# A Case-Based Approach to Network Formation

Economics Resarch Seminar Presentation, Temple University

Dimitrios Diamantaras, Temple U     Robert P. Gilles, Virginia Tech

25 April 2007

**Outline**

# Contents



# 1 Case-Based Decision Making

## 1.1 Basics of Case-Based Decision Making (Gilboa and Schmeidler 2001)

**Who?**

- Itzhak Gilboa and David Schmeidler developed case-based decision making in the 1990s.
- They give a nice summary in their 2001 book *A Theory of Case-Based Decisions*.

**Why?**

- We all know what a case is. Whole courses in business schools focus on cases.
- But in standard decision theory, we insist that somehow all the decision-maker's previous knowledge (presumably gained from similar decisions to the one presently faced) are distilled into
  1. states of nature,
  2. a prior probability distribution, and
  3. a von-Neumann-Morgenstern expected utility function.
- Often hard to perform this distillation.
- Case-based decision theory is an alternative, especially apt for these difficult cases.

**Cases defined**

- Basic objects:
- $P$ is a set of problems,
- $A$ is a set of acts, and
- $R$ is a set of outcomes.
- A *case* is a triple $(p, a, r)$ such that $p \in P$, $a \in A$, and $r \in R$.

**Similarity of cases**

- The DM has some previous experience with cases.
- These past cases may be more or less relevant to the present case.
- We capture this by the notion of similarity of cases.
- The *similarity function* is a function $s : P \times P \to [0, 1]$.
- We interpret its typical value $s(p, q)$ as telling us how similar $p$ and $p$ are in the DM's mind.

**Utility**

- The DM has a utility function $u : R \to \mathbb{R}$.
- This function acts on the third part of a case $(p, a, r)$, $r$, which is the DM's outcome from this case.
- We interpret $u(r)$ as usual: it measures how much the DM likes the outcome $r$.

**Case-based decision theory at work**

- The DM is facing a problem $p \in P$.
- The DM knows all the acts in $A$.
- The DM remembers some of the cases previously encountered.
- The DM's memory: $M \subset C$. Possibly a proper subset.
- The DM evaluates acts according to

$$U_{p,M}(a) = \sum_{(q,a,r) \in M} s(p, q) u(r).$$

**Case-based decision theory at work continued**

- $U(a)$ is a weighted sum of utilities of outcomes.
- Similar to Expected Utility, but the weights here are similarities.
- Consider some act $a' \in A$. Say it appears in some cases in $M$, namely, the ones of the form $(q, a', r)$. So in each one of these cases, the act $a'$ was chosen.
- If $u(r) > 0$ for the $r$ appearing in the case $(q, a', r) \in M$, then the presence of the case $(q, a', r)$ in the DM's memory makes the act $a'$ more attractive in the problem $p$ faced now, in proportion to the similarity $s(p, q)$.

# 2 Social Networks

## 2.1 Basics of Social Networks

**Introducing social networks**

- Finite set of players $N = \{1, \ldots, n\}$.
- A *link* between $i \in N$ and $j \in N$ is $\{i, j\}$, denoted simply as $ij$. (Defined for $i \neq j$.)
- A link means there is some relationship between $i$ and $j$ that generates value; perhaps via trading or information exchange.
- We consider *undirected* links, meaning $ij = ji$.
- Set of all potential links $g_N = \{ij \,|\, i, j \in N \text{ and } i \neq j\}$.
- A *social network* is any collection of links $g \subset g_N$.
- A network is often represented by $n \times n$ matrix of 1 and 0 entries, the adjacency matrix. Undirected graphs have symmetric adjacency matrices.

**Importance of social networks**

- Social networks can model well the process of social influence and the transmission of information and commodities. They have been widely studied in sociology and lately by physicists.
- They are a natural object of study of any economist who cares about public or collective goods, and arguably any economist, period.

**Importance of social networks continued**

- Examples of fruitful study of networks in economics (from Jackson 2006):
  1. Networks of contacts in labor markets (Granovetter, weak ties);
  2. Networks and social interactions as factors in crime;
  3. Networks of trade;
  4. Networks in social insurance;
  5. Networks in disease transmission;
  6. Networks in language and culture;
  7. Networks of interactions between firms.

**More on networks**

- Adding a link to network $g$ is denoted by $g + ij$. Subtracting a link $ij$ is denoted by $g - ij$. Formally, $g + ij \equiv g \cup \{ij\}$ and $g - ij \equiv g \setminus \{ij\}$.
- Collection of all links $i$ is involved in, in network $g$: $L_i(g) \equiv \{ij \mid ij \in g \text{ and } i \neq j\}$.
- *Set of neighbors* of $i$ in network $g$: $N_i(g) \equiv \{ j \in N \mid ij \in L_i(g)\}$.
- *Degree* of $i$ in network $g$: the number of neighbors of $i$ in $g$, i.e., the cardinality of the set $N_i(g)$.

**Paths**

- A *path* in network $g$ between $i$ and $j$ is a sequence $i_1, \ldots, i_K$ such that for all $k \in \{1, \ldots, K-1\}$, $i_k i_{k+1} \in g$ and $i_1 = i$ and $i_K = j$.
- The *length* of such a path is $K - 1$, the number of links in the path.
- A network is *connected* if, for every $i, j$ in the network, there is a path connecting them.
- A *component* of a network $g$ is a maximal connected subgraph of $g$.
- The *distance* between two nodes $i$ and $j$ in a network, $d(i, j)$, is the minimum path length between $i$ and $j$ (set to infinity if $i$ and $j$ are not connected by any path in $g$).

**Diameter**

- The *diameter* of a network $g$ is the maximum distance between any two nodes in $g$. For a non-connected network, the diameter is infinite.

**Clustering**

- If $i$ has links to $j$ and $k$, how likely is it that there is a link $jk$? The percentage of "yes" answers is $i$'s *clustering coefficient*.
- The average clustering coefficient averages each node's clustering coefficient.

**Small Worlds**

- Milgram's 1967 Small Worlds experiment (Nebraska to East Coast). Maximum number of links in a successful chain: 12; median: 5.
- Many social networks exhibit Small Worlds (small average path length). But so do many other networks (power grids, neurons).

**High Clustering**

- Social networks usually exhibit higher clustering than other networks. E.g., Adamic (1999): average clustering coefficient of a portion of the WWW, 0.11 (it would be 0.0002 for a randomly generated network).
- Other examples:
  - Citations, 0.07.
  - Co-Author, 0.16.
  - Ham Radio, 0.06.

**Degree distributions**

- If a network is formed randomly, with each possible link occurring with probability $p$, then the degree distribution of the network should be approximately a binomial one, well approximated by a Poisson distribution.
- But some networks have "fat tails". Classic modern study of this, that started off the statistical physicists: Albert, Jeong and Barabasi (1999). For a portion of the WWW in the Notre Dame ".edu" domain, they found an approximately scale-free distribution, following the power law in which the relative frequency of nodes with degree $k$ was $k^{-\gamma}$, where $\gamma > 1$ was a parameter.
- (Reason for scale-free name: consider degree $k$ and degree $ck$, where $c$ and free are some positive integers. Then the relative frequencies are $k^{-\gamma}/(ck)^{-\gamma} = c^{\gamma}$. But this is also true for any other $k'$ and the same $c$.)

# 3 Network Formation Models

## 3.1 Two Approaches (Exposition from Jackson 2006)

**Two main approaches**

1. Random graph approach, leading to statistical physics techniques;

2. Economics approach, based on game theory.

## 3.2 Mathematics- and Physics-inspired Models

**Erdös-Rényi random graphs**

- Consider a set of nodes and then consider each possible link. Independently, make each link happen with probability $p$ and not happen with probability $1 - p$.

- For large $n$ and small $p$, a Poisson distribution is a good approximation of the degree distribution.

- Interesting feature: phase transisions. With $p < 1/n$, we get disjoint components; once $p > 1/n$, a giant component. When $p$ exceeds $\log(n)/n$, the graph becomes connected.

- Past the $p = \log(n)/n$ threshold, the diameter of the network declines from $\log(n)/\log\log(n)$ as $p$ increases.

- For large $n$ and not very large $p$, there is very little clustering. So these random graphs are not good models for many social networks.

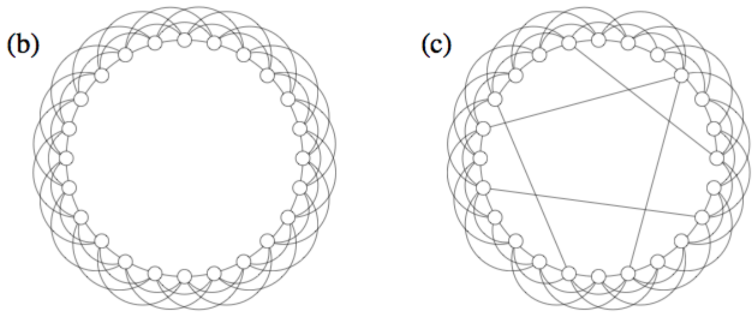
**Rewired lattices**

- Watts and Strogatz (1998) wanted a network formation process that yields "small worlds" (low diameter) and high clustering.

- Start with nodes on a lattice. E.g., on a circle, and each node linked to its 4 closest nodes. Now each node has clustering coefficient $1/2$ but the network diameter is too large, on the order of $n/4$.

- Watts and Strogatz start with such a network on a lattice, then randomly change a few links, so now some nodes are linked with some faraway nodes. With just a few such changes, the clustering declines a little and the diameter declines a lot.

**Rewired lattices illustrated**




From Newman, Mark, Models of the Small World, 2000.

**Preferential attachment**

- Preferential attachment is a process of network formation designed to grow networks with "fat tails", a concept dating back to Pareto and his namesake distribution (1896). The name "preferential attachment" is due to Barabasi and Albert (2001), whom it made famous.
- Time is discrete. A new node is born every period. (Only) upon birth, each node links to $m$ existing nodes. It selects them randomly, but gives probability proportional to the number of nodes the other node already has.
- For example, new websites often link much more often to established websites than to obscure ones.

## 3.3 Game-Theoretic Models of Network Formation

**Characteristics of game-theoretic models**

- Agents are the nodes and they derive utility from the network;
- Links are formed by the agents, and we predict the resulting network by some notion of equilibrium or by a possibly stochastic dynamic process.

**Early models**

- Myerson (1977). Studied a cooperative game, in which coalitions are the theory's primitives, and augmented these games with a graph structure. He came up with a natural modification of the Shapley value, now called the Myerson value.
- Aumann and Myerson (1998) set up and studied an extensive form game where links are considered in some fixed order. Hard to analyze.

**Jackson and Wolinsky 1996**

- Jackson and Wolinsky started a large trend in economic theory with their 1996 JET paper "A Strategic Model of Social and Economic Networks".
- Here the networks are the primitives. Once a network $g$ exists, each agent $i$ gets utility $u_i(g)$ from it. Note the domain of the utility functions: the space of all possible networks.
- Basic notion: pairwise stability. A network is *pairwise stable* if no agent wants to sever a link and no two players together want to add a link.

**Jackson and Wolinsky continued**

- Utilities are assumed additive across agents.
- Total value of network $g$ to society: $v(g) = \sum_i u_i(g)$.
- Network $g$ is *efficient relative to* $g'$ if $v(g) \geq v(g')$. Finite number of possible networks implies that an overall efficient network exists.
- Network $g$ is *Pareto efficient relative to* $(u_1, \ldots, u_n)$ by the standard definition (there is no other possible network $g'$ with $u_i(g') \geq u_i(g)$ with at least one strict inequality).

**Stability-efficiency conflict**

- Since each $u_i(g)$ depends on the entire networks, externalities are unavoidable.
- It follows that stable networks are not expected to be efficient.
- More surpising: it may not be possible to rectify this by transfers across agents.

**Connections model**

- The connections model is an example of a possible source of the utility of the network. Gilles, James, Barkhi and Diamantaras (2007) have looked at this model in the context of case-based network formation.
- In the connections model, each link is a social relationship between the linked agents, such as friendship.
- Agent $i$ gets a benefit from each direct link, and a lesser benefit from agents $i$ is indirectly linked to, friends of friends.

**Connections model utility function**

- Number of links in the shortest path from $i$ to $j$ in network $g$: $p_{ij}(g)$.
- Cost of establishing link $ij$: $c_{ij}$.
- Friendship benefit factor: $\delta_{ij} \in (0, 1)$.
- Utility function:

$$u_i(g) = \sum_{j \neq i : i \text{ and } j \text{ path-connected in g}} (\delta_{ij})^{p_{ij}(g)} - \sum_{j \neq i : ij \in g} c_{ij}.$$

## 3.4 Case-based Network Formation

**Source of memory**

- Basic idea behind the work presented here and the previously mentioned Gilles et al. paper: *an agent's memory is the network*.

**Link benefits, link costs**

- Simplest possible specification of benefits and costs for links.
- Agent $i$'s benefit from link $ij$: $\beta_i(ij) \geq 0$.
- Agent $i$'s cost from establishing link $ij$: $\gamma_i(ij) \geq 0$.
- Agent $i$'s net benefit from link $ij$: $\delta_i(ij) = \beta_i(ij) - \gamma_i(ij)$.

**Network utility function**

- From the net benefits already defined, we finally get the baseline utility function for each agent. For reasons that will be clear in a minute, we call it the network *payoff* function.
- Summing over the collection of all links that $i$ has in network $g$, the network payoff function is

$$\pi_i(g) = \sum_{ij \in L_i(g)} \delta_i(ij) = \sum_{ij \in L_i(g)} [\beta_i(ij) - \gamma_i(ij)].$$

**Strong pairwise stability**

**Definition 1.** A network $g$ is *strong pairwise stable* if it satisfies:

**Strong link deletion proofness** For every link $ij \in g$ we have $\pi_i(g - h) \leq \pi_i(g)$, where $h \subset L_i(g)$ and $g - h$ is the network that results after the removal of all links in $h$; and

**Modified link addition proofness** For every $ij \notin g$ we have that $\pi_i(g + ij) \geq \pi_i(g)$ implies $\pi_j(g + ij) < \pi_j(g)$.

**Case-based network formation ingredients**

- $N = \{1, 2, \ldots, n\}$, the set of players;
- $(\beta, \gamma)$, the link-based benefit and cost functions;
- $s : N \times N \times \mathcal{G}^{\mathcal{N}} \to \mathbb{R}_{++}$, the similarity function, which assigns to all players $i, j \in N$ a case-based similarity value $s(i, j, g) > 0$ within the context of a network $g$.

**Case-based utilities**

- If $ij \in g$, $U(i, j, g) = \delta_i(ij) = \beta_i(ij) - \gamma_i(ij)$.
- If $ij \notin g$, $U(i, j, g) = \sum_{h \in N_{i(g)}} s(i, h, g) \cdot \delta_h(hj)$.
- $U_i(g) = \sum_{j \neq i} U(i, j, g)$.

**Case-based decision making in network formation**

1. If $U(i, j, g) \geq 0$, $i$ says to $j$ "I want to link with you". If $j$ says the same to $i$, the link $ij$ is formed.
2. If the link $ij$ exists and either $U(i, j, g) < 0$ or $U(j, i, g) < 0$, the link is deleted.

**Simulations**

- It seems that the best way to understand this model is by numerical simulations on the computer.
- One of the authors (DD) learned the open-source computer language Python for this purpose, because it is known to be easy to learn, works on all platforms, and has a library of functions tailor-made for the study of networks.
- There is much more work to be done, but the preliminary results were intriguing (next few slides).
- Currently, our focus is on modularizing the computer code, adding a graphical user interface to make it easier to conduct experiments, planning a systematic exploration of the parameter space, and exploring different specifications of payoffs and similarities.

**Preliminary results**

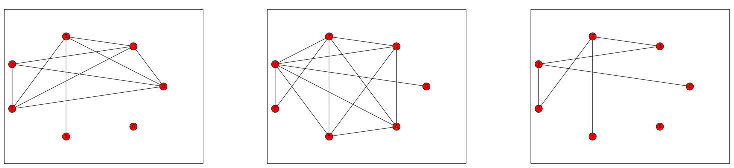
- Simulations started with a randomly generated network at $t = 0$.
- Utilities $U_i(g_t)$ were computed at each $t = 0, 1, 2, \ldots$
- The similarity function was $s(i, j, g_t) = [|U_i(g_t) - U_j(g_t)| + 1]^{-1}$.

- Links were added or deleted according to the double-agreement rule, where $i$ agrees to $ij$ if $U(i,j,g_t) \geq 0$ and both $i$ and $j$ must agree for addition. If the link $ij$ existed and either $U(i,j,g) < 0$ or $U(j,i,g) < 0$, the link was deleted.
- This resulted in network $g_{t+1}$.

**Preliminary results continued**

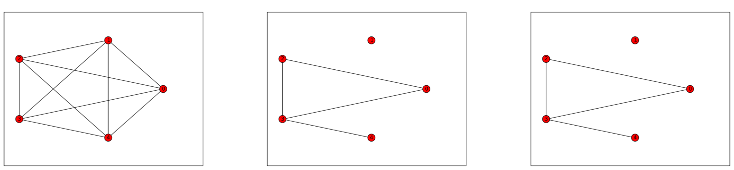
- We observed rapid convergence to a two-network cycle in all runs with few nodes (less than 20).
- The strong pairwise stable network was a subset of both networks in the cycle, and equal to one of the networks in the cycle in most cases. The other network in the cycle was the full network in most cases.
- With more nodes (100), we again saw rapid convergence to a two-cycle. One of the networks in the cycle was very similar in average clustering to the strong pairwise stable network.
- Although not speed-optimized yet, the program handles 100-node networks easily in one or two minutes per network, for several steps of network formation.

**Example 1**



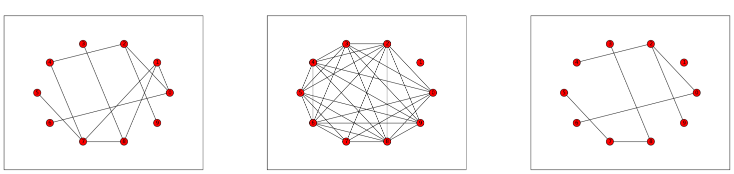
Cycle-1, Cycle-2, SPS (2006-06-10)

**Example 2**



Cycle-1, Cycle-2, SPS (2006-06-10)

**Example 3**



Cycle-1, Cycle-2, SPS (2006-06-10)

**Discussion of results**

- Results from 100 networks were similar, but harder to illustrate. We intend to find efficient ways to illustrate such networks.
- Reasonable explanation for the inclusion of the SPS graph in cycle networks: the leniency with which link proposals were accepted. For example, an isolated node has zero payoff; others with similar, nonnegative payoff see that and propose links to them, which are accepted.
- The payoff and similarity functions were simplistic. Future work will look at variations of these.
- The network may become only part of the memory of each agent; currently it is all the memory there is.

**Discussion continued**

- The network formation process studied was deterministic. Conjecture: with a small probability of experimentation for each node in each iteration, the SPS graph will result at the limit of the process. Such a result may also be provable analytically.