6.207/14.15: Networks
Lectures 22-23: Social Learning in Networks

Daron Acemoglu and Asu Ozdaglar
MIT

December 2 end 7, 2009

Outline

- Recap on Bayesian social learning
- Non-Bayesian (myopic) social learning in networks
- Bayesian observational social learning in networks
- Bayesian communication social learning in networks

- Reading:
- Jackson, Chapter 8.
- EK, Chapter 16.

Introduction

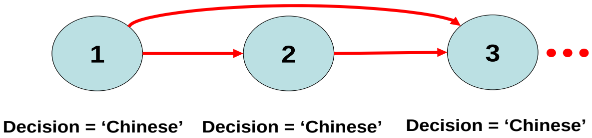
- How does network structure and "influence" of specific individuals affect opinion formation and learning?
- To answer this question, we need to extend the simple example of herding from the previous literature to a network setting.
- Question: is Bayesian social learning the right benchmark?
  - Pro: Natural benchmark and often simple heuristics can replicate it
  - Con: Often complex
- Non-Bayesian myopic learning: (rule-of-thumb)
  - Pro: Simple and often realistic
  - Con: Arbitrary rules-of-thumb, different performances from different rules, how to choose the right one?

What Kind of Learning?

- *What do agents observe?*
  - Observational learning: observe past actions (as in the example)
    - Most relevant for markets
  - Communication learning: communication of beliefs or estimates
    - Most relevant for friendship networks (such as Facebook)
  - The model of social learning in the previous lecture was a model of Bayesian observational learning.
  - It illustrated the possibility of herding, where everybody copies previous choices, and thus the possibility that dispersely held information may fail to aggregate.

Recap of Herding

- Agents arrive in town sequentially and choose to dine in an Indian or in a Chinese restaurant.
- A restaurant is strictly better, underlying state $\theta \in \{Chinese, Indian\}$.
- Agents have independent binary private signals.
- Signals indicate the better option with probability $p > 1/2$.
- Agents observe prior decisions, but not the signals of others.
- Realization: Assume $\theta = Indian$
  - Agent 1 arrives. Her signal indicates 'Chinese'. She chooses Chinese.
  - Agent 2 arrives. His signal indicates 'Chinese'. He chooses Chinese.
  - Agent 3 arrives. Her signal indicates 'Indian'. She disregards her signal and copies the decisions of agents 1 and 2, and so on.



Potential Challenges

- Perhaps this is too "sophisticated".
- What about communication? Most agents not only learn from observations, but also by communicating with friends and coworkers.
- Let us turn to a simple model of myopic (rule-of-thumb) learning and also incorporate network structure.

Myopic Learning

- First introduced by DeGroot (1974) and more recently analyzed by Golub and Jackson (2007).
- Beliefs updated by taking weighted averages of neighbors' beliefs
- A finite set $\{1, \ldots, n\}$ of agents
- Interactions captured by an $n \times n$ nonnegative interaction matrix T
    - $T_{ij} > 0$ indicates the trust or weight that i puts on j
    - T is a stochastic matrix (row sum=1; see below)
- There is an underlying state of the world $\theta \in \mathbb{R}$
- Each agent has initial belief $x_i(0)$; we assume $\theta = 1/n \sum_{i=1}^{n} x_i(0)$
- Each agent at time k updates his belief $x_i(k)$ according to

$$x_i(k+1) = \sum_{j=1}^{n} T_{ij} x_j(k)$$

What Does This Mean?

- Each agent is updating his or her beliefs as an average of the neighbors' beliefs.
- Reasonable in the context of one shot interaction.
- Is it reasonable when agents do this repeatedly?

Stochastic Matrices

Definition

T is a stochastic matrix, if the sum of the elements in each row is equal to 1, i.e.,

$$\sum_j T_{ij} = 1 \text{ for all } i.$$

Definition

T is a doubly stochastic matrix, if the sum of the elements in each row and each column is equal to 1, i.e.,

$$\sum_j T_{ij} = 1 \text{ for all } i \text{ and } \sum_i T_{ij} = 1 \text{ for all } j.$$

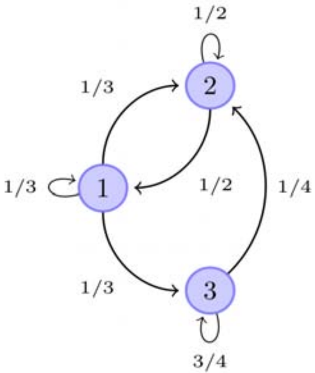
- Throughout, assume that T is a stochastic matrix. Why is this reasonable?

Example

- Consider the following example

$$T = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 \\ 0 & 1/4 & 3/4 \end{pmatrix}$$

- Updating as shown



Example (continued)

- Suppose that initial vector of beliefs is

$$x(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

- Then updating gives

$$x(1) = Tx(0) = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 \\ 0 & 1/4 & 3/4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/2 \\ 0 \end{pmatrix}$$

Example (continued)

- In the next round, we have

$$x(2) = Tx(1) = T^2 x(0) = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 \\ 0 & 1/4 & 3/4 \end{pmatrix} \begin{pmatrix} 1/3 \\ 1/2 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 5/18 \\ 5/12 \\ 1/8 \end{pmatrix}$$

- In the limit, we have

$$x(n) = T^n x(0) \to \begin{pmatrix} 3/11 & 3/11 & 5/11 \\ 3/11 & 3/11 & 5/11 \\ 3/11 & 3/11 & 5/11 \end{pmatrix} x(0) = \begin{pmatrix} 3/11 \\ 3/11 \\ 3/11 \end{pmatrix}.$$

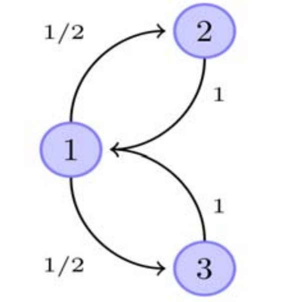
- Note that the limit matrix, $T^* = \lim_{n\to\infty} T^n$ has identical rows.
- Is this kind of convergence general? Yes, but with some caveats.

Example of Non-convergence

- Consider instead

$$T = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

- Pictorially



Example of Non-convergence (continued)

- In this case, we have
  - For n even:

$$T^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}.$$

  - For n odd:

$$T^n = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

  - Thus, **non-convergence.**

Convergence

- Problem in the above example is **periodic** behavior.
- It is sufficient to assume that $T_{ii} > 0$ for all i to ensure aperiodicity. Then we have:

Theorem

Suppose that T defines a strongly connected network and $T_{ii} > 0$ for each i, then $\lim_n T^n = T^*$ exists and is unique. Moreover, $T^* = e\pi'$, where e is the unit vector and π is an arbitrary row vector.

- In other words, T^* will have identical rows.
- An immediate corollary of this is:

Proposition

In the myopic learning model above, if the interaction matrix T defines a strongly connected network and $T_{ii} > 0$ for each i, then there will be consensus among the agents, i.e., $\lim_{n\to\infty} x_i(n) = x^*$ for all i.

Learning

- But consensus is not necessarily a good thing.
- In the herding example, there is consensus (of sorts), but this could lead to the wrong outcome.
- We would like consensus to be at

$$x^* = \frac{1}{n}\sum_{i=1}^{n} x_i(0) = \theta,$$

so that individuals learn the underlying state. If this happens, we say that the society is wise.

When Will There Be Learning?

- Somewhat distressing result:

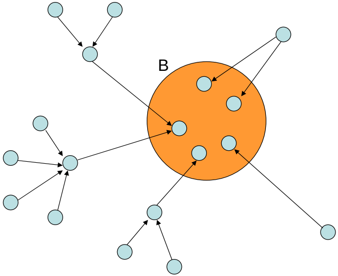
Proposition

In the myopic learning model, the society is wise if and only if T is doubly stochastic.

- Intuition: otherwise, there is no balance in the network, so some agents are influential; their opinion is listened to more than they listen to other people's opinion.
- Is this a reasonable model for understanding the implications of influence?

Influential Agents and Learning

- A set of agents B is called an influential family if the beliefs of all agents outside B is affected by beliefs of B (in finitely many steps)



- The previous proposition shows that the presence of influential agents implies no asymptotic learning
  - The presence of influential agents is the same thing as lack of doubly stochasticity of T
  - Interpretation: Information of influential agents overrepresented
- Distressing result since influential families (e.g., media, local leaders) common in practice

Towards a Richer Model

- Too myopic and mechanical: If communicating with same people over and over again (deterministically), some recognition that this information has already been incorporated.
- No notion of misinformation or extreme views that can spread in the network.
- No analysis of what happens in terms of quantification of learning without doubly stochasticity

A Model of Misinformation

- Misinformation over networks from Acemoglu, Ozdaglar, ParandehGheibi (2009)
- Finite set $\mathcal{N} = \{1, \dots, n\}$ of agents, each with initial belief $x_i(0)$.
- Time continuous: each agent recognized according to iid Poisson processes.
- $x_i(k)$: belief of agent i after k^{th} communication.
- Conditional on being recognized, agent i meets agent j with probability p_{ij}:
  - With probability β_{ij}, the two agents agree and exchange information
  $$x_i(k+1) = x_j(k+1) = (x_i(k) + x_j(k))/2.$$
  - With probability γ_{ij}, disagreement and no exchange of information.
  - With probability α_{ij}, i is influenced by j
  $$x_i(k+1) = \epsilon x_i(k) + (1-\epsilon)x_j(k)$$
  for some $\epsilon > 0$ small. Agent j's belief remains unchanged.
- We say that j is a forceful agent if $\alpha_{ij} > 0$ for some i.

Evolution of Beliefs

- Letting $x(k) = [x_1(k), \ldots, x_n(k)]$, evolution of beliefs written as

$$x(k+1) = W(k)x(k),$$

where $W(k)$ is a random matrix given by

$$W(k) = \begin{cases} A_{ij} \equiv I - \frac{(e_i - e_j)(e_i - e_j)'}{2} & \text{with probability } p_{ij}\beta_{ij}/n, \\ J_{ij} \equiv I - (1-\epsilon)\, e_i(e_i - e_j)' & \text{with probability } p_{ij}\alpha_{ij}/n, \\ I & \text{with probability } p_{ij}\gamma_{ij}/n, \end{cases}$$

where e_i is the ith unit vector (1 in the ith position and 0s everywhere else).

- The matrix $W(k)$ is a (row) stochastic matrix for all k, and is iid over all k, hence

$$E[W(k)] = \tilde{W} \qquad \text{for all } k \geq 0.$$

- We refer to the matrix $\tilde{W}$ as the mean interaction matrix.

Social Network and Influence Matrices

- Using the belief update model, we can decompose $\tilde{W}$ as:

$$\begin{aligned}\tilde{W} &= \frac{1}{n}\sum_{i,j} p_{ij}\left[\beta_{ij}A_{ij} + \alpha_{ij}J_{ij} + \gamma_{ij}I\right] \\ &= \frac{1}{n}\sum_{i,j} p_{ij}\left[(1-\gamma_{ij})A_{ij} + \gamma_{ij}I\right] + \frac{1}{n}\sum_{i,j} p_{ij}\alpha_{ij}\left[J_{ij} - A_{ij}\right] \\ &= T + D.\end{aligned}$$

- Matrix T represents the underlying social interactions: social network matrix
- Matrix D represents the influence structure in the society: influence matrix
- Decomposition of $\tilde{W}$ into a doubly stochastic and a remainder component
- Social network graph: the undirected (and weighted) graph $(\mathcal{N}, \mathcal{A})$, where $\mathcal{A} = \{\{i,j\} \mid T_{ij} > 0\}$, and the edge $\{i,j\}$ weight given by $T_{ij} = T_{ji}$

Assumptions

- Suppose, in addition, that the graph $(\mathcal{N}, \mathcal{E})$, where $\mathcal{E} = \{(i,j) \mid p_{ij} > 0\}$, is strongly connected; otherwise, no consensus is automatic.
- Moreover, suppose that

$$\beta_{ij} + \alpha_{ij} > 0 \qquad \text{for all } (i,j) \in \mathcal{E}.$$

  - Positive probability that even forceful agents obtain information from the other agents in the society.
  - Captures the idea that "no man is an island"

Convergence to Consensus

Theorem

The beliefs $\{x_i(k)\}$, $i \in \mathcal{N}$ converge to a consensus belief, i.e., there exists a random variable $\bar{x}$ such that

$$\lim_{k\to\infty} x_i(k) = \bar{x} \qquad \text{for all } i \text{ with probability one.}$$

Moreover, there exists a probability vector $\bar{\pi}$ with $\lim_{k\to\infty} \tilde{W}^k = e\bar{\pi}'$, such that

$$E[\bar{x}] = \sum_{i=1}^{n} \bar{\pi}_i x_i(0) = \bar{\pi}' x(0).$$

- Convergence to consensus guaranteed; consensus belief is a random variable.
- We are interested in providing an upper bound on

$$E\left[\bar{x} - \frac{1}{n}\sum_{i\in\mathcal{N}} x_i(0)\right] = \sum_{i\in\mathcal{N}} \left(\bar{\pi}_i - \frac{1}{n}\right) x_i(0).$$

- $\bar{\pi}$: consensus distribution, and $\bar{\pi}_i - \frac{1}{n}$: excess influence of agent i

Global Bounds on Consensus Distribution

Theorem

Let π denote the consensus distribution. Then,

$$\left\|\pi - \frac{1}{n}e\right\|_2 \leq \frac{1}{1-\lambda_2}\frac{\sum_{i,j} p_{ij}\alpha_{ij}}{n},$$

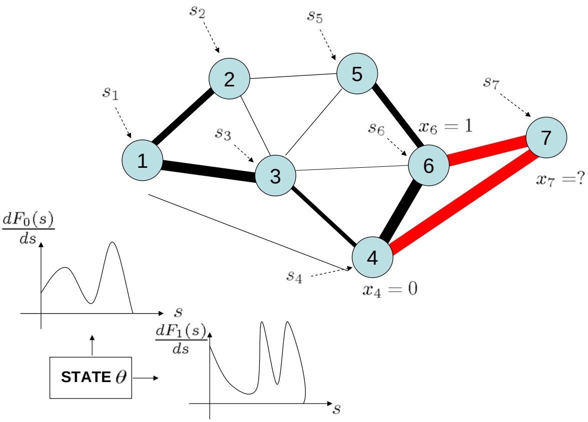
where λ_2 is the second largest eigenvalue of the social network matrix T.

- Proof using perturbation theory of Markov Chains
  - View $\tilde{W}$ as a perturbation of matrix T by the influence matrix D
- λ_2 related to mixing time of a Markov Chain
  - When the spectral gap $(1-\lambda_2)$ is large, we say that the Markov Chain induced by T is fast-mixing
- In fast-mixing graphs, forceful agents will themselves be influenced by others (since $\beta_{ij}+\alpha_{ij}>0$ for all i, j)
  - Beliefs of forceful agents moderated by the society before they spread

Bayesian Social Learning

- Learning over general networks; Acemoglu, Dahleh, Lobel, Ozdaglar (2008).
- Two possible states of the world $\theta \in \{0, 1\}$, both equally likely
- A sequence of agents $(n = 1, 2, ...)$ making decisions $x_n \in \{0, 1\}$.
- Agent n obtains utility 1 if $x_n = \theta$, and utility 0 otherwise.
- Each agent has an iid private signal s_n in S. The signal is generated according to distribution $\mathbb{F}_\theta$ (signal structure)
- Agent n has a neighborhood $B(n) \subseteq \{1, 2, ..., n-1\}$ and observes the decisions x_k for all $k \in B(n)$.
    - The set $B(n)$ is private information.
- The neighborhood $B(n)$ is generated according to an arbitrary distribution $\mathbb{Q}_n$ (independently for all n) (network topology)
    - The sequence $\{\mathbb{Q}_n\}_{n \in \mathbb{N}}$ is common knowledge.
- Asymptotic Learning: Under what conditions does $\lim_{n\to\infty} \mathbb{P}(x_n = \theta) = 1$?

An Example of a Social Network



Perfect Bayesian Equilibria

- Agent n's information set is $I_n = \{s_n, B(n), x_k \text{ for all } k \in B(n)\}$
- A strategy for individual n is $\sigma_n : \mathcal{I}_n \to \{0, 1\}$
- A strategy profile is a sequence of strategies $\sigma = \{\sigma_n\}_{n \in \mathbb{N}}$.
  - A strategy profile σ induces a probability measure $\mathbb{P}_\sigma$ over $\{x_n\}_{n \in \mathbb{N}}$.

Definition

A strategy profile σ^* is a pure-strategy Perfect Bayesian Equilibrium if for all n

$$\sigma_n^*(I_n) \in \arg\max_{y \in \{0,1\}} \mathbb{P}_{(y, \sigma^*_{-n})}(y = \theta \mid I_n)$$

- A pure strategy PBE exists. Denote the set of PBEs by Σ^*.

Definition

We say that asymptotic learning occurs in equilibrium σ if x_n converges to θ in probability,

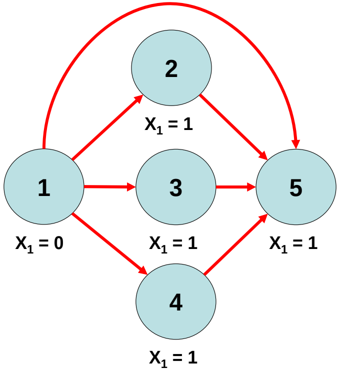
$$\lim_{n \to \infty} \mathbb{P}_\sigma(x_n = \theta) = 1$$

Some Difficulties of Bayesian Learning

- No following the crowds

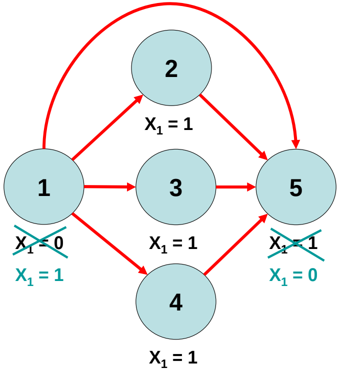
Some Difficulties of Bayesian Learning

- No following the crowds



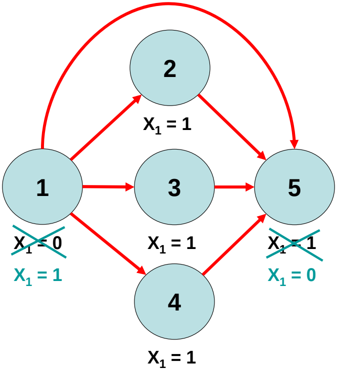
Some Difficulties of Bayesian Learning

- No following the crowds

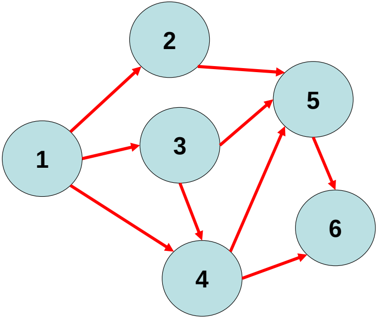


Some Difficulties of Bayesian Learning

- No following the crowds

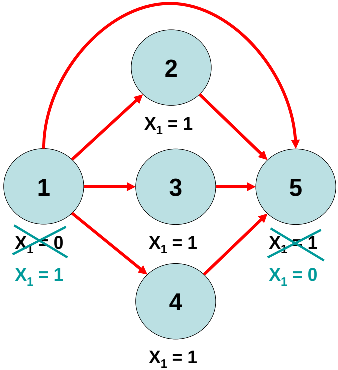


- Less can be more

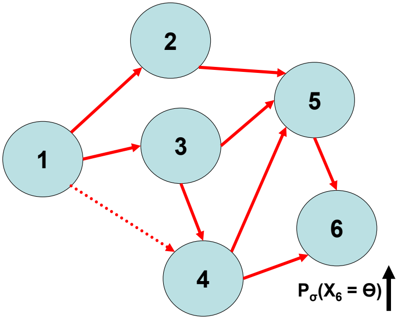


Some Difficulties of Bayesian Learning

- No following the crowds



- Less can be more.



Equilibrium Decision Rule

Lemma

The decision of agent n, $x_n = \sigma(\mathcal{I}_n)$, satisfies

$$x_n = \begin{cases} 1, & \text{if } \mathbb{P}_\sigma(\theta = 1 \mid s_n) + \mathbb{P}_\sigma\big(\theta = 1 \mid B(n), x_k \text{ for all } k \in B(n)\big) > 1, \\ 0, & \text{if } \mathbb{P}_\sigma(\theta = 1 \mid s_n) + \mathbb{P}_\sigma\big(\theta = 1 \mid B(n), x_k \text{ for all } k \in B(n)\big) < 1, \end{cases}$$

and $x_n \in \{0, 1\}$ otherwise.

- Implication: The belief about the state decomposes into two parts:
  - the Private Belief: $\mathbb{P}_\sigma(\theta = 1 \mid s_n)$;
  - the Social Belief: $\mathbb{P}_\sigma(\theta = 1 \mid B(n), x_k \text{ for all } k \in B(n))$.

Private Beliefs

- Assume $\mathbb{F}_0$ and $\mathbb{F}_1$ are mutually absolutely continuous.
- The private belief of agent n is then

$$p_n(s_n) = \mathbb{P}_\sigma(\theta = 1 | s_n) = \left(1 + \frac{d\mathbb{F}_0(s_n)}{d\mathbb{F}_1(s_n)}\right)^{-1}.$$

Definition

The signal structure has unbounded private beliefs if

$$\inf_{s \in S} \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(s) = 0 \quad \textit{and} \quad \sup_{s \in S} \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(s) = \infty.$$

- If the private beliefs are unbounded, then there exist agents with beliefs arbitrarily strong in both directions.
  - Gaussian signals yield unbounded beliefs; discrete signals yield bounded beliefs.

Properties of Network Topology

Definition

A network topology $\{\mathbb{Q}_n\}_{n\in\mathbb{N}}$ has expanding observations if for all K,

$$\lim_{n\to\infty} \mathbb{Q}_n\left(\max_{b\in B(n)} b < K\right) = 0.$$

- Nonexpanding observations equivalent to a group of agents that is excessively influential. This is stronger than being influential.
- More concretely, the first K agents are excessively influential if there exists $\epsilon > 0$ and an infinite subset $\mathcal{N} \in \mathbb{N}$ such that

$$\mathbb{Q}_n\left(\max_{b\in B(n)} b < K\right) \geq \epsilon \quad \text{for all} \quad n \in \mathcal{N}.$$

  - For example, a group is excessively influential if it is the source of all information for an infinitely large component of the network.
- Expanding observations $\Leftrightarrow$ no excessively influential agents.

Learning Theorem – with Unbounded Beliefs

Theorem

Assume that the network topology $\{\mathbb{Q}_n\}_{n\in\mathbb{N}}$ has nonexpanding observations. Then, there exists no equilibrium $\sigma \in \Sigma^*$ with asymptotic learning.

Theorem

Assume unbounded private beliefs and expanding observations. Then, asymptotic learning occurs in every equilibrium $\sigma \in \Sigma^*$.

- Implication: Influential, but not excessively influential, individuals do not prevent learning.
  - This contrasts with results in models of myopic learning.
  - Intuition: The weight given to the information of influential individuals is adjusted in Bayesian updating.

Proof of Theorem – A Roadmap

- Characterization of equilibrium strategies when observing a single agent.
- Strong improvement principle when observing one agent.
- Generalized strong improvement principle.
- Asymptotic learning with unbounded private beliefs and expanding observations.

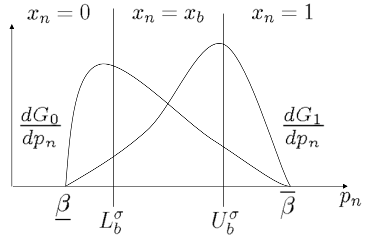
Observing a Single Decision

Proposition

Let $B(n) = \{b\}$ for some agent n. There exists L_b^σ and U_b^σ such that agent n's decision x_n in $\sigma \in \Sigma^*$ satisfies

$$x_n = \begin{cases} 0, & \text{if } p_n < L_b^\sigma; \\ x_b, & \text{if } p_n \in (L_b^\sigma, U_b^\sigma); \\ 1, & \text{if } p_n > U_b^\sigma. \end{cases}$$

- Let $\mathbb{G}_j(r) = \mathbb{P}(p \leq r \mid \theta = j)$ be the conditional distribution of the private belief with $\underline{\beta}$ and $\overline{\beta}$ denoting the lower and upper support



Strong Improvement Principle

- Agent n has the option of copying the action of his neighbor b:

$$\mathbb{P}_\sigma(x_n = \theta \mid B(n) = \{b\}) \geq \mathbb{P}_\sigma(x_b = \theta).$$

- Using the equilibrium decision rule and the properties of private beliefs, we establish a strict gain of agent n over agent b.

Proposition (Strong Improvement Principle)

Let $B(n) = \{b\}$ for some n and $\sigma \in \Sigma^*$ be an equilibrium. There exists a continuous, increasing function $\mathcal{Z} : [1/2, 1] \to [1/2, 1]$ with $\mathcal{Z}(\alpha) \geq \alpha$ such that

$$\mathbb{P}_\sigma(x_n = \theta \mid B(n) = \{b\}) \geq \mathcal{Z}\left(\mathbb{P}_\sigma(x_b = \theta)\right).$$

Moreover, if the private beliefs are unbounded, then:

- $\mathcal{Z}(\alpha) > \alpha$ *for all* $\alpha < 1$.
- *Thus* $\alpha = 1$ *is the unique fixed point of* $\mathcal{Z}(\alpha)$.

Generalized Strong Improvement Principle

- With multiple agents, learning no worse than observing just one of them.
- Equilibrium strategy is better than the following heuristic:
  - Discard all decisions except the one from the most informed neighbor.
  - Use equilibrium decision rule for this new information set.

Proposition (Generalized Strong Improvement Principle)

For any $n \in \mathbb{N}$, any set $\mathfrak{B} \subseteq \{1, ..., n-1\}$ and any $\sigma \in \Sigma^*$,

$$\mathbb{P}_\sigma\left(x_n = \theta \mid B(n) = \mathfrak{B}\right) \geq \mathcal{Z}\left(\max_{b \in \mathfrak{B}} \mathbb{P}_\sigma(x_b = \theta)\right).$$

Moreover, if the private beliefs are unbounded, then:

- $\mathcal{Z}(\alpha) > \alpha$ *for all* $\alpha < 1$.
- *Thus* $\alpha = 1$ *is the unique fixed point of* $\mathcal{Z}(\alpha)$.

Proof of Theorem

- Under expanding observations, one can construct a sequence of agents along which the generalized strong improvement principle applies
- Unbounded private beliefs imply that along this sequence $\mathcal{Z}(\alpha)$ strictly increases
- Until unique fixed point $\alpha = 1$, corresponding to asymptotic learning

No Learning with Bounded Beliefs

Theorem

Assume that the signal structure has bounded private beliefs. Assume that the network topology satisfies one of the following conditions:

(a) $B(n) = \{1, ..., n-1\}$ *for all* n,

(b) $|B(n)| \leq 1$ *for all* n,

(c) *there exists some constant* M *such that* $|B(n)| \leq M$ *for all* n *and*

$$\lim_{n\to\infty} \max_{b\in B(n)} b = \infty \text{ with probability 1,}$$

then asymptotic learning does not occur.

- **Implication:** No learning from observing neighbors or sampling the past.

Proof Idea -Part (c): Learning implies social beliefs converge to 0 or 1 a.s.

- With bounded beliefs, agents decide on the basis of social belief alone. Then, positive probability of mistake–contradiction

Learning with Bounded Beliefs

Theorem

(a) *There exist random network topologies for which learning occurs in all equilibria for any signal structure (bounded or unbounded).*

(b) *There exist signal structures for which learning occurs for a collection of network topologies.*

- Important since it shows the role of stochastic network topologies and also the possibility of many pieces of very limited information to be aggregated.

Learning with Bounded Beliefs (Continued)

Example

Let the network topology be

$$B(n) = \begin{cases} \{1, \ldots, n-1\}, & \text{with probability } 1 - \frac{1}{n}, \\ \emptyset, & \text{with probability } \frac{1}{n}. \end{cases}$$

Asymptotic learning occurs in all equilibria $\sigma \in \Sigma^*$ for any signal structure $(\mathbb{F}_0, \mathbb{F}_1)$.

- Proof Idea:
  - The rate of contrary actions in the long run gives away the state.

Heterogeneity and Learning

- So far, all agents have the same preferences.
  - They all prefer to take action $= \theta$, and with the same intensity.
- In realistic situations, not only diversity of opinions, but also diversity of preferences.
- How does diversity of preferences/priors affect social learning?
- Naive conjecture: diversity will introduce additional noise and make learning harder or impossible.
- Our Result: in the line topology, diversity always facilitates learning.

Model with Heterogeneous Preferences

- Assume $B(n) = \{1, ..., n-1\}$.
- Let agent n have private preference t_n independently drawn from some $\mathbb{H}$.
- The payoff of agent n given by:

$$u_n(x_n, t_n, \theta) = \begin{cases} I(\theta = 1) + 1 - t_n & \text{if } x_n = 1 \\ I(\theta = 0) + t_n & \text{if } x_n = 0 \end{cases}$$

- Assumption: $\mathbb{H}$ has full support on $(\underline{\gamma}, \overline{\gamma})$, $\mathbb{G}_1$, $\mathbb{G}_0$ have full support in $(\underline{\beta}, \overline{\beta})$.
- As before, private beliefs are unbounded if $\underline{\beta} = 0$ and $\overline{\beta} = 1$ and bounded if $\underline{\beta} > 0$ and $\overline{\beta} < 1$.
- Heterogeneity is unbounded if $\underline{\gamma} = 0$ and $\overline{\gamma} = 1$ and bounded if $\underline{\gamma} > 0$ and $\overline{\gamma} < 1$.

Main Results

Theorem

With unbounded heterogeneity, i.e., $[0,1] \subseteq supp(\mathbb{H})$*, asymptotic learning occurs in all equilibria* $\sigma \in \Sigma^*$ *for any signal structure* $(\mathbb{F}_0, \mathbb{F}_1)$*.*

- Greater heterogeneity under $\mathbb{H}_1$ than under $\mathbb{H}_2$ if $\underline{\gamma}_1 < \underline{\gamma}_2$ and $\overline{\gamma}_1 > \overline{\gamma}_2$

Theorem

With bounded heterogeneity (i.e., $[0,1] \nsubseteq supp(\mathbb{H})$*) and bounded private beliefs, there is no learning, but greater heterogeneity leads to "greater social learning".*

- Heterogeneity pulls learning in opposite directions:
  - Actions of others are less informative (direct effect)
  - Each agent uses more of his own signal in making decisions and, therefore, there is more information in the history of past actions (indirect effect).
- Indirect effect dominates the direct effect!

Some Observations

- Preferences immediately imply that each agent will use a threshold rule as a function of this type t_n.

$$x_n = \begin{cases} 1, & \text{if } \mathbb{P}_\sigma(\theta = 1 | I_n) > t_n; \\ 0, & \text{if } \mathbb{P}_\sigma(\theta = 1 | I_n) < t_n. \end{cases}$$

- Similar arguments lead to a characterization in terms of private and social beliefs.
- Private belief: $p_n = \mathbb{P}(\theta = 1 | s_n)$
- Social belief: $q_n = \mathbb{P}(\theta = 1 | x_1, ..., x_{n-1})$.

Preliminary Lemmas

Lemma

In equilibrium, agent n chooses action $x_n = 0$ if and and if

$$p_n \leq \frac{t_n(1-q_n)}{t_n(1-2q_n)+q_n}.$$

- This follows by manipulating the threshold decision rule.

Lemma

The social belief q_n converges with probability 1.

- This follows from a famous result in stochastic processes, Martingale Convergence Theorem (together with the observation that q_n is a martingale).
- Let the limiting belief (random variable) be $\hat{q}$.

Key Lemmas

Lemma

The limiting social belief $\hat{q}$ satisfies

$$\hat{q} \notin \left(\left[1 + \left(\frac{\overline{\beta}}{1-\overline{\beta}} \right) \left(\frac{1-\underline{\gamma}}{\underline{\gamma}} \right) \right]^{-1}, \left[1 + \left(\frac{\underline{\beta}}{1-\underline{\beta}} \right) \left(\frac{1-\overline{\gamma}}{\overline{\gamma}} \right) \right]^{-1} \right)$$

with probability 1.

Lemma

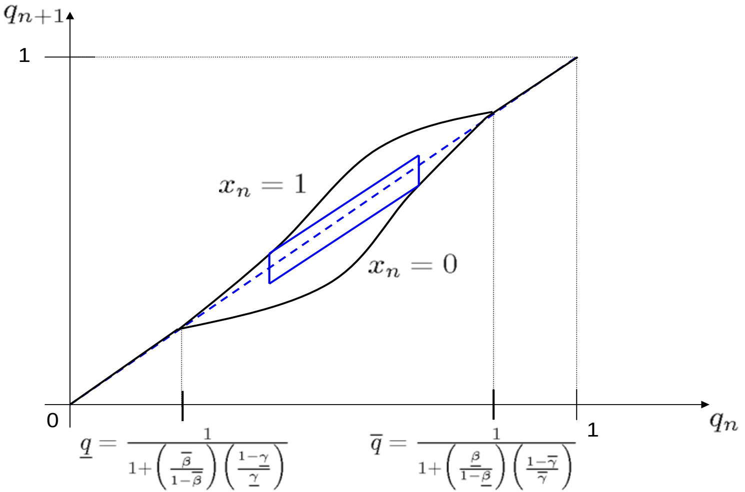
The limiting social belief $\hat{q}$ satisfies

$$\hat{q} \notin \left[0, \left[1 + \frac{1-\underline{\beta}}{\underline{\beta}} \frac{\overline{\beta}}{1-\overline{\beta}} \frac{1-\underline{\gamma}}{\underline{\gamma}} \right]^{-1} \right) \bigcup \left(\left[1 + \frac{1-\overline{\beta}}{\overline{\beta}} \frac{\underline{\beta}}{1-\underline{\beta}} \frac{1-\overline{\gamma}}{\overline{\gamma}} \right]^{-1}, 1 \right]$$

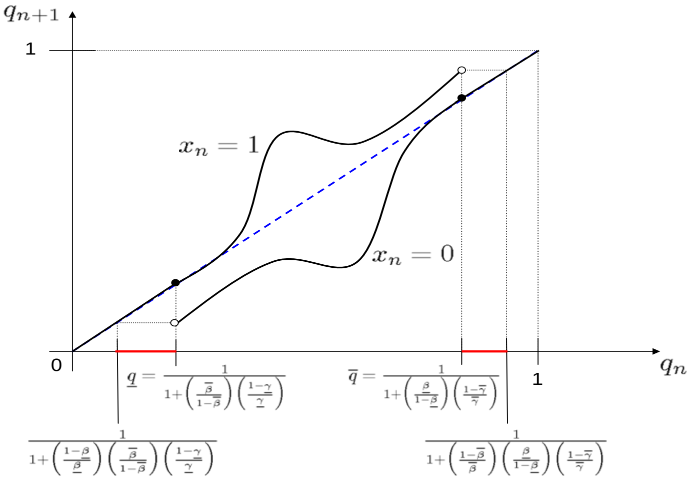
with probability 1.

- This characterization is "tight" in the sense that simple examples reach any of the points not ruled out by these lemmas.

Sketch of the Proof of the Lemmas



Sketch of the Proof of the Lemmas (continued)



Main Results As Corollaries

- Setting $\underline{\beta} = 0$ and $\overline{\beta} = 1$, and we conclude that $\hat{q}$ must converge almost surely either to 0 or 1.
- Since $q_n/(1 - q_n)$ conditional on $\theta = 0$ and $(1 - q_n)/q_n$ conditional on $\theta = 1$ are also martingales and converge to random variables with finite expectations, when $\theta = 0$, we cannot almost surely converge to 1 and vice versa.
- Therefore, there is asymptotic learning with unbounded private beliefs (as before).
- Similarly, setting $\underline{\gamma} = 0$ and $\overline{\gamma} = 1$, we obtain the first theorem—with unbounded heterogeneity, there is always asymptotic learning regardless of whether privates beliefs are unbounded.
- In this case, asymptotic learning with unbounded private beliefs and homogeneous preferences has several "unattractive features"—large jumps in beliefs.
- Learning with unbounded heterogeneous preferences takes a much more "plausible" form—smooth convergence to the correct opinion.

Main Results As Corollaries (continued)

- Finally, when $\underline{\beta} > 0$, $\overline{\beta} < 1$, $\underline{\gamma} > 0$ and $\overline{\gamma} < 1$, then no social learning.
- But in this case, the region of convergence shifts out as heterogeneity increases: Why does this correspond to more social learning?
- Because it can be shown that the ex-ante probability of making the right choice
$$\frac{1}{2}\mathbb{P}\left[\underline{q}|\theta = 0\right] + \frac{1}{2}\mathbb{P}\left[\overline{q}|\theta = 1\right],$$
is decreasing in $\underline{\gamma}$ and increasing $\overline{\gamma}$—greater social learning.

A Model of Bayesian Communication Learning

- Effect of communication on learning: Acemoglu, Bimpikis, Ozdaglar (2009)
- Two possible states of the world, $\theta \in \{0, 1\}$
- A set $\mathcal{N} = \{1, \ldots, n\}$ of agents and a friendship network given

Stage 1: Network Formation

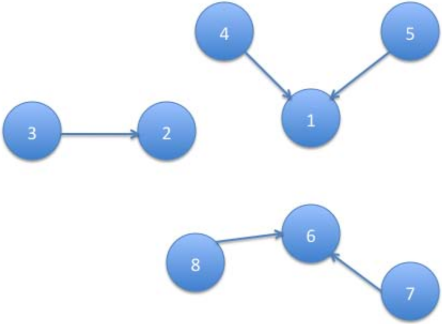
- Additional link formation is costly, c_{ij}^n : cost incurred by i to link with j
- Induces the communication network $G^n = (\mathcal{N}, \mathcal{E}^n)$

Stage 2: Information Exchange (over the communication network G^n)

- Each agent receives an iid private signal, $s_i \sim \mathbb{F}_\theta$
- Agents receive all information acquired by their direct neighbors
- At each time period t they can choose:
  (1) irreversible action 0 (2) irreversible action 1 (3) wait

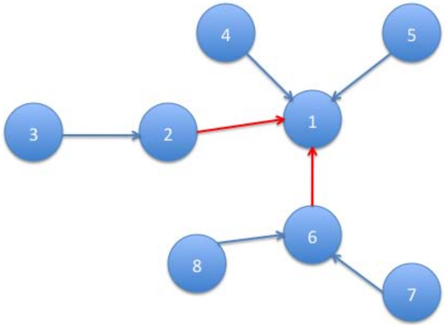
Stage 1: Forming the communication network

Friendship network

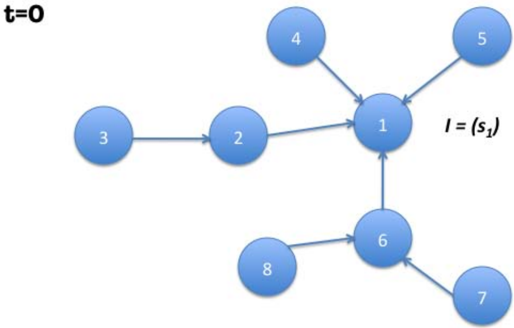


Stage 1: Forming the communication network

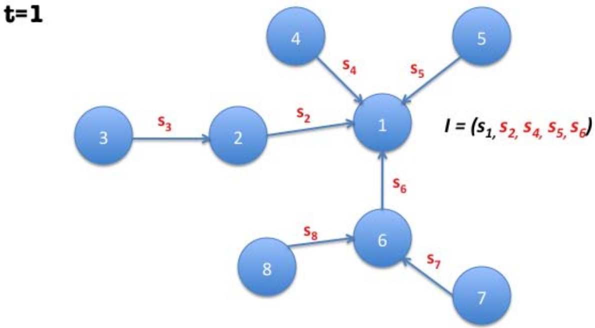
Friendship network + Additional Links=Communication network



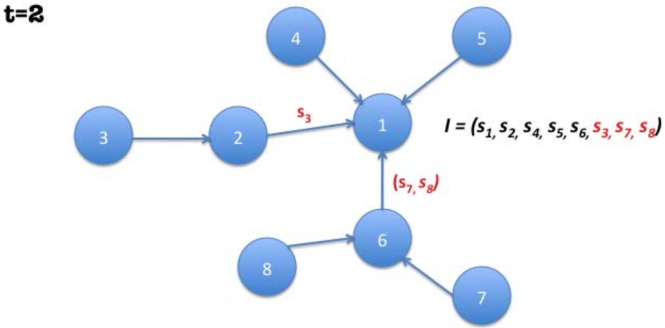
Stage 2: Information Exchange



Stage 2: Information Exchange



Stage 2: Information Exchange



Model

- In this lecture: Focus on stage 2
- Agent i's payoff is given by

$$u_i(\mathbf{x_i^n}, \theta) = \begin{cases} \delta^{\tau}\pi & \text{if } x^n_{i,\tau} = \theta \text{ and } x^n_{i,t} = \text{ ``wait'' for } t < \tau \\ 0 & \text{otherwise} \end{cases}$$

  - $\mathbf{x_i^n} = [x^n_{i,t}]_{t \geq 0}$: sequence of agent i's decisions, $x^n_{i,t} \in \{0, 1, \text{``}wait\text{''}\}$
  - δ: discount factor ($\delta < 1$)
  - τ: time when action is taken (agent collects information up to τ)
  - π: payoff - normalized to 1
- Preliminary Assumptions (relax both later):
  - Information continues to be transmitted after exit.
  - Communication between agents is not strategic
- Let
  $B^n_{i,t} = \{j \neq i \mid \exists \text{ a directed path from } j \text{ to } i \text{ with at most } t \text{ links in } G^n\}$
  - All agents that are at most t links away from i in G^n
- Agent i's information set at time t: $I^n_{i,t} = \{s_i, s_j \text{ for all } j \in B^n_{i,t}\}$.

Equilibrium and Learning

- Given a sequence of communication networks $\{G^n\}$ (society):
  - Strategy for agent i at time t is $\sigma^n_{i,t} : \mathcal{I}^n_{i,t} \to \{\text{“wait”}, 0, 1\}$

Definition

A strategy profile $\sigma^{n,*}$ is a Perfect-Bayesian Equilibrium if for all i and t,

$$\sigma^{n,*}_{i,t} \in \arg\max_{y \in \{\text{“wait”},0,1\}} \mathbb{E}_{(y,\sigma^{n,*}_{-i,t})}\left(u_i(\mathbf{x_i^n}, \theta) | I^n_{i,t}\right).$$

- Let

$$M^n_{i,t} = \begin{cases} 1 & \text{if } x_{i,\tau} = \theta \text{ for some } \tau \leq t \\ 0 & \text{otherwise} \end{cases}$$

Definition

We say that asymptotic learning occurs in society $\{G^n\}$ if for every $\epsilon > 0$

$$\lim_{n\to\infty} \lim_{t\to\infty} \mathbb{P}_{\sigma^{n,*}}\left(\left[\frac{1}{n}\sum_{i=1}^{n}\left(1 - M^n_{i,t}\right)\right] > \epsilon\right) = 0$$

Agent Decision Rule

Lemma

Let $\sigma^{n,*}$ be an equilibrium and $I^n_{i,t}$ be an information set of agent i at time t. Then, the decision of agent i, $x^n_{i,t} = \sigma^{n,*}_{i,t}(I^n_{i,t})$ satisfies

$$x^n_{i,t} = \begin{cases} 0, & \textit{if } \log L(s_i) + \sum_{j \in B^n_{i,t}} \log L(s_j) \leq -\log A^{n,*}_{i,t}, \\ 1, & \textit{if } \log L(s_i) + \sum_{j \in B^n_{i,t}} \log L(s_j) \geq \log A^{n,*}_{i,t}, \\ \text{“wait”}, & \textit{otherwise}, \end{cases}$$

where $L(s_i) = \frac{dP_\sigma(s_i \mid \theta=1)}{dP_\sigma(s_i \mid \theta=0)}$ is the likelihood ratio of signal s_i, and $A^{n,*}_{i,t} = \frac{p^{n,*}_{i,t}}{1-p^{n,*}_{i,t}}$, is a time-dependent parameter.

- $p^{n,*}_{i,t}$: belief threshold that depends on time and graph structure
- For today:
  - Focus on binary private signals $s_i \in \{0,1\}$
  - Assume $L(1) = \frac{\beta}{1-\beta}$ and $L(0) = \frac{1-\beta}{\beta}$ for some $\beta > 1/2$.

Minimum Observation Radius

Lemma

The decision of agent i, $x_{i,t}^n = \sigma_{i,t}^{n,*}(I_{i,t}^n)$ satisfies

$$x_{i,t}^n(I_{i,t}^n) = \begin{cases} 0, & \text{if } k_{i,0}^t - k_{i,1}^t \geq \log A_{i,t}^{n,*} \cdot \left(\log \frac{\beta}{1-\beta}\right)^{-1}, \\ 1, & \text{if } k_{i,1}^t - k_{i,0}^t \geq \log A_{i,t}^{n,*} \cdot \left(\log \frac{\beta}{1-\beta}\right)^{-1}, \\ \text{“wait”}, & \text{otherwise}, \end{cases}$$

where $k_{i,1}^t$ ($k_{i,0}^t$) denotes the number of 1's (0's) agent i observed up to time t.

Definition

We define the minimum observation radius of agent i, denoted by d_i^n, as

$$d_i^n = \arg\min_t \left\{ \left|B_{i,t}^n\right| \;\middle|\; \left|B_{i,t}^n\right| \geq \log A_{i,t}^{n,*} \cdot \left(\log \frac{\beta}{1-\beta}\right)^{-1} \right\}.$$

- Agent i receives at least $|B_{i,d_i^n}^n|$ signals before she takes an irreversible action
- $B_{i,d_i^n}^n$: Minimum observation neighborhood of agent i

A Learning Theorem

Definition

For any integer $k > 0$, we define the k-radius set, denoted by V_k^n, as

$$V_k^n = \{j \in \mathcal{N} \mid |B^n_{j,d_j^n}| \leq k\}$$

- Set of agents with "finite minimum observation neighborhood"
- Note that any agent i in the k-radius (for k finite) set has positive probability of taking the wrong action.

Theorem

Asymptotic learning occurs in society $\{G^n\}$ if and only if

$$\lim_{k\to\infty} \lim_{n\to\infty} \frac{|V_k^n|}{n} = 0.$$

- A "large" number of agents with finite obs. neigh. precludes learning.

Interpreting the Learning Condition

Definition

Agent i is called an (information) maven of society $\{G^n\}_{n=1}^{\infty}$ if i has an infinite in-degree. Let $\mathcal{MAVEN}(\{G^n\}_{n=1}^{\infty})$ denote the set of mavens of society $\{G^n\}_{n=1}^{\infty}$.

- For any agent j, let $d_j^{\mathcal{MAVEN},n}$ the shortest distance defined in communication network G^n between j and a maven $k \in \mathcal{MAVEN}(\{G^n\}_{n=1}^{\infty})$.
- Let W^n be the set of agents at distance at most equal to their minimum observation radius from a maven in G^n, i.e., $W^n = \{j \mid \; d_j^{\mathcal{MAVEN},n} \leq d_j^n\}$.

Corollary

Asymptotic learning occurs in society $\{G^n\}_{n=1}^{\infty}$ if $\lim_{n\to\infty} \frac{1}{n} \cdot |W^n| = 1$.

- "Mavens" as information hubs; most agents must be close to a hub.

Interpreting the Learning Condition (Continued)

Definition

Agent i is a social connector of society $\{G^n\}_{n=1}^{\infty}$ if i has an infinite out-degree.

Corollary

Consider society $\{G^n\}_{n=1}^{\infty}$ such that the sequence of in- and out-degrees is non-decreasing for every agent (as n increases), and

$$\lim_{n\to\infty} \frac{\left|\mathcal{MAVEN}(\{G^n\}_{n=1}^{\infty})\right|}{n} = 0.$$

Then, asymptotic learning occurs if the society contains a social connector within a short distance to a maven, i.e.,

$$d_i^{\mathcal{MAVEN},n} \leq d_i^n, \ \textit{for some social connector } i.$$

- Unless a non-negligible fraction of the agents belongs to the set of mavens and the rest can obtain information directly from a maven, information aggregated at the mavens spreads through the out-links of a connector.

Relaxing the Information Flow Assumption

Theorem

Asymptotic learning occurs in society $\{G^n\}$ even when information flows are interrupted after exit if

$$\lim_{k\to\infty} \lim_{n\to\infty} \frac{\left|V_k^n\right|}{n} = 0.$$

- Intuition: When there is asymptotic learning, no interruption of information flow for a non-negligible fraction of agents.
- The corollaries apply as above.

Relaxing the Nonstrategic Communication Assumption

Theorem

Asymptotic learning in society $\{G^n\}$ is an ϵ-equilibrium if

$$\lim_{k\to\infty}\lim_{n\to\infty}\frac{|V_k^n|}{n}=0.$$

- Intuition: Misrepresenting information to a hub (maven) not beneficial, and thus at most a small benefit for most agents from misrepresenting their information.
- Therefore, if there is asymptotic learning without strategic communication, then there exists an equilibrium with strategic communication in which agents taking the right action without strategic communication have no more than ϵ to gain by misrepresenting, and thus there exists an ϵ-equilibrium with asymptotic learning.

Learning in Random Graph Models

- Focus on networks with bidirectional communication (corresponding to undirected graphs).
- Recall that asymptotic learning occurs if and only if for all but a negligible fraction of agents, the shortest path to a hub/maven is shorter than minimum observation radius.
- Then the following proposition is intuitive:

Proposition

Asymptotic Learning fails for

(a) *Bounded Degree Graphs, e.g., expanders.*

(b) *Preferential Attachment Graphs (with high probability).*

- *Intuition: Edges form with probability proportional to degree, but there exist many low degree nodes.*

Learning in Random Graph Models

Proposition

Asymptotic Learning occurs for

(a) *Complete and Star Graphs.*

(b) *Power Law Graphs with exponent $\gamma \leq 2$ (with high probability).*

- *Intuition: The average degree is infinite - there exist many hubs.*

(c) *Hierarchical Graphs.*

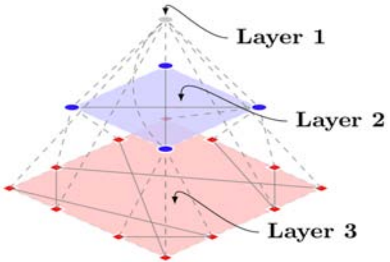


Figure: Hierarchical Society.

14.15J / 6.207J Networks
Fall 2009