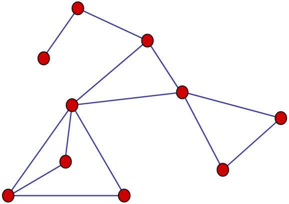
Non-Bayesian Social Learning

Presented by Arastoo Fazeli

November 30, 2009

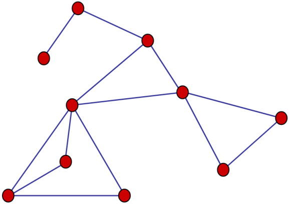
Learning in Complex Networks: Model and Abstractions

- Each vertex represents an agent
- Each edge represents information flow between two agents
- Agents have access to their neighbors' information



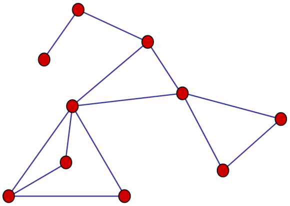
Learning in Complex Networks: Model and Abstractions

- Each vertex represents an agent
- Each edge represents information flow between two agents
- Agents have access to their neighbors' information



Θ	parameter space
$\theta^* \in \Theta$	the unobservable true state of the world
$s_t = (s_t^1, \ldots, s_t^n)$	random signals observed by the agents

Bayesian Learning over Networks



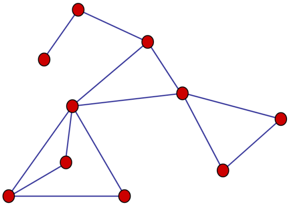
$$\mu_{i,t}(\theta) = \mathbb{P}\left[\theta = \theta^* | \mathcal{F}_{i,t}\right]$$

where

$$\mathcal{F}_{i,t} = \sigma\left(s_1^i, \ldots, s_t^i, \{\mu_{j,k} : j \in \mathcal{N}_i, k \leq t\}\right)$$

is the information available to agent i up to time t.

Bayesian Learning over Networks



$$\mu_{i,t}(\theta) = \mathbb{P}\left[\theta = \theta^* | \mathcal{F}_{i,t}\right]$$

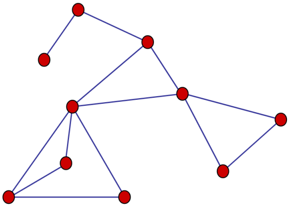
where

$$\mathcal{F}_{i,t} = \sigma\left(s_1^i, \ldots, s_t^i, \{\mu_{j,k} : j \in \mathcal{N}_i, k \leq t\}\right)$$

is the information available to agent i up to time t.

Agents need to make rational deductions about everybody's beliefs based on only observing neighbors' beliefs:

Bayesian Learning over Networks



$$\mu_{i,t}(\theta) = \mathbb{P}\left[\theta = \theta^* | \mathcal{F}_{i,t}\right]$$

where

$$\mathcal{F}_{i,t} = \sigma\left(s_1^i, \ldots, s_t^i, \{\mu_{j,k} : j \in \mathcal{N}_i, k \leq t\}\right)$$

is the information available to agent i up to time t.

Agents need to make rational deductions about everybody's beliefs based on only observing neighbors' beliefs:

Computationally hard!

The Problem with Bayesian Learning

1. Incomplete network information

The Problem with Bayesian Learning

1. Incomplete network information
2. Incomplete information about other agents' signal structures

The Problem with Bayesian Learning

1. Incomplete network information
2. Incomplete information about other agents' signal structures
3. Higher order beliefs matter

The Problem with Bayesian Learning

1. Incomplete network information
2. Incomplete information about other agents' signal structures
3. Higher order beliefs matter
4. The source of each piece of information is not immediately clear

$\boxed{\text{Intractable and not local.}}$

Non-Bayesian Social Learning

Need a local and computationally tractable update, which hopefully delivers asymptotic social learning.

Non-Bayesian Social Learning

Need a local and computationally tractable update, which hopefully delivers asymptotic social learning.

Agent i is

- Bayesian when it comes to her observation
- non-Bayesian when incorporating others information

Model

$\mathcal{N} = \{1, 2, \ldots, n\}$ individuals in the society

Model

$\mathcal{N} = \{1, 2, \ldots, n\}$	individuals in the society
$G = (\mathcal{N}, \mathcal{E})$	social network

Model

$\mathcal{N} = \{1, 2, \ldots, n\}$	individuals in the society
$G = (\mathcal{N}, \mathcal{E})$	social network
Θ	finite parameter space

Model

$\mathcal{N} = \{1, 2, \dots, n\}$	individuals in the society
$G = (\mathcal{N}, \mathcal{E})$	social network
Θ	finite parameter space
$\theta^* \in \Theta$	the unobservable true state of the world

Model

$\mathcal{N} = \{1, 2, \ldots, n\}$	individuals in the society
$G = (\mathcal{N}, \mathcal{E})$	social network
Θ	finite parameter space
$\theta^* \in \Theta$	the unobservable true state of the world
$s_t = (s_t^1, \ldots, s_t^n)$	s_t^i is the signal observed by agent i at time t

Model

$\mathcal{N} = \{1, 2, \ldots, n\}$	individuals in the society
$G = (\mathcal{N}, \mathcal{E})$	social network
Θ	finite parameter space
$\theta^* \in \Theta$	the unobservable true state of the world
$s_t = (s_t^1, \ldots, s_t^n)$	s_t^i is the signal observed by agent i at time t
$S = S_1 \times S_2 \times \cdots \times S_n$	signal space

Model

$\mathcal{N} = \{1, 2, \ldots, n\}$	individuals in the society
$G = (\mathcal{N}, \mathcal{E})$	social network
Θ	finite parameter space
$\theta^* \in \Theta$	the unobservable true state of the world
$s_t = (s_t^1, \ldots, s_t^n)$	s_t^i is the signal observed by agent i at time t
$S = S_1 \times S_2 \times \cdots \times S_n$	signal space
$\ell(s\|\theta)$	global signal structure

Model

$\mathcal{N} = \{1, 2, \ldots, n\}$	individuals in the society
$G = (\mathcal{N}, \mathcal{E})$	social network
Θ	finite parameter space
$\theta^* \in \Theta$	the unobservable true state of the world
$s_t = (s_t^1, \ldots, s_t^n)$	s_t^i is the signal observed by agent i at time t
$S = S_1 \times S_2 \times \cdots \times S_n$	signal space
$\ell(s\|\theta)$	global signal structure
$\ell_i(s^i\|\theta)$	agent i's signal structure

Model

$\mu_{i,t}(\theta)$	time t beliefs of agent i (a probability measure on Θ)
$\mu_{i,0}(\theta)$	agent i's prior belief

Model

$\mu_{i,t}(\theta)$	time t beliefs of agent i (a probability measure on Θ)
$\mu_{i,0}(\theta)$	agent i's prior belief
$\mathbb{P}^* = \otimes_{t=1}^{\infty} \ell(\cdot \vert \theta^*)$	the true probability measure

Model

$\mu_{i,t}(\theta)$	time t beliefs of agent i (a probability measure on Θ)
$\mu_{i,0}(\theta)$	agent i's prior belief
$\mathbb{P}^* = \otimes_{t=1}^{\infty} \ell(\cdot\|\theta^*)$	the true probability measure

Agent i's time t forecasts of the next observation:

$$m_{i,t}(s_{t+1}^i) = \int_\Theta \ell_i(s_{t+1}^i|\theta) d\mu_{i,t}(\theta)$$

What Do We Mean by Learning?

Definition

The Forecasts of agent i are eventually correct on a path $\{s_t\}_{t=1}^{\infty}$ if, along that path,

$$m_{i,t}(\cdot) \to \ell_i(\cdot|\theta^*) \quad \text{as} \quad t \to \infty.$$

What Do We Mean by Learning?

Definition

The Forecasts of agent i are eventually correct on a path $\{s_t\}_{t=1}^{\infty}$ if, along that path,

$$m_{i,t}(\cdot) \to \ell_i(\cdot|\theta^*) \quad \text{as} \quad t \to \infty.$$

Definition

Agent i asymptotically learns the true parameter θ^* on a path $\{s_t\}_{t=1}^{\infty}$ if, along that path,

$$\mu_{i,t}(\theta^*) \to 1 \quad \text{as} \quad t \to \infty.$$

What Do We Mean by Learning?

Definition

The Forecasts of agent i are eventually correct on a path $\{s_t\}_{t=1}^{\infty}$ if, along that path,

$$m_{i,t}(\cdot) \to \ell_i(\cdot|\theta^*) \quad \text{as} \quad t \to \infty.$$

Definition

Agent i asymptotically learns the true parameter θ^* on a path $\{s_t\}_{t=1}^{\infty}$ if, along that path,

$$\mu_{i,t}(\theta^*) \to 1 \quad \text{as} \quad t \to \infty.$$

- Asymptotic learning, in this setup, is stronger.

Model: Belief Update

$$\mu_{i,t+1}(\theta) = a_{ii}\,\mathrm{BU}(\mu_{i,t}; s^i_{t+1})(\theta) + \sum_{j\in\mathcal{N}_i} a_{ij}\mu_{j,t}(\theta)$$

where

$$\mathrm{BU}(\mu_{i,t}; s^i_{t+1})(\theta) = \mu_{i,t}(\theta)\frac{\ell_i(s^i_{t+1}|\theta)}{m_{i,t}(s^i_{t+1})}$$

$$a_{ij} \geq 0 \quad , \quad \sum_{j\in\mathcal{N}_i} a_{ij} = 1$$

- Individuals rationally update their beliefs after observing the signal
- exhibit a bias towards the average belief in the neighborhood

Model: Belief Update

$$\mu_{i,t+1}(\theta) = a_{ii}\mu_{i,t}(\theta)\frac{\ell_i(s^i_{t+1}|\theta)}{m_{i,t}(s^i_{t+1})} + \sum_{i\neq j} a_{ij}\mu_{j,t}(\theta) \qquad \forall\theta\in\Theta$$

Model: Belief Update

$$\mu_{i,t+1}(\theta) = a_{ii}\mu_{i,t}(\theta)\frac{\ell_i(s^i_{t+1}|\theta)}{m_{i,t}(s^i_{t+1})} + \sum_{i \neq j} a_{ij}\mu_{j,t}(\theta) \qquad \forall \theta \in \Theta$$

- Does not require knowledge about the network.

Model: Belief Update

$$\mu_{i,t+1}(\theta) = a_{ii}\mu_{i,t}(\theta)\frac{\ell_i(s^i_{t+1}|\theta)}{m_{i,t}(s^i_{t+1})} + \sum_{i \neq j} a_{ij}\mu_{j,t}(\theta) \qquad \forall\theta \in \Theta$$

- Does not require knowledge about the network.
- Does not require deduction about the beliefs of others.

Model: Belief Update

$$\mu_{i,t+1}(\theta) = a_{ii}\mu_{i,t}(\theta)\frac{\ell_i(s^i_{t+1}|\theta)}{m_{i,t}(s^i_{t+1})} + \sum_{i\neq j} a_{ij}\mu_{j,t}(\theta) \qquad \forall\theta\in\Theta$$

- Does not require knowledge about the network.
- Does not require deduction about the beliefs of others.
- Does not require knowledge about other agents' signal structures.

Model: Belief Update

$$\mu_{i,t+1}(\theta) = a_{ii}\mu_{i,t}(\theta)\frac{\ell_i(s^i_{t+1}|\theta)}{m_{i,t}(s^i_{t+1})} + \sum_{i \neq j} a_{ij}\mu_{j,t}(\theta) \qquad \forall \theta \in \Theta$$

- Does not require knowledge about the network.
- Does not require deduction about the beliefs of others.
- Does not require knowledge about other agents' signal structures.
- The update is local and tractable.

Model: Belief Update

$$\mu_{i,t+1}(\theta) = a_{ii}\mu_{i,t}(\theta)\frac{\ell_i(s^i_{t+1}|\theta)}{m_{i,t}(s^i_{t+1})} + \sum_{i \neq j} a_{ij}\mu_{j,t}(\theta) \qquad \forall \theta \in \Theta$$

- Does not require knowledge about the network.
- Does not require deduction about the beliefs of others.
- Does not require knowledge about other agents' signal structures.
- The update is local and tractable.
- If signals are uninformative, reduces to the model of DeGroot(1974).

Model: Belief Update

$$\mu_{i,t+1}(\theta) = a_{ii}\mu_{i,t}(\theta)\frac{\ell_i(s^i_{t+1}|\theta)}{m_{i,t}(s^i_{t+1})} + \sum_{i \neq j} a_{ij}\mu_{j,t}(\theta) \qquad \forall \theta \in \Theta$$

- Does not require knowledge about the network.
- Does not require deduction about the beliefs of others.
- Does not require knowledge about other agents' signal structures.
- The update is local and tractable.
- If signals are uninformative, reduces to the model of DeGroot(1974).
- Reduces to the benchmark Bayesian case if agents assign weight zero to the beliefs of their neighbors.

First Result: Correct Forecasts

$$\mu_{i,t+1}(\theta) = a_{ii}\mu_{i,t}(\theta)\frac{\ell_i(s^i_{t+1}|\theta)}{m_{i,t}(s^i_{t+1})} + \sum_{i\neq j} a_{ij}\mu_{j,t}(\theta) \qquad \forall\theta\in\Theta$$

Proposition

Suppose that

(i) social network is strongly connected,

(ii) all agents have strictly positive self-confidence,

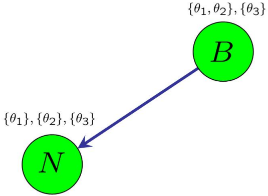
(iii) there exists an agent with strictly positive prior belief on θ^*.

Then, all agents eventually forecast their private observations accurately with $\mathbb{P}^*$-probability one.

Why Strong Connectivity?

What if the network has a directed spanning tree but is not strongly connected?

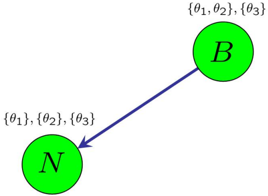
- $\mathcal{N} = \{B, N\}$
- $\Theta = \{\theta_1, \theta_2, \theta_3\}$
- $\theta^* = \theta_2$



Why Strong Connectivity?

What if the network has a directed spanning tree but is not strongly connected?

- $\mathcal{N} = \{B, N\}$
- $\Theta = \{\theta_1, \theta_2, \theta_3\}$
- $\theta^* = \theta_2$



$$\mu_{N,t+1}(\theta) = \lambda \mu_{N,t}(\theta) \frac{\ell_N(s_{t+1}^N | \theta)}{m_{N,t}(s_{t+1}^N)} + (1-\lambda)\mu_{B,t}(\theta) \qquad \forall \theta \in \Theta$$

N is mislead by listening to the less informed agent B.

Convergence of Beliefs & Agreement

Proposition

Under the assumptions of previous proposition, the beliefs of all agents converge with $\mathbb{P}^*$-probability one.

Corollary

Under the assumptions of the proposition, all agents have asymptotically equal beliefs $\mathbb{P}^*$-almost surely.

Convergence of Beliefs & Agreement

Proposition

Under the assumptions of previous proposition, the beliefs of all agents converge with $\mathbb{P}^*$-probability one.

Corollary

Under the assumptions of the proposition, all agents have asymptotically equal beliefs $\mathbb{P}^*$-almost surely.

$\boxed{\text{Consensus!}}$

Social Learning: Information Aggregation

Theorem

Suppose that

(i) social network is strongly connected,

(ii) all agents have strictly positive self-confidence,

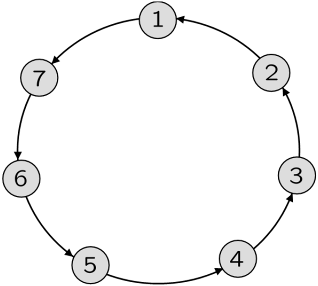
(iii) there exists an agent with strictly positive prior on θ^*.

(iv) for any agent i there exists a signal $\hat{s}^i \in S_i$ such that $\frac{l_i(\hat{s}^i|\theta)}{l_i(\hat{s}^i|\theta^*)} < 1$ $\forall \theta \notin \bar{\Theta}_i$ where $\bar{\Theta}_i = \{\theta \in \Theta : l_i(s^i|\theta) = l_i(s^i|\theta^*), \ \ \forall s^i \in S^i\}$

(v) there is no state $\theta \neq \theta^*$ that is observationally equivalent to θ^* from the point of view of all agents in the network, i.e., $\bar{\Theta}_1 \cap \ldots \cap \bar{\Theta}_n = \{\theta^*\}$

Then, *all agents in the social network learn the true state of the world* P^* *almost surely; that is,* $\mu_i(\theta^*) \to 1$ *with* P^* *probability 1* $\forall i \in \mathcal{N}$

Information Aggregation: An Example

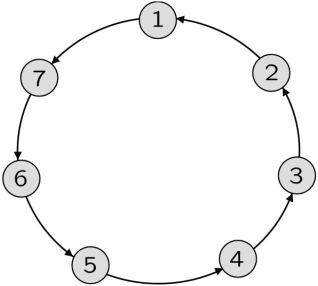


$$\Theta = \{\theta^*, \theta_1, \theta_2, \ldots, \theta_7\}$$

$$S_i = \{H, T\}$$

$$\ell_i(H|\theta) = \begin{cases} \frac{i}{i+1} & \text{if } \theta = \theta_i \\ \frac{1}{(i+1)^2} & \text{otherwise} \end{cases}$$

Information Aggregation: An Example



$$\Theta = \{\theta^*, \theta_1, \theta_2, \ldots, \theta_7\}$$

$$S_i = \{H, T\}$$

$$\ell_i(H|\theta) = \begin{cases} \frac{i}{i+1} & \text{if } \theta = \theta_i \\ \frac{1}{(i+1)^2} & \text{otherwise} \end{cases}$$

Agents learn as if they had access to all information and updated their beliefs rationally.

Summary and Potential Future Directions

A non-Bayesian social learning model:

- Local and tractable
- No information about network topology or signal structures required
- Can handle repeated interactions and information flow over time

Remaining questions:

- The effect of network topology on the learning
- What if actions are observable, and not beliefs?

Summary and Potential Future Directions

A non-Bayesian social learning model:

- Local and tractable
- No information about network topology or signal structures required
- Can handle repeated interactions and information flow over time

Remaining questions:

- The effect of network topology on the learning
- What if actions are observable, and not beliefs?