6.207/14.15: Networks
Lecture 1: Introduction

Daron Acemoglu and Asu Ozdaglar
MIT

September 9, 2009

Outline

- What are networks?
- Examples.
- Small worlds.
- Economic and social networks.
- "Network effects".
- Networks as graphs.
- Strong triadic closure.
- Power in a network.
- Decisions and games in networks.
- Implications of strategic behavior.
- Rest of the course.

Reading:

- EK, Chapter, 1 (also skim Chapters 3-5).
- Jackson, Chapter 1.

Introduction

- What are networks? Why study networks? Which networks and which commonalities? Which tools?
- Networks are a representation of **interaction** structure among units.
  - In the case of social and economic networks, these units (nodes) are individuals or firms.
- At some broad level, the study of networks can encompass the study of all kinds of interactions.
  - Information transmission.
  - Web links.
  - Information exchange.
  - Trade.
  - Credit and financial flows.
  - Friendship.
  - Trust.
  - Spread of epidemics.
  - Diffusion of ideas and innovation.

Visual Examples—1

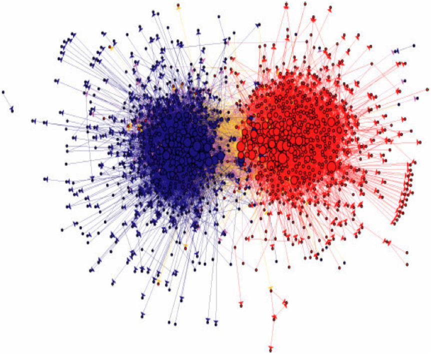


Figure: The network structure of political blogs prior to the 2004 U.S. Presidential election reveals two natural and well-separated clusters (Adamic and Glance, 2005)

Visual Examples—2

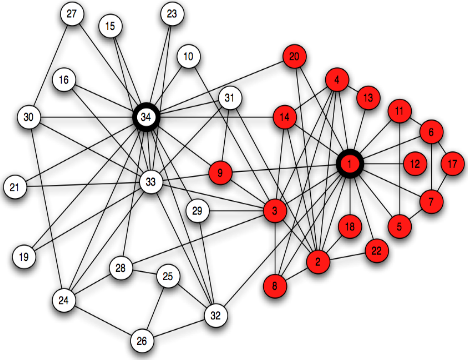


Figure: The social network of friendships within a 34-person karate club provides clues to the fault lines that eventually split the club apart (Zachary, 1977)

Visual Examples—3

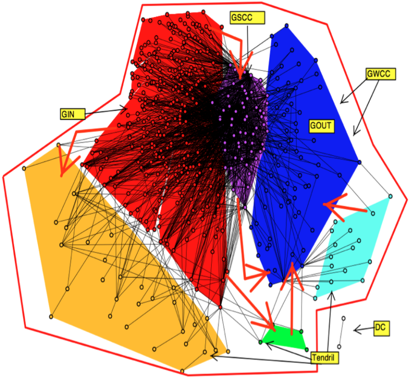


Figure: The network of loans among financial institutions can be used to analyze the roles that different participants play in the financial system, and how the interactions among these roles affect the health of individual participants and the system as a whole. (Bech and Atalay 2008)

Visual Examples—4

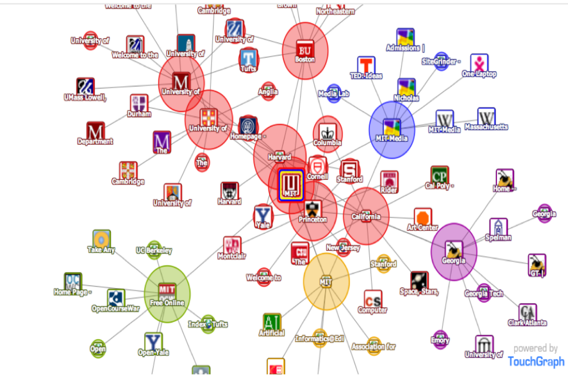


Figure: The web link structure centered at http://web.mit.edu (touchgraph)

Visual Examples—5

Figure: The spread of an epidemic disease (such as the tuberculosis outbreak shown here) is another form of cascading behavior in a network. The similarities and contrasts between biological and social contagion lead to interesting research questions. (Andre et al. 2007)

Visual Examples—6

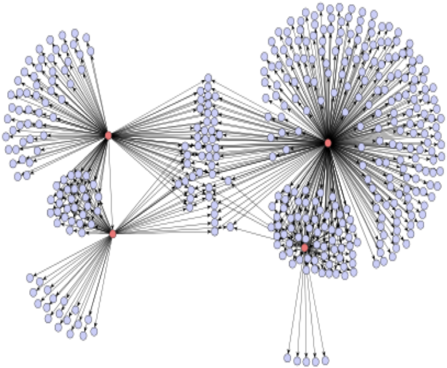


Figure: When people are influenced by the behaviors of their neighbors in the network, the adoption of a new product or innovation can cascade through the network structure. Here, e-mail recommendations for a Japanese graphic novel spread in a kind of informational or social contagion. (Leskovec et al. 2007)

Visual Examples—7

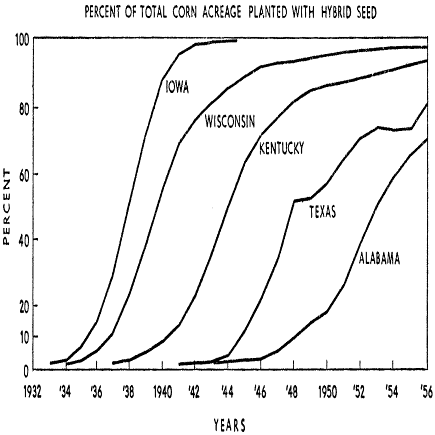


Figure: Percentage of total corn acreage planted with hybrid seed. (USDA Agricultural Statistics)

Do We Live in a Small World?

- Early 20th century Hungarian poet and writer Frigyes Karinthy first came up with the idea that we live in "small world". He suggested, in a play, that any two people among the one and a half billion inhabitants of the earth then were linked through at most five acquaintances.
- The sociologist Stanley Milgram made this famous in his study "The Small World Problem" (1967)—though this study is now largely discredited.
- He asked certain residents of Wichita and Omaha to contact and send a folder to a target person by sending it to an acquaintance, who would then do likewise etc., until the target person was reached. This would allow Milgram to measure how many "intermediate nodes" would be necessary to link the original sender and the target.
- 42 of the 160 letters supposedly made it to their target, with a median number of intermediates equal to 5.5.

Do We Live in a Small World? (continued)

- Hence was born the idea of six degrees of separation.
- Can you think why Milgram's procedure could give misleading results? Or why we may not wish to take these results on faith?
- There are similar studies for other types of networks.
- For example, Albert, Jeong, and Barabasi (1999) "Diameter of the World Wide Web" estimated that in 1998 it took on average 11 clicks to go from one random website to another (at the time there were 800 million websites).
- What do these kind of "small world" results imply?

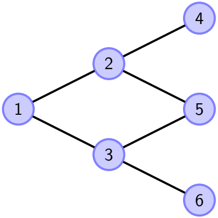
Interpreting Small Worlds

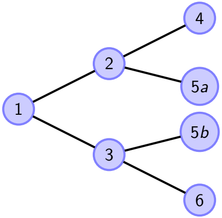
- Suppose that each node has λ neighbors (e.g., each website has links to λ other websites).
- Each of my λ neighbors will then have λ neighbors themselves.
- Suppose (unrealistically) that my neighbors don't have any neighbors in common (i.e., the λ websites that are linked to my website are not linked among themselves). Then in two steps, I can reach λ^2 other nodes.
- Repeating the same reasoning (and maintaining the same unrealistic assumption), in d steps I can reach λ^d other nodes.
- Now imagine that this network has $n = \lambda^d$ nodes.
- This implies that the "degrees of separation" (average distance) is

$$d = \frac{\ln n}{\ln \lambda}.$$

Interpreting Small Worlds (continued)

- But our unrealistic assumption rules out the reasonable triadic relations and clustering phenomena, which are common both in social networks, web links, and other networks.





- Interestingly, however, in Poisson (Erdos-Renyi) random graphs, we will see that average distance can be approximated for large n by $d = \ln n / \ln \lambda$ (where λ is the expected degree of a node).
- This is because triadic relations shown in the figure are relatively rare in such graphs.

Interpreting Small Worlds (continued)

- This last result in fact can be interpreted as stating that Poisson (Erdos-Renyi) random graphs, though mathematically convenient, will not be good approximations to social networks.
- This can be seen from the above numbers as well.
- The Karinthy conjecture, under the Poisson assumption, would require that each person should have had approximately 68 "independent" friends. ($\exp[\ln(1,500,000,000)/5] \simeq 68.5$).
- The Milgram conjecture, of six degrees of separation in the 1960s, would require that each person should have had approximately 41 friends.
- Instead, most people would be connected to others in remote parts through "special links" (or "connectors"), such as their political representatives, village head, or cousin in a different city etc.
- Models of small world networks try to capture this pattern (albeit not always perfectly).

Social and Economic Networks

- Most "networked" interactions involve a human element, hence much of network analysis must have some focus on social and economic networks (even when the main interest may be on understanding communication networks).
  - E.g., social network structures, such as Facebook, superimposed over the Internet.
- In this course, social and economic networks will be our main focus.
- An important feature of social and economic networks is that they are not only characterized by a pattern of linkages, but also by the interactions that take place over the network structure.
  - Will you lend money to your friend? Will you follow their advice? Will you imitate their behavior? Will you trade with other firms that you are potentially "connected to"?
- Most of these decisions are strategic, hence the use of game theory.

A Central Question

- What are the commonalities in different (social, economic and other) networks?
- Diffusion of new technologies and spread of epidemics have certain common features when one looks at their dynamics.
  - Does this mean that they obey the same logic?
  - Should we have a single theory to explain both?
  - Should we use the same mathematical tools to analyze both?

Importance of Networks in Economics and Sociology?

- Sociology largely about group interactions, thus network structure naturally important.
  - Notions such as social capital, power, or leadership may be best understood by studying the network of interactions within groups.
- Sociology largely descriptive and nonmathematical. Can the study of networks bring more analytic power to sociology?
- For example, what is "social power" related to? What kind of relationships and linkages does a leader need to have in a community?
- Or about dynamics of groups: is the karate club depicted above likely to splinter into two groups?

Importance of Networks in Economics and Sociology? (continued)

- Economics about "allocation of scarce resources"— broadly construed: "trades," cooperation vs. competition, information exchange and aggregation, technology adoption, etc.
- Much of this allocation takes place in networked situations. But much of economics studies either one of two extremes: (1) markets, where all interactions are anonymous (implicitly anybody can trade with anybody else); (2) games among few players— with the identities of the players predetermined.
  - Example: competitive equilibrium at the one end, and bargaining and auctions at the other.
- Can we develop new insights by systematically analyzing (and economically representing) the network of relations underlying "trades"?

Examples of "Network Effects"

- How do people find jobs?
  - Myers and Shultz (1951) *The Dynamics of a Labor Market* and Rees and Shultz (1970) *Workers in an Urban Labor Market* documented that most workers find (have found) their jobs through "a social contact".
  - Granovetter (1973) "The Strength of Weak Ties": most people find jobs through acquaintances *not* close friends.
  - Is this a puzzle?
  - Yes and no. No because people have many more acquaintances than friends, but also because of strong triadic closure; if 1 and 2 are close friends, and 2 and 3 are close friends, then 1 and 3 are very likely to know each other. Therefore, you are more likely to get referrals to a manager whom you don't know through an acquaintance than a close friend → importance of weak ties.
- Weak ties may also be very important in understanding "social capital".

Examples of "Network Effects" (continued)

- How do people start and run their businesses?
  - In many developing economies (but also even in societies with very strong institutions), networks of "acquaintances and contacts" shape business behavior.
  - Munshi (2009) "Strength in Numbers: A Network-Based Solution to Occupational Traps": Indian diamond industry, which makes up about 14% of total merchandise exports, is dominated by a few small subcasts, *the Marwaris, the Palanpuris, the Kathiawaris*—in the same way that Antwerp and New York diamond trade used to be dominated by ultra-Orthodox Jews.
  - Initially, *the Marwaris* and *the Palanpuris* dominated Indian diamond trade. But in the 1980s, the lower agricultural subcast, *the Kathiawaris*, started dominating much of Indian exports.

Examples of "Network Effects" (continued)

- What explains the rise of *the Kathiawaris*?
  - India does not produce rough diamonds, so mostly brought from Antwerp. But legal contracts difficult to enforce, particularly for small traders, thus trust relations especially important.
  - *The Marwaris* and *the Palanpuris* institutionalized their relationship with Antwerp (often opening branches of their firms there). Moreover, over time, lower intermarriage rates for these groups. Network relationships seem to matter less.
  - *The Kathiawaris* initially a lower, agricultural subcast, some of them working as cutters for *the Marwaris* and *the Palanpuris.* Strong network ties, intermarriage rates etc. After the increase in the world supply of rough diamonds in the 1970s (following the opening the Australia's Argyle Mines), *the Kathiawaris* slowly dominate the business. Mutual support, referrals, long-term relationships based on networks.

Examples of "Network Effects" (continued)

- How do people learn about new products?
  - Example: the Japanese graphic novel.
  - "Cult following" for movies or records.
- How does a new technology spread?
  - More important examples: the diffusion of new technologies and agriculture. Famous example: hybrid corn in the United States in the early 20th century. Spreading with a clear special pattern. Word-of-mouth from the early adopters important.
  - Similar patterns seen in prescription of new medication by doctors in the Midwest in the 1960s.
- How do people form their political, social and religious opinions?
  - Imitate family, friends and neighbors? Wisdom of the crowds?
  - More sophisticated information aggregation by talking and observing friends and news sources?
  - Does the social network matter?

Another Pertinent Question

- Have the tremendous advances in information and communication technology changed the nature of social networks?
  - Recall that Frigyes Karinthy had suggested that the world had become small only recently at the beginning of the 20th century. Perhaps it has become small now?
- Do new communication mediums such as Facebook, MySpace, Twitter change what type of information we obtain and how we process information?
  - Most people use new mediums to communicate with a small group.
  - Recall the political blogs. Certainly, the web does not seem to automatically guarantee greater that each individual will obtain a greater diversity of opinions.
  - Perhaps greater access to information can increase "herding"— excessive copying of others behavior and information instead of "wisdom of crowds" phenomena.

Networks As Graphs

- We will typically (mathematically) represent networks with graphs, which formalize the patterns of links between different units, or nodes.
- Graphs can be directed or undirected, depending on what kinds of relationship they represent. For example, blog links are directed.
- They can also be weighted or unweighted, depending on whether links differ in terms of their importance, capacity, likelihood of materializing, etc.. For example, weak vs. strong ties in referrals.
- At the simplest level, a directed (unweighted) graph is

$$G = (N, E)$$

N = the set of nodes in the graph (e.g., in the blogs example, the nodes are the weblogs)
E = the set of edges, linking nodes in the graph.

Networks As Graphs (continued)

- We write $j \in N$ if j is a node in this network, and $(i, j) \in E$ if there is a link from i to j.
- If this is a directed graph, then this does not necessarily imply that $(j, i) \in E$.
- For undirected graphs, sometimes we use the notation $\{i, j\} \in E$ to denote an edge between i and j, but we will not do so in this lecture.
- We can also use the notation $g_{ij} = 1$ if $(i, j) \in E$ and $g_{ij} = 0$ otherwise (and use g as the matrix of g_{ij}'s to do matrix algebra to derive properties of networks).
- For a weighted graph, we could also use the notation $g_{ij} > 0$ if $(i, j) \in E$ and $g_{ij} = 0$ otherwise.
    - In this case, the magnitude of g_{ij} would correspond to the strength of the link.

The Strong Triadic Closure

- Recall job referral patterns.
- Let us represent a weighted (undirected) graph in an economical fashion as "augmented" undirected graph, $G = (N, E, E')$, where $E' \subset E$ represents "strong ties". Thus, $(i, j) \in E$ means that i and j are acquaintances, while $(i, j) \in E'$ means that i and j are close friends.
- The strong triadic closure property is the following:

$$\text{if } (i, j) \in E' \text{ and } (i, k) \in E', \text{ then } (j, k) \in E.$$

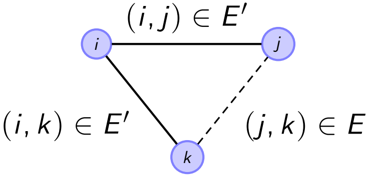


Figure: Triadic Closure

The Strong Triadic Closure (continued)

- Naturally, this property in this strong form is often violated, so one may wish to have a "probabilistic" version of this, where we would say that the conditional probability that $(j,k) \in E$ given $(i,j) \in E'$ and $(i,k) \in E'$ is greater than the unconditional probability that $(j,k) \in E$, i.e.,

$$\mathbb{P}\left((j,k) \in E \mid (i,j) \in E' \text{ and } (i,k) \in E'\right) > \mathbb{P}\left((j,k) \in E\right).$$

An alternative probabilistic version would be

$$\begin{aligned}&\mathbb{P}((j,k) \in E \text{ and } (i,j) \in E' \mid (i,k) \in E') \\ &\qquad > \mathbb{P}((j,k) \in E \text{ and } (i,j) \in E' \mid (i,k) \in E \setminus E')\end{aligned}$$

The Strong Triadic Closure (continued)

- Now suppose that i can get a job with manager k with a referral through j if j is close friends with k but i and k do not know each other.
- Assume the probabilistic version of the strong triadic closure.
- Then close friends are less useful for finding jobs than acquaintances.
- More formally, let $\{(i, j) \in R\}$ be the event that i obtained the job through a referral by j, and $\mathbb{P}((i, j) \in R)$ denote the probability of this event.

The Strong Triadic Closure (continued)

- Then

$$
\begin{aligned}
\mathbb{P}((i,j) \in R \quad | \quad (i,j) \in E' \ ) \\
&= \mathbb{P}\left((i,k) \notin E \text{ and } (j,k) \in E' \mid (i,j) \in E' \ \right) \\
&< \mathbb{P}\left((i,k) \notin E \text{ and } (j,k) \in E' \mid (i,j) \in E \setminus E' \ \right) \\
&= \mathbb{P}\left((i,j) \in R \mid (i,j) \in E \setminus E'\right),
\end{aligned}
$$

potentially explaining Granovetter's findings. In fact, with the non-probabilistic version of the property, $\mathbb{P}\left((i,j) \in R \mid (i,j) \in E' \ \right) = 0$!

Power in a Network

- The Medicis emerged as the most influential family in 15th century Florence. Cosimo de Medici ultimately formed the most politically powerful and economically prosperous family in Florence, dominating Mediterranean trade.
- The Medicis, to start with, were less powerful than many other important families, both in terms of political dominance of Florentine institutions and economic wealth.
- How did they achieve their prominence?
- It could just be luck (in social science, we have to be very careful to distinguish luck from a systematic pattern, and correlation from causation).
- An interesting explanation, eschewing luck, is offered by Padgett and Ansell (1993) "Robust Action and the Rise of the Medici"— they were the most powerful family because of their **situation in the social network** of Florence.

Power in a Network (continued)

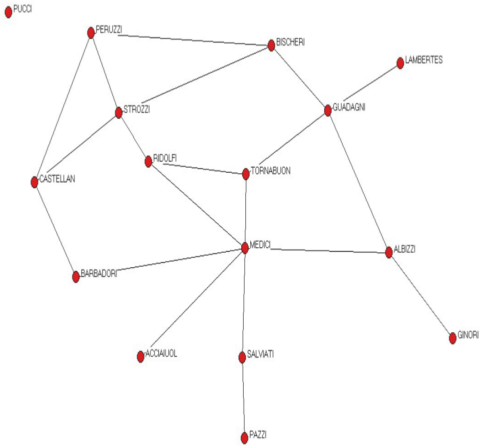


Figure: Political and friendship blockmodel structure (Padgett and Ansell 1993)

Power in a Network (continued)

- One measure of power that takes into account the "location" of the family with the network is the "betweenness" measure defined as follows.
- Let $P(i,j)$ be the number of shortest paths connecting family i to family j.
- Let $P_k(i,j)$ be the number of shortest paths connecting these two families that include family k.
- The measure of betweenness (for a network with n nodes) is then defined as

$$B_k \equiv \sum_{(i,j)\in E: i\neq j, k\notin\{i,j\}} \frac{P_k(i,j)/P(i,j)}{(n-1)(n-2)/2},$$

  with the convention that $P_k(i,j)/P(i,j) = 0$ if $P(i,j) = 0$.
- Intuitively, this measure gives, for each pair of families, the fraction of the shortest paths that go through family k.

Power in a Network (continued)

- It turns out that this measure over betweenness B_k is very high for the Medicis, 0.522.
- No other family has B_k greater than 0.255.
- So the Medicis may have played a central role in holding the network of influential families in Florence together and thus gained "power" via this channel.
- Is this a good measure of "social power"? Of political power? Is this a plausible explanation?

Decisions and Games on Networks

- Social networks are interesting because they represent **interactions** among different agents in a social situation.
- Thus decisions that define these interactions (trade, trust, friendship, imitation) are key.
- But this implies that interactions will be strategic, and we have to think of how social networks shape strategic interactions, and also how they are formed and evolved as a result of such strategic interactions.
- Let us next consider some examples where strategic interactions may significantly change the way we may wish to think of network relations.

Example I: Are More Links Always Better?

- Let us return to business networks. Clearly, links are "good" in this context, since they represent trust in trade relationships. But could less be more?
- Recall that Munshi's argument was that network connections helped *the Kathiawaris* pull ahead of the richer and more established *Marwaris* and *Palanpuris.*
- But why don't (didn't) *the Marwaris* and *the Palanpuris* exploit their well-established positions and greater links (especially in Antwerp) to form even stronger network ties?

Example I: Are More Links Always Better? (continued)

- Perhaps the answer is that more links are not always better.
- *The Marwari* and *the Palanpuri* businessmen were sufficiently more established, so they did not depend on their subcast links, so implicitly reneging on their long-term relationships within their cast would have carried relatively limited costs for them.
- But if so, then there would be little "trust" in the network of *the Marwaris* and *the Palanpuris*. (What does "trust" mean here?).
- In contrast, *the Kathiawaris* strongly depended on their network, so any reneging (or appearance of reneging) would lead to their exclusion from the business community supporting them forever—-and this support is very valuable to *the Kathiawaris*.
- Thus in this example, after a certain level, fewer links may be better—to make one more dependent on his network and thus more trustworthy.

Example II: "Acting White"—Are More Links Encouraged?

- In many minority groups, in the United States and in developing countries, those perceived as "acting white," that is, adopting norms and behavior patterns of majority groups, including high achievement in schooling, receive social sanctions.
  - For example, for the caste system in India playing this role, see Munshi and Rosenzweig (2006) "Traditional Institutions Meet the Modern World: Cast, Gender and Schooling Choice in a Globalizing Economy".
- Why would this be?
- One possible explanation may be that these kind of sanctions severe the outside links (reduce the outside options) of minority kids and make them more dependent on (and more dependable for) the minority network.
  - See Austen-Smith and Fryer (2006) "An Economic Analysis of Acting White" for a richer model with imperfect information and signaling.

Example III: The Role of Diversity—Is Greater Diversity Useful?

- "Wisdom of the crowds": combining the information of many, particularly of those with different perspectives and diverse experiences, would lead to better decisions.
- Francis Galton and Marquis de Condorcet: "average of a group is wiser than its members"
  - Galton: people in the market guessing the weight of an ox: "The average competitor was probably as well fitted for making a just estimate of the dressed weight of the ox, as an average voter is of judging the merits of most political issues on which he votes." And he found out that the average competitor could do very well.
  - Condorcet jury theorem: apply the law of large numbers to opinions that are independent draws from a random distribution with mean equal to the "truth".
- These perspectives suggest that large groups ("large networks") can reach better and more accurate decisions.

Example III: The Role of Diversity (continued)

- But “groupthink” (a version of herding) in large groups (networks).
- Also cooperation and coordination much harder in groups with greater diversity.
- Perhaps more importantly, group decision-making fraught with difficulties.
- Arrow’s Impossibility Theorem: it is impossible for a group to have a decision rule that is efficient and non-dictatorial (and that satisfies the independence of irrelevant alternatives).
- Why? Because of conflict of interest among group members.
- The challenge to group decision-making.
- Other challenges: difficulty of coordination, free-rider problems, and communication problems.

The Rest of the Course

- More on graph theory and social networks.
- Random graphs and their applications.
- Rich get richer phenomena on networks.
- Spread of epidemics.
- Game theory for the analysis of strategic interactions.
- Game theory on networks.
- Trust, cooperation and trade on networks.
- Diffusion of innovation and ideas on networks.
- Network effects.
- Learning and information diffusion on networks: wisdom of the crowds or wisdom of the few?
- Markets vs. networks.
- Decision-making in organizations, committees and societies.