Network Structure and Naive Sequential Learning

Krishna Dasaratha
Kevin He

June 15, 2018

Sequential Learning on Networks

Environment: sequence of agents guessing a state based on both private info and guesses of predecessors

Examples:

- consumers choosing between rival products
- doctors choosing a treatment
- individuals judging accuracy of a rumor

Observation network: typically each agent only sees guesses of a subset of predecessors

Question: how does the structure of the observation network affect the probability of correct social learning in the long run?

Model and Contributions

Inferential Naiveté: people do not fully take into account that predecessors' actions reflect a combination of their private info and inference they drew from observing still others

Our model: a sequence of naive agents on a social network

- each believes predecessors' actions only reflect private info

Results and contributions:

1. Compare networks in terms of prob of long-run correct learning
   - literature: binary classification of whether learning is perfect w/p 1
   - zoom in on networks with imperfect learning, the leading case

2. Explain observations at odds with other social learning models (e.g. rational, DeGroot, ...)
   - theory: denser networks lead to more mislearning
   - experiment: subjects' accuracy gain from social learning twice as high on sparser networks than denser networks
   - persistent disagreement can happen under partial segregation

Literature

Effect of network structure on learning

- DeGroot model:
  - DeMarzo, Vayanos, and Zwiebel (2003)
  - Golub and Jackson (2010)
  - Golub and Jackson (2012)
- behavioral microfoundations for naive learning:
  - Molavi, Tahbaz-Salehi, and Jadbabaie (2018)
  - Mueller-Frank and Neri (2017WP)
  - Levy and Razin (2018)

Sequential social learning

- rational agents:
  - Acemoglu, Dahleh, Lobel, and Ozdaglar (2011)
  - Lobel and Sadler (2015)
- behavioral agents:
  - Eyster and Rabin (2010)
  - Bohren (2016)
  - Bohren and Hauser (2017WP)
  - Eyster and Rabin (2014)

The Social Learning Game

Basic setup

- binary state of the world $\omega \in \{0, 1\}$, equally likely
- sequence of agents indexed by $i = 1, 2, 3, \ldots$, move in turn

On agent i's turn

- observe private signal s_i
- observe actions of some previous agents (next slide)
- form belief about ω (next slide)
- play $a_i \in [0, 1]$ to maximize $\mathbb{E}[-(a_i - \omega)^2]$

Gaussian private signals (for this talk)

- $s_i \sim \mathcal{N}(1, \sigma^2)$ when $\omega = 1$
- $s_i \sim \mathcal{N}(-1, \sigma^2)$ when $\omega = 0$
- signals conditionally i.i.d. given ω
- more general signal structure in paper

The Social Learning Game

Network observation

- agent i observes actions of neighbors $N_i \subseteq \{1, 2, ..., i-1\}$
- neighborhoods N_i for $i = 1, 2, 3, ...$ define a directed network
- adjacency matrix M defined by $M_{ij} = 1$ if $j \in N_i$, $M_{ij} = 0$ else

Inferential naiveté

- i observes $(a_j)_{j \in N_i}$ and thinks $a_j = \mathbb{P}[\omega = 1 \mid s_j]$ for each $j \in N_i$
- agents are Bayesians except for this mistake
- can be thought of as...
  - i has the misspecified model $N_j = \emptyset$ for each $j \in N_i$
  - i neglects redundancy/correlation in observed actions
- Eyster and Rabin (2010) considers this assumption on a complete network

Actions and Network Paths

Change of variable:

$\tilde{s}_i := \ln\left(\frac{\mathbb{P}(\omega=1|s_i)}{\mathbb{P}(\omega=0|s_i)}\right)$, log-likelihood ratio of states given signal s_i

$\tilde{a}_i := \ln\left(\frac{a_i}{1-a_i}\right)$, under log-likelihood ratio $\tilde{a}_i$, action a_i is optimal

Proposition (log-linear expression of actions)

$$\begin{pmatrix} \tilde{a}_1 \\ \vdots \\ \tilde{a}_n \end{pmatrix} = (I - M)^{-1} \cdot \begin{pmatrix} \tilde{s}_1 \\ \vdots \\ \tilde{s}_n \end{pmatrix}$$

where $(I - M)^{-1} = I + M + M^2 + M^3 + \ldots$

Note: $(M^L)_{i,j}$ counts paths of length L from i to j in network M.

So $\tilde{a}_i$ is a linear combination of $(\tilde{s}_j)_{j \leq i}$ with coefficients equal to number of paths (of any length) from i to j in the network M

Weighted Networks

Now let adjacency matrix entries be non-integral, $M_{ij} \in [0, 1]$

- think of M_{ij} as the weight i puts on j's action

Formally, log-linear expression of actions obtains under either of following interpretations of the weights M_{ij}

- **Noisy observations**: Instead of observing $(\tilde{a}_j)_{j \in N_i}$, agent i observes $(\tilde{a}_j + \epsilon_{ij})_{j \in N_i}$ where $\epsilon_{ij} \sim \mathcal{N}(0, \frac{1}{M_{ij}} - 1)$
- **Generations interpretation**:
  - replace each agent i with a continuum of agents (generation i)
  - all members of generation i receive the same signal s_i
  - stochastic social info: each member observes one random member of generation j with probability M_{ij}, for each $j < i$
  - interpret $\tilde{a}_i$ in Proposition as mean of log action distribution in generation i

Probability of Correct Learning

Let $b_{i,j} := (I - M)^{-1}_{ij}$ count number of weighted paths from i to j in M — number of times j's signal indirectly enters i's social info

Theorem (probability that n is correct about the state)

$$\mathbb{P}[a_n > \frac{1}{2} \mid \omega = 1] = \Phi\left(\frac{1}{\sigma} \cdot \frac{\left\|\vec{b}_n\right\|_1}{\left\|\vec{b}_n\right\|_2}\right)$$

Φ is Gaussian cdf, $\vec{b}_n := (b_{n,1}, ..., b_{n,n})$, $||\cdot||_1$ is 1-norm, $||\cdot||_2$ is 2-norm.

Proof idea: with Gaussian signals, log-likelihoods $\tilde{s}_i$ also Gaussian. By previous Proposition,

$$\tilde{a}_n = \sum_{i=1}^{n} b_{n,i} \cdot \tilde{s}_i \sim \frac{2}{\sigma^2} \cdot \mathcal{N}(\left\|\vec{b}_n\right\|_1, \left\|\vec{b}_n\right\|_2^2).$$

Remark: $\liminf_{n\to\infty} \mathbb{P}[a_n > \frac{1}{2} \mid \omega = 1] < 1$ for almost all common models of network M — imperfect learning is the leading case.

Uniform Weights

Society mislearns when $a_n \to 0$ but $\omega = 1$, or $a_n \to 1$ but $\omega = 0$.

Proposition (denser networks lead to more mislearning)

Consider the network where each link has weight $0 < q \leq 1$. Then

1. *Agents' actions converge almost surely to 0 or 1;*
2. *Probability of mislearning increases in q.*

Intuition: On sparse networks, early agents do not influence each other much. So social consensus incorporates more independent sources of information and is more likely to be correct.

Related intuition: as q grows, agents' beliefs about network structure deviate more and more from the truth.

By contrast, density has no effect on long-run learning accuracy for rational agents or DeGroot agents (in large networks).

Uniform Weights — Proof Sketch

Main technique (also used in paper to analyze other network structures): recursive expression of path counts $b_{n,i}$.

Weight-preserving bijection between

- $P_1 =$ paths from n to i that do not pass through $n-1$
- $P_2 =$ paths from $n-1$ to i

Get the recursion $b_{n,i} = (1+q) \cdot b_{n-1,i}$, because:

- each path from n to i is either in P_1, or consists of a path in P_2 prefixed with the edge $n \to (n-1)$
- edge $n \to (n-1)$ applies multiplicative factor q to path weight

Can explicitly solve for $\vec{b}_n$ and calculate $\frac{||\vec{b}_n||_1}{||\vec{b}_n||_2}$ as a function of q

By previous Theorem, this determines the probability of mislearning.

Experiment: Setup

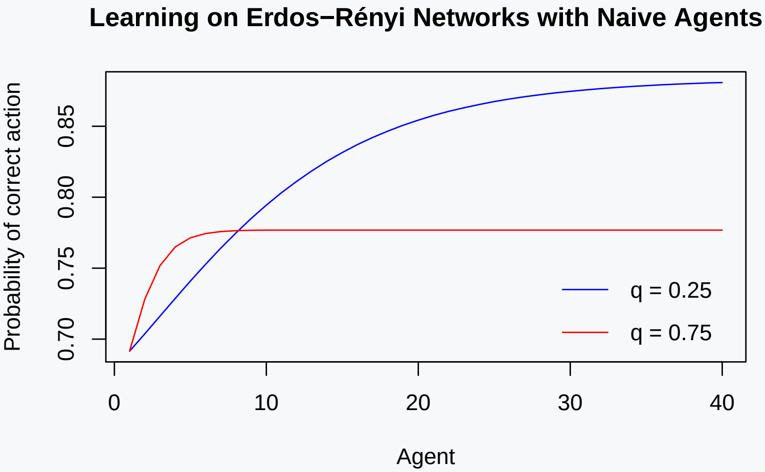
The Game:

- Binary state L or R drawn at the start of each game
- 40 agents move in order
- On his turn, agent receives private signal about state, sees the guesses of some predecessors, then enters a guess L or R

Two Network Densities:

- Each predecessor's guess is observed with probability q
- Each game g has either $q_g = 0.25$ (sparse) or $q_g = 0.75$ (dense)
- Compare average learning accuracy across games with different densities

Predictions under Inferential Naiveté



Predictions for Rational Agents

Results of Acemoglu, Dahleh, Lobel, and Ozdaglar (2011) imply asymptotic learning of state regardless of q

Is 40 agents enough for this limit?

Using technique similar to Lobel and Sadler (2015)'s neighborhood choice function, can compute explicit lower bound on the accuracy of rational agents

This lower bound is 97% for 33rd agent on dense network — so sparse network cannot improve accuracy much, if at all

Experiment: Logistics

Subjects

- Experiment done on Amazon Mechanical Turk, 1040 subjects
- Must pass a three-question comprehension check
- Each plays 10 games with same density in same position
- Subjects know network-generating process
- \$0.25 per correct guess. Average: \$2.08 for less than 10 min

Signals and accuracy

- Private signal $\sim \mathcal{N}(-1, 4)$ in state L, $\sim \mathcal{N}(1, 4)$ in state R
- Can have accuracy 69% from using private signal alone
- $\tilde{y}_g$ — fraction of last 8 agents who guess correctly in game g

Dataset: 260 games, half with each density. Regress across games

$$\tilde{y}_g = \beta_0 + \beta_1 q_g + \epsilon_g$$

The experiment — including sample size, measure of long-run accuracy, and statistical analysis — was pre-registered on the AsPredicted registry prior to data collection.

Experiment: Results

	Dependent variable:
	FractionCorrect
NetworkDensity	−0.092**
	(0.041)
Constant	0.802***
	(0.022)
Observations	260
R^2	0.020
Adjusted R^2	0.016
Residual Std. Error	0.164 (df = 258)
F Statistic	5.166** (df = 1; 258)
Note:	*p<0.1; **p<0.05; ***p<0.01

Experiment: Results

Accuracy gain from social learning:

- In dense networks, last 8 agents guess correctly 5.7% more often than if they had no social observations
- This accuracy gain is 12.6% in sparse networks, more than twice as large (p-value 0.0239)

Source of this difference:

- Agent goes against signal if guess L with a positive signal or guess R with a negative signal
- Among agents in the last 8 positions, 138 instances of this in sparse networks, 136 instances in dense networks
- Accuracy conditional on going against signal:
  - 82% in sparse networks
  - 71% in dense networks
- So driven by differential effectiveness of social learning

Disagreement with Partial Segregation

Modification of baseline model:

- each agent has a binary action set (as in experiment)
- chooses $a_i = 0$ or $a_i = 1$ to maximize $\mathbb{P}[a_i = \omega]$

Network with two groups:

- Odd-numbered agents in one group, even-numbered agents in another group
- Random network generated by stochastic block model:
  - prob q_s of observing each predecessor in the same group
  - prob q_d of observing each predecessor in the opposite group
- Result also true for weighted network M with $M_{i,j} \in \{q_s, q_d\}$ depending on if $i \equiv j \pmod 2$

Proposition (persistent disagreement with coarse actions)

Suppose $q_s > q_d > 0$. Then there is a positive probability that all odd-numbered agents choose action 0 while all even-numbered agents choose action 1.

Disagreement with Partial Segregation

Proposition (persistent disagreement with coarse actions)

Suppose $q_s > q_d > 0$*. Then there is a positive probability that all odd-numbered agents choose action* 0 *while all even-numbered agents choose action* 1.

Disagreement persists even though there are infinitely many connections between the two groups.

By contrast, asymptotic agreement is a robust prediction of both rational and DeGroot models on connected networks.

Proof idea: if two groups get opposite signals early on and each agent mostly sees neighbors from own group, groups never agree.

Remark: with continuous actions, we prove disagreement almost surely does not persist (in both random network and weighted network versions of this model). But binary actions lead to more information loss, so disagreement persists with positive probability.

Conclusion

Question: how does the structure of the observation network affect the probability of correct social learning in the long-run?

For agents who suffer from inferential naiveté, we have studied how network influences long-run learning outcome

- exact expression for the accuracy of each agent
- mislearning more likely on denser networks
- experimentally measured learning accuracy, found the accuracy gain from social learning twice as large on sparse networks
- disagreement can persist forever with partial segregation

Thank you!