6.207/14.15: Networks
Lecture 16: Cooperation and Trust in Networks

Daron Acemoglu and Asu Ozdaglar
MIT

November 4, 2009

Outline

- The role of networks in cooperation
- A model of social norms
- Cohesion of groups and social norms
- Trust in networks

- Reading:
- Osborne, Chapters 14 and 15.

The Role of Social Networks

- Recall the importance of "social contacts" in finding jobs. Especially of "weak ties" (e.g., Granovetter (1973) "*The Strength of Weak Ties*": most people find jobs through acquaintances not close friends.
- The idea is that recommendations from people you know are more trusted.
- Similarly, social networks important in starting businesses?
- Recall that in many developing economies (but also even in societies with very strong institutions), networks of "acquaintances and contacts" shape business behavior. (e.g., Munshi (2009) "*Strength in Numbers: A Network-Based Solution to Occupational Traps*").
- The Indian diamond industry 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.

Trust in Networks

- The rise of the Kathiawaris most likely related to their close-knit network.
- When 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.
- Recall that Munshi's argument was that network connections helped the Kathiawaris pull ahead of the richer and more established Marwaris and Palanpuris.

Trust in Networks (continued)

- Perhaps trust is more difficult when the network is larger.
- 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.
- 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.

Social Norms

- Even in broader social groupings, some types of implicit understanding on expected behavior important.
- We sometimes refer to these as *social norms*: how to dress, how to interact with others, limits on socially costly selfish behavior, etc.
- How are they supported?
- This lecture: using repeated games to understand social norms and trust in social networks.

Modeling Social Norms

- We will think of social norms as the convention—expected play—in the game. The key question is whether a particular social norm is **sustainable** as the equilibrium in society.
- Consider a society consisting of N players playing an infinitely-repeated **symmetric** two-player strategic form game $G = \langle \mathcal{I}, A, u \rangle$.
- Throughout N is a large number.
- Here A denotes the set of actions at each stage, and thus

$$u_i : A \times A \to \mathbb{R}.$$

- That is, $u\left(a_i^t, a_j^t\right)$ is the state payoff to player i when action profile $a^t = \left(a_i^t, a_j^t\right)$ is played at stage t between players i and $j \neq i$.
- We will think of a social norm simply as an action $a^* \in A$ that all players are expected to play.

Modeling Social Norms (continued)

- Suppose, to start with, that players are **matched randomly** at each date (you may wish to think that N is even).
- Let $\mathbf{a}_i$ be the sequence of plays for player i, i.e., $\mathbf{a}_i = \left\{ \left(a_i^t, a_{j(i,t)}^t \right) \right\}_{t=0}^{\infty}$, where $j(i,t)$ denotes the player matched to i at time t.
- The payoff of player i is then

$$U(\mathbf{a}_i) = \sum_{t=0}^{\infty} \delta^t u(a_i^t, a_{j(i,t)}^t)$$

where $\delta \in [0,1)$ is again the discount factor.

Full Monitoring

- **Full monitoring** applies when players observe the entire history of past actions.
- For example, they observe the entire history of play in each random match.
- With full monitoring, the following **personalized trigger strategies** are possible.
- If individual i deviates from the social norm a^* at time t, everybody observes this, and will play some punishment action $\underline{a} \in A$ against i (they can still cooperate with other players).
- Then the arguments from standard repeated games (in particular the folk theorems) immediately imply the following theorem.

Full Monitoring Theorem

Theorem

Let a^{NE} be a static equilibrium of the stage game. With full monitoring, for any $a \in A$ with $u(a, a) > u(a^{NE}, a^{NE})$, there exists some $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, there exists an equilibrium supporting social norm a.

Proof (essentially identical to the proof of the folks theorems):

- Deviation has some benefit $\bar{u}$ now and thus overall return

$$\bar{u} + \delta \frac{u(a^{NE}, a^{NE})}{1-\delta},$$

  since all other players will punish the deviator (e.g., playing the NE).
- Cooperation has return $\frac{u(a,a)}{1-\delta}$.
- Therefore,

$$\delta \geq \underline{\delta} \equiv \frac{\bar{u} - u(a^{NE}, a^{NE})}{\bar{u} - u(a, a)} \in (0, 1)$$

  guarantees that the social norm of cooperation is sustainable.

Application

- Recall the prisoners' dilemma:

	Cooperate	Defect
Cooperate	$1, 1$	$-1, 2$
Defect	$2, -1$	$0, 0$

- In this game, (C, C), that is "cooperation," can be supported as the social norm in society when $\delta \geq 1/2$.

Problems with Full Monitoring

- Full monitoring too "unrealistic". It means that social norms are supported by each individual knowing what everybody else in the society does.
- More likely, individuals know their own experiences, and perhaps what is happening to their neighbors, friends and coworkers.
- This implies **social network** structure will matter.

Pure Private Histories

- The other extreme from full monitoring is a situation in which each individual only observes what has happened to themselves.
- For example, in the prisoners' dilemma game, a player will be matched with a different partner every period, and then play cooperate or defect, and will only know his or her experience in the past.
- With N large, one might first conjecture that cooperation is impossible to sustain in this society.
  - Either individuals do not even remember the identity of their past partners (fully anonymous), in which case one might conjecture that the strategy of "defect" will have no future cost for a player (if there was indeed the social norm of cooperation).
  - Or individuals remember the identity of their past partners, but in this case, the cost of having "a bad reputation" against a single player is not too high, since with N large, this player will not be met again in the future.

Pure Private Histories (continued)

- However, Kandoori (1994) showed that **contagion strategies** can support cooperation.
- Contagion strategies involve each player defecting in all future periods if they observe any deviation in their private history.
- The name "contagion" comes from the fact that, under this strategy, defection will spread in a contagious manner and "invade" cooperative behavior.
- Therefore, a player who defects will recognize that ultimately everybody in the society will start defecting because of contagion, and this will have a negative effect on her future payoffs.
- If the discount factor is sufficiently close to 1, defecting will not be profitable.

Pure Private Histories (continued)

Theorem

Let a^{NE} be a static equilibrium of the stage game. With private monitoring, for any $a \in A$ with $u(a, a) > u\left(a^{NE}, a^{NE}\right)$ and any $N < \infty$, there exists some $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, the social norm of playing a can be supported.

- A proper proof of this theorem requires us to consider dynamic games of incomplete information, since when a player experiences a defection (an action different from a), she does not know whether this is because her partner is deviating or whether because she is in the middle of a contagious phase.
- But verifying that it is a Nash equilibrium can be done with the concepts we have developed so far.

Proof Sketch

- First, we give the idea of the proof. Suppose all other players are cooperating, in the sense of playing a. Then cooperation has payoff

$$\frac{u(a,a)}{1-\delta}.$$

- If a player deviates, she will obtain $u(a',a) > u(a,a)$ today and also again in the future against all others who are still playing a. Against others who have been reached by the contagious deviation, she will obtain no more than $\underline{u} = \max\left\{u\left(a', a^{NE}\right), u\left(a^{NE}, a^{NE}\right)\right\} < u(a,a)$.
- Loosely speaking, because the society is finite ($N < \infty$), almost all players will be ultimately reached by the contagious deviation in finite time.
- Therefore for δ arbitrarily close to 1, cooperation is better than defection.

Proof Sketch (continued)

- More explicitly, the contagion process induces the number of agents who are playing a^{NE} to increase as in the Bass model of disease diffusion.
- In particular, let $x(t)$ denote the number of agents playing a^{NE}.
- Each agent currently playing a (who has not been reached by the contagion) has probability equal to $x(t)/(N-1)$ of matching against an agent playing a^{NE} and thus switching his behavior thereafter.
- With a mean field type approximation, the law of motion of $x(t)$ can therefore be written as

$$\begin{aligned} x(t+1) &\simeq x(t) + (N - x(t)) \times \frac{x(t)}{N-1}, \\ &\simeq x(t) + x(t) \times \frac{N - x(t)}{N}. \end{aligned}$$

Proof Sketch (continued)

- With a differential equation approximation, we have
$$\dot{x}(t) \simeq x(t)\left(1 - \frac{x(t)}{N}\right)$$
with initial condition $x(0) = 1$.
- The solution to this differential equation is
$$x(t) = \frac{N\left(e^{t+c}\right)}{1 + e^{t+c}},$$
where c is the constant of integration, given by
$$c = \log\left(\frac{1}{N-1}\right).$$
- Thus alternatively,
$$x(t) = \frac{Ne^{t}}{N - 1 + e^{t}},$$
- Clearly, $x(t) \to N$ as $t \to \infty$.

Proof Sketch (continued)

- Moreover, for $T \geq T'$, we have $x(t) \geq N - M$, where
$$T' = \log(N - M) + \log N - \log M.$$
- An upper bound on the payoff to deviating is
$$U^d = \frac{1 - \delta^{T'+1}}{1 - \delta} u\left(a', a^{NE}\right) + \delta^{T'} \tilde{u},$$
where
$$\tilde{u} = \max_{\tilde{a} \in A} \left\{ \frac{N - M}{N} u\left(\tilde{a}, a^{NE}\right) + \frac{M}{N} u(\tilde{a}, a) \right\}$$
is the maximum payoff that the deviator can obtain after time T', where at least $N - M$ people have switched to the Nash equilibrium of the stage game with probability arbitrarily close to 1. For N sufficiently large and M sufficiently small, we have $\tilde{u} < u(a, a)$.
- This is an upper bound on deviation payoff, since in reality the deviator will not obtain $u\left(a', a^{NE}\right)$ for the T' periods.

Proof Sketch (continued)

- Now comparing U^d to the payoff from following the social norm, $U^c = u(a, a) / (1 - \delta)$, we have that the social norm of playing a will be sustainable if

$$u(a, a) \geq \left(1 - \delta^{T'+1}\right) u\left(a', a^{NE}\right) + (1 - \delta)\, \delta^{T'} \tilde{u}.$$

- Clearly, this inequality is satisfied as $\delta \to 1$. Therefore, there exists $\underline{\delta} < 1$ such that if

$$\delta \geq \underline{\delta},$$

this inequality is still satisfied and the social norm of cooperation can be supported with private histories.

Value of Small Groups

- The example of the Kathiawaris suggests that perhaps the social norm of cooperation in small groups is easier to sustain.
- Intuitively, this is easy to answer.
- Imagine a society has N members. Contagion will take a long time if N is large. But if N is divided into N/M groups, most interacting within themselves, then contagion of your partners will be faster.
- As a result, cooperation can be sustained for smaller values of the discount factor δ.

Value of Local Interactions

- Is it just the size of the groups or the structure of interactions?
- In social networks, local interactions are important.
- For example, society consisting of N members will exhibit different behavior when there is random matching vs. when individuals just interact with their neighbors over a circle.
- Typically, local interactions facilitate sustaining the social norm of cooperation.

Example

- Consider the prisoners' dilemma played on the circle. At each date, a player will play this stage game with one of his two neighbors.

	Cooperate	Defect
Cooperate	$1, 1$	$-1, 2$
Defect	$2, -1$	$0, 0$

- Optimal deviation would be to defect against your opponent at $t = 0$ and then against your other opponent whenever you are matched with them.
- What is the likelihood that your other neighbor starts defecting before you are matched with them?
- For a sufficiently large Circle, there is probability approximately equal to 1 that your other neighbor is still playing cooperate by the time you match with them.

Example (continued)

- Then expected utility from defection is

$$
\begin{aligned}
U^d &= 2 + \frac{1}{2}\delta\left(2 + 0 + ...\right) \\
&\quad + \frac{1}{2}\delta\left(0 + \frac{1}{2}\delta\left(2 + 0 + ...\right) + \frac{1}{2}\delta\left(0 + \frac{1}{2}\delta\left(2 + 0 + ...\right) + ...\right)\right) \\
&= 2 + \frac{\delta}{1 - \delta/2}.
\end{aligned}
$$

- Intuition: each period there is probability $1/2$ that you will match with your other neighbor who is still playing cooperate.
- Expected utility from cooperation is

$$U^c = \frac{1}{1-\delta}.$$

- Therefore, cooperation can be sustained as a social norm if $\delta \geq 2/3$.

Example (continued)

- Now imagine the same structure with three neighbors. With the same reasoning

$$\begin{aligned} U^d &= 2 + \frac{2}{3}\delta\left(2 + \frac{1}{3}\delta\left(2 + 0 + \ldots\right) + \frac{2}{3}\delta\left(0 + \frac{1}{3}\delta\left(2 + 0 + \ldots\right)\right)\ldots\right) \\ &\quad + \frac{1}{3}\delta\left(0 + \frac{2}{3}\delta\left(2 + \frac{1}{3}\delta\left(2 + 0\right) + \ldots\right) + \ldots\right) \end{aligned}$$

- With similar computations, cooperation cannot be sustained as the social norm if δ is greater than approximately $8/9$.
- So having three neighbors instead of two significant increases the threshold.

Cohesiveness

- Does the "cohesiveness" of a group matter?
- Different ways of thinking of cohesiveness.
- One possibility is that cohesive groups do not interact well with other groups.
- This type of cohesiveness may facilitate cooperation.

Modified Prisoners' Dilemma

- Consider a society consisting of M groups.
- Histories are observable within groups, so if an individual defects, all group members learn this immediately and can play accordingly.
- Other groups do not possess this information.
- Suppose the game is given by

	Cooperate	Defect	Punish
Cooperate	$3x, 3x$	$-x, 6x$	$0, 0$
Defect	$6x, -x$	$2x, 2x$	$0, 0$
Punish	$0, 0$	$0, 0$	$0, 0$

- Cohesiveness is captured by the fact that $x = x_h$ when playing against your own group and $x = x_l < x_h$ when playing against another group ($x_l > 0$).

The Effects of Cohesiveness

- Suppose that an individual can play this game within his group or leave his group and play against a member of one of the other groups.
- The minmax payoff here is 0. So the group can coordinate to hold down a defector to a payoff of zero.
- Individuals that decide to leave their group are randomly matched to a member of a different group and are anonymous, and the outcome of this game is not observed to other group members of either player. This implies that minmax strategies are not possible when two individuals from different groups play.
- Therefore, optimal to quit the group after defection.
- When can cooperation be sustained?

The Effects of Cohesiveness (continued)

- Without defection in any group, each individual will play within his group, and thus opting payoff of

$$U^c = \frac{3x_h}{1-\delta}.$$

- After defection, individual will leave and play against somebody from a different group. Because this game is anonymous, both players will defect, and thus the payoff to defection is

$$U^d = 6x_h + \delta \frac{2x_l}{1-\delta}.$$

The Effects of Cohesiveness (continued)

- Therefore, $U^c \geq U^d$ if

$$\frac{3x_h}{1-\delta} \geq 6x_h + \delta \frac{2x_l}{1-\delta}$$

or if

$$\delta \geq \frac{3}{6 - 2x_l/x_h}.$$

- If $x_l/x_h = 0$, cohesiveness is strong enough that *outside options* are as bad as staying in the group after defecting, and the social norm of cooperation can be sustained for any $\delta \geq 1/2$. If x_l/x_h is close to 1, then outside options after cheating are good, and the social norm of cooperation can be sustained only if $\delta \geq 3/4$.

Trust in Networks

- Social networks also can act as conduits of "trust".
- Consider a network consisting of three individuals, 1, 2 and 3. 1 and 2 and 2 and 3 interact frequently and trust each other.
- Suppose now that there is a transaction between 1 and 3.
- Can they leverage the fact that they both know 2 and use this to trust each other?
- How is this trust sustained? We would require that 2 observes the outcome of the interaction between 1 and 3, and switches from cooperating with whoever defects in the relationship between 1 and 3.
- Idea related to importance of **social capital**.

Example

- Let us again take prisoners' dilemma:

	Cooperate	Defect
Cooperate	$1,1$	$-1,2$
Defect	$2,-1$	$0,0$

- Suppose that at each date there is probability p_{ij} that i and j match and play this game.
- Assume that all draws are independent and a player could match with two others within the same date (thus no correlation between matches).
- All players have discount factor δ. Suppose that p_{13} is small.
- All histories are commonly observed.
- We will contrast the situation in which players 1 and 3 leverage their relationship with 2 vs. the one in which they do not.

Example (continued)

- If there is no leveraging of trust from the network, the pairwise relationship between 1 and 3 would work as follows.
- If they cooperate, player 1 would obtain a return of

$$U^c = 1 + \delta p_{13} + p_{13}\delta^2 + p_{13}\delta^3 + ... = 1 + \frac{\delta p_{13}}{1-\delta}.$$

- Defection has payoff

$$U^d = 2.$$

- Therefore, if

$$\delta < \frac{1}{1+p_{13}},$$

cooperation between 1 and 3 is not possible. For p_{13} small enough, meaning a sufficiently weak link between these two players, there will not be cooperation (provided that $\delta < 1$).

Example (continued)

- Now imagine that there is leveraging of trust.
- This means that if player 1 or 3 defect against each other, 2 will also defect against them in the future. Clearly, this trigger strategy is subgame perfect.
- Now we have to analyze the relationship between all players simultaneously. In this case, for player $i = 1$ or 3, we have

$$\begin{aligned}\bar{U}_i^c &= 1 + \delta\left(p_{i2} + p_{ij}\right) + \left(p_{i2} + p_{ij}\right)\delta^2 + \left(p_{i2} + p_{ij}\right)\delta^3 + \ldots \\ &= 1 + \frac{\delta\left(p_{i2} + p_{ij}\right)}{1 - \delta},\end{aligned}$$

where $j \neq i, 2$.

Example (continued)

- Defection for player i against player j then gives

$$\bar{U}_i^d = 2,$$

since player 2 will also play defect thereafter.

- Now, cooperation between 1 and 3 requires

$$\bar{U}_i^c \geq \bar{U}_i^d,$$

or

$$\delta \geq \frac{1}{1 + (p_{i2} + p_{ij})}.$$

- Even if p_{13} is small, this condition will be satisfied provided that δ, p_{12}, and p_{23} are sufficiently high.
- This is an example of leveraging the network to obtain trust between two weakly connected individuals.

General Insights

- What types of networks will foster trust?
- We can repeat the same analysis with a general weighted graph representing interaction structures within a society (group). Then cooperation (trust) within the society can be sustained if

$$\delta \geq \frac{1}{1 + \sum_{j \neq i} p_{ij}} \text{ for all } i.$$

- Therefore, **generalized trust** can be supported if the social network has sufficient interactions for all players.

General Insights (continued)

- Alternatively, we can have trust only among some players.
- Suppose the set of edges $\{i, j\} \in E$ between which trust can be supported is denoted by E^T (with the convention that $\{i, i\} \notin E^T$), then we would require that: for all $\{i, j\} \in E^T$,

$$\delta \geq \frac{1}{1 + \left(p_{ij} + \sum_{\{i,k\} \in E^T, k \neq j} p_{ik}\right)} \text{ and } \geq \frac{1}{1 + \left(p_{ij} + \sum_{\{j,k\} \in E^T, k \neq i} p_{jk}\right)}$$

- This would ensure that neither of the two players wish to deviate for any $\{i, j\} \in E^T$.
- But this also implies that E^T must be a completely connected subgraph of the original graph, in the sense that if $\{i, j\} \in E^T$ and $\{i, k\} \in E^T$, then we also have $\{j, k\} \in E^T$. (Why?)