Aggregative Efficiency of Bayesian Learning in Networks

Krishna Dasaratha   Kevin He

ACM EC'21

Social-Learning Dynamics in Different Networks

- **Social learning**: info about unknown state dispersed among society of agents, agents act based on private signals and observations of social neighbors
- How does social network affect efficiency of info aggregation?
- Esp. relevant today as technology reshapes networks
- Existing theoretical work focuses on complete network
- Less known about how rational social learning compares across networks, and existing results say agents **eventually** learn completely on **all** (reasonable) networks
- Open question: impact of network on **how well signals are aggregated** — hence how quickly rational agents learn

Golub and Sadler (2016): "A significant gap in our knowledge concerns short-run dynamics and rates of learning in these models."

Environment and Key Results

Introduce **tractable model of rational sequential learning** that lets us compare learning dynamics across different networks

- fine-grained ranking of networks wrt social-learning efficiency

Highlight **network-based informational confounds**

- suppose P2 and P3 see P1, but P4 sees only P2 and P3
- from P4's perspective, P1's action confounds the info content of P2 and P3's behavior
- "intransitivity" that appears in almost all realistic observation networks can lead to arbitrarily inefficient social learning

Generations network – observe subset of agents in previous gen

- express learning rate as simple function of network parameters
- extent of info loss: under a symmetry condition, learning aggregates **no more than 2 signals per gen** asymptotically

Related Social-Learning Literature

Sequential learning: Banerjee (1992), Bikhchandani, Hirshleifer, Welch (1992)

Obstructions to the efficient learning rate in sequential social learning

- Coarse action space: Harel, Mossel, Strack, Tamuz (2020), Rosenberg and Vieille (2019), Hann-Caruthers, Martynov, Tamuz (2018)
- Endogenous info: Burguet and Vives (2000), Mueller-Frank and Pai (2016), Ali (2018), Lomys (2020), Liang and Mu (2020)
- This paper: **network-based** obstructions to efficient learning

Network structure and social learning

- Network does not matter (within "reasonable" class) for long-run learning: Acemoglu, Dahleh, Lobel, and Ozdaglar (2011), Lobel and Sadler (2015), Rosenberg and Vieille (2019)
- Examples and numerical simulations suggesting network affects learning in finite populations: Sgroi (2002), Lobel, Acemoglu, Dahleh, and Ozdaglar (2009), Arieli and Mueller-Frank (2019)
- This paper: **analytic ranking of networks on rate of learning**

Speed of learning under heuristics: Ellison and Fudenberg (1993), Golub and Jackson (2012), Molavi, Tahbaz-Salehi, Jadbabaie (2018). This paper: **rational Bayesian learning**

Model and Notations

- Two equally likely states $\omega \in \{0, 1\}$
- Agents $i = 1, 2, 3, ...$ move in order, each acting once
  - i observes **private signal** $s_i \in \mathbb{R}$ and actions of **neighbors**, $N(i) \subseteq \{1, ...i-1\}$
  - picks **action** $a_i \in [0, 1]$ to maximize expectation of $-(a_i - \omega)^2$
- Signals are Gaussian and conditionally i.i.d. given state, $s_i \sim \mathcal{N}(1, \sigma^2)$ when $\omega = 1$ and $s_i \sim \mathcal{N}(-1, \sigma^2)$ when $\omega = 0$
- Neighborhoods (observation network) = common knowledge
- A **strategy** for i specifies i's play as a function of:
  1. observed actions from neighbors $N(i)$, and
  2. private signal s_i.
- Can only observe earlier agents $\Rightarrow$ there is a unique Bayesian-Nash **equilibrium** strategy profile

Signal-Counting Interpretation of Eqm Accuracy

If i's only info is $n \in \mathbb{N}_+$ indep signals, $\ln\left(\frac{a_i}{1-a_i}\right) \sim \mathcal{N}\left(\pm n \cdot \frac{2}{\sigma^2}, n \cdot \frac{4}{\sigma^2}\right)$

Definition

Social learning **aggregates** $r \in \mathbb{R}_+$ **signals by agent** i if eqm log-action $\ln\left(\frac{a_i^*}{1-a_i^*}\right) \sim \mathcal{N}\left(\pm r \cdot \frac{2}{\sigma^2}, r \cdot \frac{4}{\sigma^2}\right)$ in two states.

- When agents use arbitrary strategy profile (even if log-linear), need not hold for **any** $r \in \mathbb{R}$
- But, **equilibrium** log-actions always admit this kind of signal-counting interpretation, suff. stat for rational accuracy

Proposition 1

For every network, there exist $(r_i)_{i \geq 1}$ so that social learning aggregates r_i signals by agent i. These $(r_i)_{i \geq 1}$ don't depend on σ^2.

- Can measure each i's eqm accuracy in units of private signals

Aggregative Efficiency

- Equilibrium actions converge to true state in probability (i.e., complete long-run learning) iff $r_i \to \infty$
- Turns out $r_i \to \infty$ for all networks satisfying a **very weak condition** (Acemoglu, Dahleh, Lobel, and Ozdaglar (2011)'s "expanding observations" for a non-random network)
- Complete long-run learning not useful for ranking networks

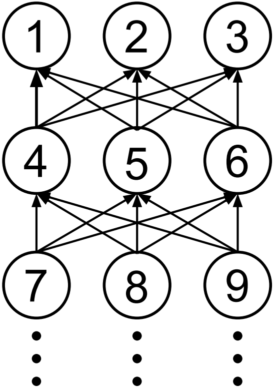
Definition

$\lim_{i\to\infty}(r_i/i)$ is the **aggregative efficiency** of the network

- What fraction of signals in the entire society do individuals aggregate under social learning?
- Can have $r_i \to \infty$ but $\lim_{i\to\infty}(r_i/i)$ near 0: complete long-run learning, but get there very slowly
- Rest of the talk: compare networks for social learning by comparing their aggregative efficiency

Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t-1$



Proposition 2

In maximal generations networks:

- *Society learns completely in the long run with any K.*
- $\lim_{i \to \infty}(r_i/i) = \frac{(2K-1)}{K^2}$.
- *In the long run, social learning aggregates...*
  - *fewer signals per agent with larger K*
  - *fewer than 2 signals per generation with any K*
- *After generation 2, social learning aggregates fewer than 3 signals per generation with any K*

Bounds on Signals Aggregated Per Generation

- Social learning must aggregate at least 1 signal per gen (improvement by combining own signal with social obs)
- This lower-bound not too far from the actual learning rate:

$$r_i \,/ \underbrace{\lceil i/K \rceil}_{\text{gen of } i} = \underbrace{\frac{(2K-1)}{K}}_{<2} + o(1)$$

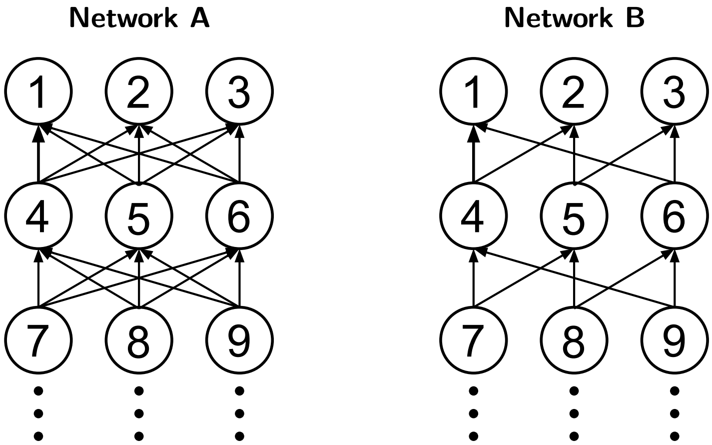
(No more than **2 signals** per gen in long-run, for any K)

$$r_i - r_{i'} \leq 3, \qquad \text{for } i, i' \text{ in gen } t, t-1 \text{ where } t \geq 3$$

(No more than **3 signals** per gen **starting with gen 3**, for any K)

- For K large, individuals only manage to aggregate an unboundedly small fraction of their private signals in eqm
- Someone in gen $t+1$ finds it hard to figure out gen t's private signals due to **info confounding**: which part of neighbors' actions come from their signals, and which part from their own social observations?

Which Network Leads to Faster Learning?



- **Network A** is the maximal generations network with $K = 3$
- **Network B** puts agents in each gen into 3 slots, $k \in \{1, 2, 3\}$. $k = 1$ sees 1 and 2, $k = 2$ sees 2 and 3, $k = 3$ sees 3 and 1.
- Fewer social observations, but also less info confounding
- Need: aggregative efficiency on more general networks

Generations Network with Partial Observations

- Generations network with K agents per gen
- Each agent observes a subset of predecessors in previous gen

Definition

The network is **symmetric** if all agents observe $d \geq 1$ neighbors and all pairs of agents in the same generation share c common neighbors.

For example, "**Network B**" is symmetric with $d = 2$, $c = 1$

Aggregative Efficiency with Partial Observations

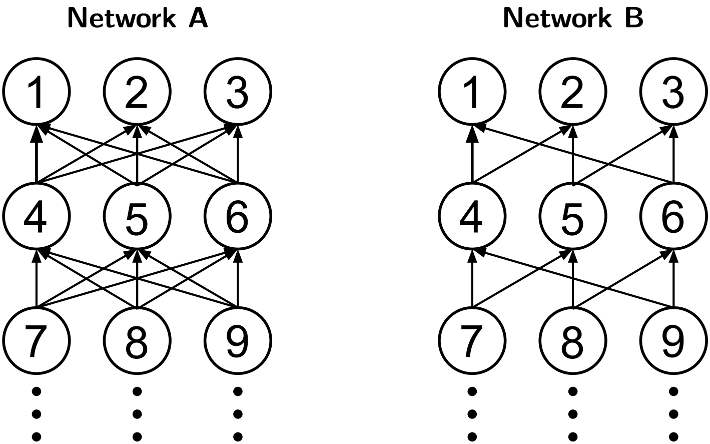
Theorem 1

In symmetric generations networks,

$$\lim_{i\to\infty}(r_i/i) = \left(1 + \frac{d^2 - d}{d^2 - d + c}\right)\frac{1}{K}.$$

- Exact expression of aggregative efficiency for a broader class of generations networks
- Term in parenthesis increases in d and decreases in c — more obs speeds up rate of learning per gen but more confounding slows it down, all else equal
- Maximal gen network has the worst rate of learning, among all symmetric gen networks with same d
  - Because actions very confounded in maximal gen network
- But Theorem 1 shows asymptotic bound of 2 signals per gen applies to **all** such networks, strengthening Proposition 2

Which Network Leads to Faster Learning?



- Applying Theorem 1, aggregative efficiency is the same in **Network A** ($d = 3, c = 3$) and **Network B** ($d = 2, c = 1$)!
- Extra social obs exactly cancel out reduced info content of each obs

Summary

- A tractable model of rational sequential learning that focuses on how the social network affects aggregative efficiency
- Generally, network confounds info content of neighbors' behavior and leads to info loss
- Exact aggregative efficiency in all generations networks with symmetric observations
- Significant info loss due to confounding: in any such network, each generation eventually aggregates no more than 2 signals

Thank you!