cutpoints divide a graph
  - Lambda sets
    - Set of actors who, if disconnected would most greatly disrupt the flow among all the actors.

Network Positions and Social Roles

- Social roles are defined by regularities in the patterns of relations among actors.
- Two actors have the same position to the extent that their pattern of relationship with other actors are the same.

Types of Equivalence

- Structural equivalence
  - Two nodes have the same relationship to all other nodes
- Automorphic equivalence
  - Actors are automorphically equivalent if we can permute the graph in such a way that exchanging the two actors has not effect on the distances among all actors in the graph
- Regular equivalence
  - Two nodes have the same profile of ties with members of other sets of actors that are also regularly equivalent

Forming Sub-groups

- Hierarchical clustering based on
  - Pearson correlation coefficients
  - Euclidean distance
    - Less weight to extreme cases
  - Percent of exact matches
    - For binary data
  - Jaccard coefficients
    - Percent of positive matches
    - For sparse connections

Forming Sub-groups 2

- CONCOR
  - Correlating each pair of actors
  - Each row is extracted and correlated with each other row,
  - Divides the data into two sets
  - Then repeats the process within each set
- Tabu search
  - Searches for the sets of actors who, if placed into a block, produce the smallest sum of within-block variances in the tie profiles

Example: Forming Sub-groups

Initial Correlation Matrix

	1	2	3	4
1	1.00	0.77	0.33	0.15
2	0.77	1.00	0.26	0.58
3	0.33	0.26	1.00	0.45
4	0.15	0.58	0.45	1.00

Relation 1 Blocked Matrix

	1	2	3	4
1	1	1		
2	1		1	
3	1	1	1	1
4			1	

PARTITION DIAGRAM

Level	1	2	3	4
1	XXX	XXX		

Density Matrix

	1	2
1	1.000	0.250
2	0.500	1.000

Borgatti, S.P., Everett, M.G. and Freeman, L.C. 2002.
Ucinet for Windows: Software for Social Network Analysis. Harvard: Analytic Technologies.

Applications

- Visualizing networks – showing nodes and meta nodes
- Studying differences of cultures and how they can be changed
- Analyzing coded texts
- Group problem solving
- Computer mediated communication
- Intra- and interorganizational studies
- Spread of illness, especially HIV

Dynamic Networks

- Limitations to traditional social network analysis
  - Focused on small bounded networks with 2-3 types of links, among one type of nodes, at one point of time, with close to perfect information
- Dynamic networks
  - Meta matrix
  - Treating ties as probabilistic
  - Combining social networks with cognitive science and multi-agent systems
    - Interaction is influenced by relative similarity, relative expertise and co-worker, agents can store information and learn
  - Networks and agents co-evolve

Change Processes in DNA

Table 2. Basic Change Processes for Nodes in the Meta-Matrix

People	Knowledge/Resources	Events/Tasks	Organizations
Birth	Innovation	Goal Change	Organizational birth
Death	Discovery	Re-engineering	Organizational death
Promotion	Forgetting	Development of new technology	Mergers
Mobility	Consumption		Acquisitions
Recruitment		Stop usage of technology	Legislation of new entity
Incarceration			
Isolation			

Carley in Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers (2003)

Change Processes in DNA

Table 3. Change Processes for Relations in the Meta-Matrix

	People	Knowledge/ Resources	Events/ Tasks	Organizations
People	Motivation to Interact Change in access	Learning Acquisition	Re-assignment	Mobility Recruitment
Knowledge/ Resources		Discovery Analogical reasoning	Innovation	IP development
Events/Tasks				Re-engineering Out-sourcing
Organizations				Alliances Coalitions

Carley in Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers (2003)

Applications of DNA

- The possible effects of biological attacks on cities – BioWar (Carley et al 2002)
- Evaluation of CIO response strategies to denial of service attacks (Chen 2002)
- Evaluation of information security within organizations (ThreatFinder Project Carley 2001)
- Evaluation of how to build stable adaptive networks with high performance and how to destabilize networks (DyNet)
  - Basic cognitive forces are learning, forgetting, goal setting and motivation for interaction
  - Basic social forces are recruitment, isolation, initiation of rumors and training

Knowledge Networks

Definitions of Knowledge

- ‘Justified true belief’ (Plato)
- ‘The whole set of insights, experiences and procedures which are considered correct and true and which, therefore, guide the thoughts, behaviors and communication of people’ (Van der Spek, 1997)
- ‘A fluid mix of framed experience, values, contextual information, and expert insight’ (Davenport and Prusak, 1998)

Types of Knowledge

- Explicit knowledge
  - Written down, recorded in some way
- Implicit knowledge
  - Vision, values, culture
- Tacit knowledge
  - Learning by doing
- Organizational knowledge
- Personal knowledge

Knowledge Production

- Production of new knowledge is ‘a series of transformations, by which standard resources, which are available in open markets [or contained within the organizations], are used and combined within the organizational context in order to produce [competences and] capabilities’ (Ciborra and Andreu, 2001)

Schools of Thought

- The technocratic school
  - focuses on information management or management technologies which are thought to assist the employees in improving their business performance.
- The economic school
  - regards knowledge as an intellectual capital/asset to be exploited.
- The behavioral school
  - endeavors to create a business culture which stimulates knowledge production, sharing and (re)use.

Measures of Knowledge Quality

- Correctness
- Uniqueness
- Timeliness

Economics of Knowledge Based Goods (S.J. Liebowitz)

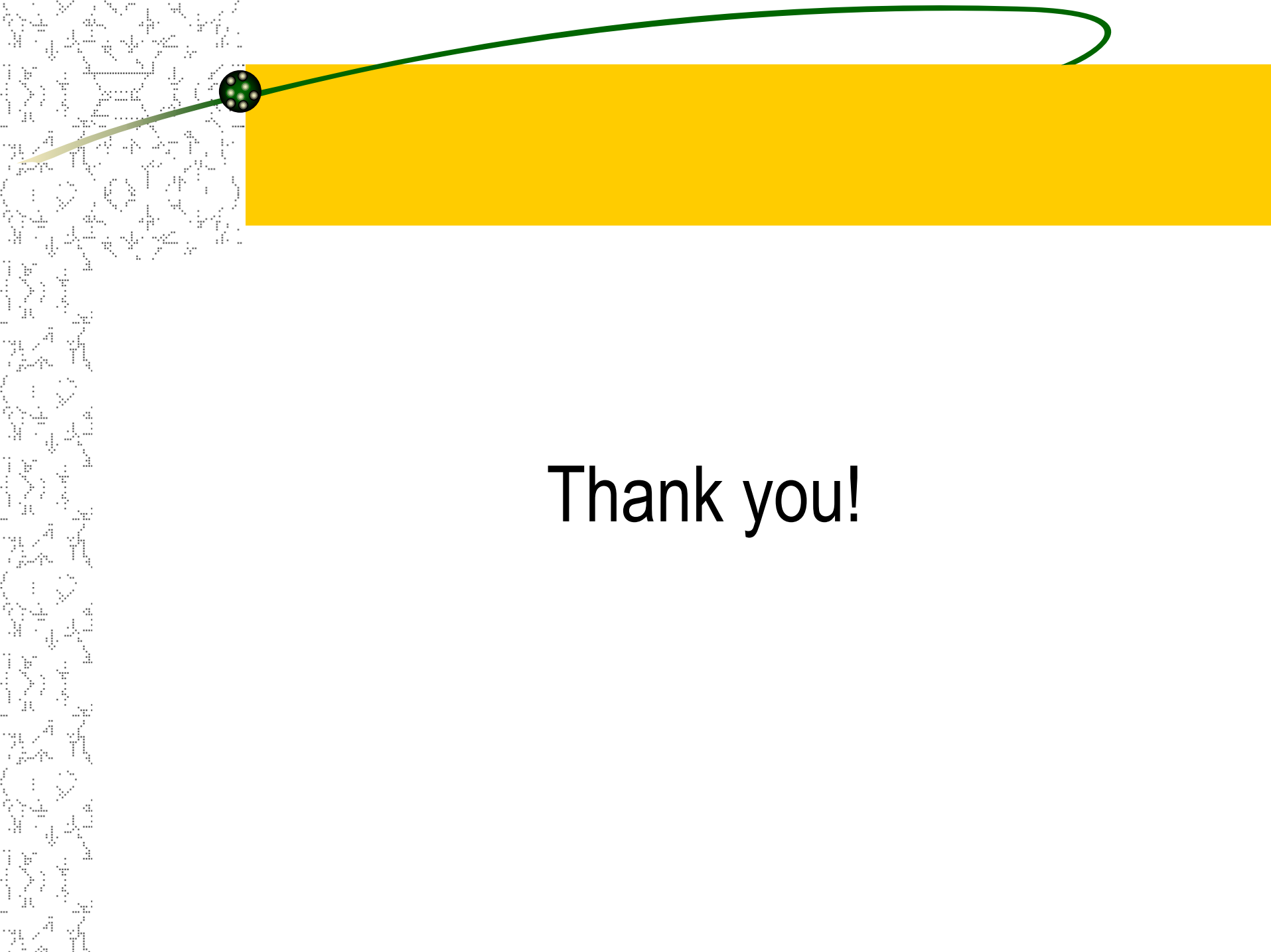
- Public goods (Samuelson)
  - Nonrival consumption: goods such that one person's consumption doesn't reduce anyone else's possible consumption
  - Non-excludability of users
  - Demand curve is the vertical sum of the demands for all consumers at the given quantity
  - The sum of the prices that all individuals would be willing to pay for a given quantity is the price that the market is willing to pay

Example of a Knowledge Based Good

- Book
  - Public good: book title
  - Private good: physical copies
  - Revenue
    - Total revenue - printing costs – costs of writing
  - Problem of copies
    - Site licensing
    - Higher journal prices to libraries

Conclusion

- Special characteristics of knowledge based products
  - Nonrival consumption
    - Demand is vertical sum
    - Problem of copies
  - Quality depends on organizational and individual knowledge
  - Standardization of production is more difficult or even impossible


Thank you!