Asymptotic Learning on Social Networks

Elchanan Mossel
Joint work with:
Allan Sly (UC Berkeley)
Omer Tamuz (Weizmann)

February 23, 2012

The Learning on Networks Challenge

- The big question: **Is learning / inference possible where agents communicate over a network?**

The Learning on Networks Challenge

- The big question: **Is learning / inference possible where agents communicate over a network?**
- We consider models from economics where agents on some network observe the opinions and actions of their neighbors.

The Learning on Networks Challenge

- The big question: **Is learning / inference possible where agents communicate over a network?**
- We consider models from economics where agents on some network observe the opinions and actions of their neighbors.
  - Agents are Bayesian, their opinions are governed solely by what information they have observed. This is the most "rational" opinion to have.

The Learning on Networks Challenge

- The big question: **Is learning / inference possible where agents communicate over a network?**
- We consider models from economics where agents on some network observe the opinions and actions of their neighbors.
  - Agents are Bayesian, their opinions are governed solely by what information they have observed. This is the most "rational" opinion to have.
  - A *communications* protocol is given through which agents receive information from their neighbors.

The Learning on Networks Challenge

- The big question: **Is learning / inference possible where agents communicate over a network?**
- We consider models from economics where agents on some network observe the opinions and actions of their neighbors.
  - Agents are Bayesian, their opinions are governed solely by what information they have observed. This is the most "rational" opinion to have.
  - A *communications* protocol is given through which agents receive information from their neighbors.
- **Question:** Do agents in this decentralized model aggregate their information effectively?

Basic Assumptions of the Model

- The world is a probability space.

Basic Assumptions of the Model

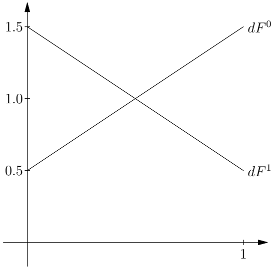
- The world is a probability space.
- An unknown **"state of the world"** $S \in \{-1, 1\}$.
  For example:
  - Which is the better purchase?
    iGadget or **Gadgetoid**?
  - Is there a bubble in the tulip market?
    Yes or **no**?
  - Who would make a better president?
    Gin-g-rich or **Rich**?

Basic Assumptions of the Model

- The world is a probability space.
- An unknown **"state of the world"** $S \in \{-1, 1\}$.
  For example:
  - Which is the better purchase?
    iGadget or **Gadgetoid**?
  - Is there a bubble in the tulip market?
    Yes or **no**?
  - Who would make a better president?
    Gin-g-rich or **Rich**?
- S takes the value ± 1 with probability $1/2$.

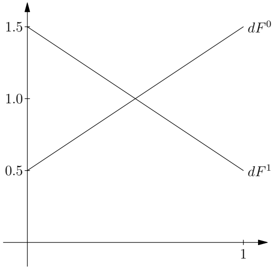
Basic Assumptions of the Model

- The world is a probability space.
- An unknown **"state of the world"** $S \in \{-1, 1\}$.
  For example:
  - Which is the better purchase?
    iGadget or **Gadgetoid**?
  - Is there a bubble in the tulip market?
    Yes or **no**?
  - Who would make a better president?
    Gin-g-rich or **Rich**?
- S takes the value ± 1 with probability $1/2$.
- n agents. Each is given a **private signal** ω_i with distribution
  - F^- if $S = -1$
  - F^+ if $S = 1$



Basic Assumptions of the Model

- The world is a probability space.
- An unknown **"state of the world"** $S \in \{-1, 1\}$. For example:
    - Which is the better purchase? **iGadget** or **Gadgetoid**?
    - Is there a bubble in the tulip market? **Yes** or **no**?
    - Who would make a better president? **Gin-g-rich** or **Rich**?
- S takes the value ± 1 with probability $1/2$.
- n agents. Each is given a **private signal** ω_i with distribution
    - F^- if $S = -1$
    - F^+ if $S = 1$
- Conditioned on S, private signals are **independent**.



Assumption & Questions

- Class of models
  - A group of **Bayesian** agents.
  - The agents need to make a **decision**.
  - The agents have private information sources and **learn** from each other.

Assumption & Questions

- Class of models
  - A group of **Bayesian** agents.
  - The agents need to make a **decision**.
  - The agents have private information sources and **learn** from each other.
- Two main models
  - **Revealed beliefs:** In each time step agent i tells her neighbors her posterior probability $X_i(t) = \mathbb{P}\left[S = 1 \mid \mathcal{F}_i(t)\right]$.

Assumption & Questions

- Class of models
  - A group of **Bayesian** agents.
  - The agents need to make a **decision**.
  - The agents have private information sources and **learn** from each other.
- Two main models
  - **Revealed beliefs:** In each time step agent i tells her neighbors her posterior probability $X_i(t) = \mathbb{P}[S = 1 \mid \mathcal{F}_i(t)]$.
  - **Revealed actions:** In each time step agent i tells her neighbors her preferred "action" $A_i(t) = \text{argmax}_{a \in \{-1,1\}} \mathbb{P}[S = a \mid \mathcal{F}_i(t)]$.

Assumption & Questions

- Class of models
  - A group of **Bayesian** agents.
  - The agents need to make a **decision**.
  - The agents have private information sources and **learn** from each other.
- Two main models
  - **Revealed beliefs:** In each time step agent i tells her neighbors her posterior probability $X_i(t) = \mathbb{P}[S = 1 \mid \mathcal{F}_i(t)]$.
  - **Revealed actions:** In each time step agent i tells her neighbors her preferred "action" $A_i(t) = \text{argmax}_{a \in \{-1,1\}} \mathbb{P}[S = a \mid \mathcal{F}_i(t)]$.
- Two Major Questions
  - When do the agents reach consensus?
    - **Generally well understood** (Aumann (1976), Sebenius and Geanakoplos (1983), Parikh and Krasucki (1990), Rosenberg, Solan and Vieille (2009); Sebenius and Geanakoplos (1983), Ménager (2006)

Assumption & Questions

- Class of models
  - A group of **Bayesian** agents.
  - The agents need to make a **decision**.
  - The agents have private information sources and **learn** from each other.
- Two main models
  - **Revealed beliefs:** In each time step agent i tells her neighbors her posterior probability $X_i(t) = \mathbb{P}[S = 1 \mid \mathcal{F}_i(t)]$.
  - **Revealed actions:** In each time step agent i tells her neighbors her preferred "action" $A_i(t) = \text{argmax}_{a \in \{-1,1\}} \mathbb{P}[S = a \mid \mathcal{F}_i(t)]$.
- Two Major Questions
  - When do the agents reach consensus?
    - **Generally well understood** (Aumann (1976), Sebenius and Geanakoplos (1983), Parikh and Krasucki (1990), Rosenberg, Solan and Vieille (2009); Sebenius and Geanakoplos (1983), Ménager (2006)
  - When do the agents learn from each other efficiently?
    - **Generally poorly understood.**

Example: Revealed Beliefs Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.

Example: Revealed Beliefs Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.

Example: Revealed Beliefs Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.

Example: Revealed Beliefs Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.
- $X_1(1) = X_2(1) = \mathbb{P}[S = + \mid 0.4, 0.7]$.

Example: Revealed Beliefs Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.
- $X_1(1) = X_2(1) = \mathbb{P}[S = + \mid 0.4, 0.7]$.
- $$\frac{\mathbb{P}[0.4, 0.7 \mid S = +]}{\mathbb{P}[0.4, 0.7 \mid S = -]} = \frac{0.8 * 1.4}{1.2 * 0.6} = \frac{14}{9}.$$

Example: Revealed Beliefs Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.
- $X_1(1) = X_2(1) = \mathbb{P}[S = + \mid 0.4, 0.7]$.
- $$\frac{\mathbb{P}[0.4, 0.7 \mid S = +]}{\mathbb{P}[0.4, 0.7 \mid S = -]} = \frac{0.8 * 1.4}{1.2 * 0.6} = \frac{14}{9}.$$
- Challenge: $G = \{0, 1\}^3$ with signals $0.11, 0.22, 0.33, 0.44, 0.55, 0.66, 0.77, 0.88$.

Example: Revealed Beliefs Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.
- $X_1(1) = X_2(1) = \mathbb{P}[S = + \mid 0.4, 0.7]$.
- $$\frac{\mathbb{P}[0.4, 0.7 \mid S = +]}{\mathbb{P}[0.4, 0.7 \mid S = -]} = \frac{0.8 * 1.4}{1.2 * 0.6} = \frac{14}{9}.$$
- Challenge: $G = \{0, 1\}^3$ with signals $0.11, 0.22, 0.33, 0.44, 0.55, 0.66, 0.77, 0.88$.
- Explicitly calculate the dynamics and what it converges to.

Example: Revealed Beliefs Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.
- $X_1(1) = X_2(1) = \mathbb{P}[S = + \mid 0.4, 0.7]$.
- $$\frac{\mathbb{P}[0.4, 0.7 \mid S = +]}{\mathbb{P}[0.4, 0.7 \mid S = -]} = \frac{0.8 * 1.4}{1.2 * 0.6} = \frac{14}{9}.$$
- Challenge: $G = \{0, 1\}^3$ with signals $0.11, 0.22, 0.33, 0.44, 0.55, 0.66, 0.77, 0.88$.
- Explicitly calculate the dynamics and what it converges to.
- Does it converge in a finite number of iterations?

The Revealed Beliefs Model

- Model Specification:
  - The agents lie on a connected undirected graph.
  - At time 0 each agents receive her own signal $X_0(t)$.

The Revealed Beliefs Model

- Model Specification:
  - The agents lie on a connected undirected graph.
  - At time 0 each agents receive her own signal $X_0(t)$.
- Interaction Protocol:
  - In each time step t agent i tells her neighbors her posterior probability $X_i(t) = \mathbb{P}[S = 1 \mid \mathcal{F}_i(t)]$.
  - $\mathcal{F}_i(t+1)$ is generated by $(X_i(s), X_j(s) : s \leq t, j \sim i)$

The Revealed Beliefs Model

- Model Specification:
  - The agents lie on a connected undirected graph.
  - At time 0 each agents receive her own signal $X_0(t)$.
- Interaction Protocol:
  - In each time step t agent i tells her neighbors her posterior probability $X_i(t) = \mathbb{P}[S = 1 \mid \mathcal{F}_i(t)]$.
  - $\mathcal{F}_i(t+1)$ is generated by $(X_i(s), X_j(s) : s \leq t, j \sim i)$
- Claim: Eventually all agents agree: $X_i(t) \to X(\infty)$ a.s.

The Revealed Beliefs Model

- Model Specification:
  - The agents lie on a connected undirected graph.
  - At time 0 each agents receive her own signal $X_0(t)$.
- Interaction Protocol:
  - In each time step t agent i tells her neighbors her posterior probability $X_i(t) = \mathbb{P}[S = 1 \mid \mathcal{F}_i(t)]$.
  - $\mathcal{F}_i(t+1)$ is generated by $(X_i(s), X_j(s) : s \leq t, j \sim i)$
- Claim: Eventually all agents agree: $X_i(t) \to X(\infty)$ a.s.
  - Proof: Exercise! (Aumann 76 etc.). Pf Sketch:

The Revealed Beliefs Model

- Model Specification:
  - The agents lie on a connected undirected graph.
  - At time 0 each agents receive her own signal $X_0(t)$.
- Interaction Protocol:
  - In each time step t agent i tells her neighbors her posterior probability $X_i(t) = \mathbb{P}[S = 1 \mid \mathcal{F}_i(t)]$.
  - $\mathcal{F}_i(t+1)$ is generated by $(X_i(s), X_j(s) : s \leq t, j \sim i)$
- Claim: Eventually all agents agree: $X_i(t) \to X(\infty)$ a.s.
  - Proof: Exercise! (Aumann 76 etc.). Pf Sketch:
  - $X_i(t) \to X_i(\infty)$ a.s. as a bounded martingale.

The Revealed Beliefs Model

- Model Specification:
  - The agents lie on a connected undirected graph.
  - At time 0 each agents receive her own signal $X_0(t)$.
- Interaction Protocol:
  - In each time step t agent i tells her neighbors her posterior probability $X_i(t) = \mathbb{P}[S = 1 \mid \mathcal{F}_i(t)]$.
  - $\mathcal{F}_i(t+1)$ is generated by $(X_i(s), X_j(s) : s \leq t, j \sim i)$
- Claim: Eventually all agents agree: $X_i(t) \to X(\infty)$ a.s.
  - Proof: Exercise! (Aumann 76 etc.). Pf Sketch:
  - $X_i(t) \to X_i(\infty)$ a.s. as a bounded martingale.
  - If $i \sim j$ then $X_i(\infty) \in \mathcal{F}_j(\infty)$ and $X_j(\infty) \in \mathcal{F}_i(\infty)$.

The Revealed Beliefs Model

- Model Specification:
  - The agents lie on a connected undirected graph.
  - At time 0 each agents receive her own signal $X_0(t)$.
- Interaction Protocol:
  - In each time step t agent i tells her neighbors her posterior probability $X_i(t) = \mathbb{P}[S = 1 \mid \mathcal{F}_i(t)]$.
  - $\mathcal{F}_i(t+1)$ is generated by $(X_i(s), X_j(s) : s \le t, j \sim i)$
- Claim: Eventually all agents agree: $X_i(t) \to X(\infty)$ a.s.
  - Proof: Exercise! (Aumann 76 etc.). Pf Sketch:
  - $X_i(t) \to X_i(\infty)$ a.s. as a bounded martingale.
  - If $i \sim j$ then $X_i(\infty) \in \mathcal{F}_j(\infty)$ and $X_j(\infty) \in \mathcal{F}_i(\infty)$.
  - If $X_i(\infty) \neq X_j(\infty)$ then

$$\mathrm{Var}\left[\frac{X_i(\infty) + X_j(\infty)}{2} - S\right] < \max_{k=i,j}(\mathrm{Var}\left[X_k(\infty) - S\right])$$

  Contradiction!

The Revealed Beliefs Model

- Model Specification:
  - The agents lie on a connected undirected graph.
  - At time 0 each agents receive her own signal $X_0(t)$.
- Interaction Protocol:
  - In each time step t agent i tells her neighbors her posterior probability $X_i(t) = \mathbb{P}[S = 1 \mid \mathcal{F}_i(t)]$.
  - $\mathcal{F}_i(t+1)$ is generated by $(X_i(s), X_j(s) : s \leq t, j \sim i)$
- Claim: Eventually all agents agree: $X_i(t) \to X(\infty)$ a.s.
  - Proof: Exercise! (Aumann 76 etc.). Pf Sketch:
  - $X_i(t) \to X_i(\infty)$ a.s. as a bounded martingale.
  - If $i \sim j$ then $X_i(\infty) \in \mathcal{F}_j(\infty)$ and $X_j(\infty) \in \mathcal{F}_i(\infty)$.
  - If $X_i(\infty) \neq X_j(\infty)$ then

$$\mathrm{Var}\left[\frac{X_i(\infty) + X_j(\infty)}{2} - S\right] < \max_{k=i,j}(\mathrm{Var}\left[X_k(\infty) - S\right])$$

    Contradiction!
  - Note: Proof doesn't require independent signals.

Example: Revealed Actions Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.

Example: Revealed Actions Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.

Example: Revealed Actions Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.

Example: Revealed Actions Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.
- $A_1(0) = -, A_2(0) = +$.

Example: Revealed Actions Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.
- $A_1(0) = -, A_2(0) = +$.
- $X_1(1) = \mathbb{P}[S = + \mid 0.4, \geq 0.5]$.

Example: Revealed Actions Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.
- $A_1(0) = -, A_2(0) = +$.
- $X_1(1) = \mathbb{P}[S = + \mid 0.4, \geq 0.5]$.
- $$\frac{\mathbb{P}[0.4, \geq 0.5 \mid S = +]}{\mathbb{P}[0.4, \geq 0.5 \mid S = -]} = \frac{0.8 * 3/4}{1.2 * 1/4} = 2 \implies A_1(1) = +.$$
  $$\frac{\mathbb{P}[0.7, \leq 0.5 \mid S = +]}{\mathbb{P}[0.7, \leq 0.5 \mid S = -]} = \frac{1.4 * 1/4}{0.6 * 3/4} = \frac{7}{9} \implies A_2(1) = -.$$

Example: Revealed Actions Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.
- $A_1(0) = -, A_2(0) = +$.
- $X_1(1) = \mathbb{P}[S = + \mid 0.4, \geq 0.5]$.
- $$\frac{\mathbb{P}[0.4, \geq 0.5 \mid S = +]}{\mathbb{P}[0.4, \geq 0.5 \mid S = -]} = \frac{0.8 * 3/4}{1.2 * 1/4} = 2 \implies A_1(1) = +.$$
  $$\frac{\mathbb{P}[0.7, \leq 0.5 \mid S = +]}{\mathbb{P}[0.7, \leq 0.5 \mid S = -]} = \frac{1.4 * 1/4}{0.6 * 3/4} = \frac{7}{9} \implies A_2(1) = -.$$
- etc.

Example: Revealed Actions Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.
- $A_1(0) = -, A_2(0) = +$.
- $X_1(1) = \mathbb{P}[S = + \mid 0.4, \geq 0.5]$.
- $$\frac{\mathbb{P}[0.4, \geq 0.5 \mid S = +]}{\mathbb{P}[0.4, \geq 0.5 \mid S = -]} = \frac{0.8 * 3/4}{1.2 * 1/4} = 2 \implies A_1(1) = +.$$
  $$\frac{\mathbb{P}[0.7, \leq 0.5 \mid S = +]}{\mathbb{P}[0.7, \leq 0.5 \mid S = -]} = \frac{1.4 * 1/4}{0.6 * 3/4} = \frac{7}{9} \implies A_2(1) = -.$$
- etc.
- Challenge: $V = \{1, 2, 3, 4\}, E = \{(1, 2), (2, 3), (3, 4)\}$ with signals $0.15, 0.41, 0.6, 0.87$.

Example: Revealed Actions Model

- Let $V = \{1, 2\}, E = \{(1, 2)\}$.
- Assume signal in $[0, 1]$ with $dF_+ = 2xdx$, $dF_- = (2 - 2x)dx$.
- Suppose $X_1(0) = 0.4$ and $X_2(0) = 0.7$.
- $A_1(0) = -, A_2(0) = +$.
- $X_1(1) = \mathbb{P}[S = + \mid 0.4, \geq 0.5]$.
- $$\frac{\mathbb{P}[0.4, \geq 0.5 \mid S = +]}{\mathbb{P}[0.4, \geq 0.5 \mid S = -]} = \frac{0.8 * 3/4}{1.2 * 1/4} = 2 \implies A_1(1) = +.$$
$$\frac{\mathbb{P}[0.7, \leq 0.5 \mid S = +]}{\mathbb{P}[0.7, \leq 0.5 \mid S = -]} = \frac{1.4 * 1/4}{0.6 * 3/4} = \frac{7}{9} \implies A_2(1) = -.$$
- etc.
- Challenge: $V = \{1, 2, 3, 4\}, E = \{(1, 2), (2, 3), (3, 4)\}$ with signals $0.15, 0.41, 0.6, 0.87$.
- Explicitly calculate the dynamics and what it converges to.
- Does it converge in a finite number of iterations?

The Revealed Actions Model

- Model Specification:
  - The agents lie on a connected undirected graph.
  - At time 0 each agents receive her own signal $X_0(t)$.

The Revealed Actions Model

- Model Specification:
  - The agents lie on a connected undirected graph.
  - At time 0 each agents receive her own signal $X_0(t)$.
- The revealed actions model:
  - In each time step t agent i tells her neighbors her preferred action$A_i(t) = \text{argmax}_{a \in \{-1,1\}} \mathbb{P}[S = a \mid \mathcal{F}_i(t)]$.
  - $\mathcal{F}_i(t+1)$ is generated by $(X_i(0), A_j(s) : s \leq t, j \sim i)$

The Revealed Actions Model

- Model Specification:
  - The agents lie on a connected undirected graph.
  - At time 0 each agents receive her own signal $X_0(t)$.
- The revealed actions model:
  - In each time step t agent i tells her neighbors her preferred action$A_i(t) = \text{argmax}_{a \in \{-1,1\}} \mathbb{P}[S = a \mid \mathcal{F}_i(t)]$.
  - $\mathcal{F}_i(t+1)$ is generated by $(X_i(0), A_j(s) : s \leq t, j \sim i)$
- Imitation principle (Gale and Kariv): If $i \sim j$ and $\lim A_i(t) \neq \lim A_j(t)$ then $X_i(\infty) = X_j(\infty) = 1/2$.

The Revealed Actions Model

- Model Specification:
  - The agents lie on a connected undirected graph.
  - At time 0 each agents receive her own signal $X_0(t)$.
- The revealed actions model:
  - In each time step t agent i tells her neighbors her preferred action$A_i(t) = \text{argmax}_{a \in \{-1,1\}} \mathbb{P}[S = a \mid \mathcal{F}_i(t)]$.
  - $\mathcal{F}_i(t+1)$ is generated by $(X_i(0), A_j(s) : s \leq t, j \sim i)$
- Imitation principle (Gale and Kariv): If $i \sim j$ and $\lim A_i(t) \neq \lim A_j(t)$ then $X_i(\infty) = X_j(\infty) = 1/2$.
- Stronger version (M-Sly-Tamuz-12): Under the non-atomic beliefs three possible limiting actions are possible
  1. For all i, $A_i(t) \to 1$ and $X_i(\infty) > \frac{1}{2}$.
  2. For all i, $A_i(t) \to -1$ and $X_i(\infty) < \frac{1}{2}$.
  3. For all i, $A_i(t)$ does not converge and $X_i(\infty) = \frac{1}{2}$.

Why do Economists care?

- A basic premise of economics is that good markets learn the state of the world. Is it true for (idealized) networked markets?

Why do Economists care?

- A basic premise of economics is that good markets learn the state of the world. Is it true for (idealized) networked markets?
- Aumann's original motivation: Bayesian Economics doesn't make sense. Doesn't allow to "agree to disagree".

Challenges - wasn't asymptotic learning studied before?

- Central problem in learning, i.e. Gale and Kariv ask: " whether the common action chosen asymptotically is optimal, in the sense that the same action would be chosen if all the signals were public information... there is no reason why this should be the case"

Challenges - wasn't asymptotic learning studied before?

- Central problem in learning, i.e. Gale and Kariv ask: " whether the common action chosen asymptotically is optimal, in the sense that the same action would be chosen if all the signals were public information... there is no reason why this should be the case"
- Instead simpler models are studied where either each agents acts only once, or agents are not truly rational.

Challenges - wasn't asymptotic learning studied before?

- Central problem in learning, i.e. Gale and Kariv ask: " whether the common action chosen asymptotically is optimal, in the sense that the same action would be chosen if all the signals were public information... there is no reason why this should be the case"
- Instead simpler models are studied where either each agents acts only once, or agents are not truly rational.
  - "different motivation is simply technical expediency" (Ellison and Fundenberg)

Challenges - wasn't asymptotic learning studied before?

- Central problem in learning, i.e. Gale and Kariv ask: " whether the common action chosen asymptotically is optimal, in the sense that the same action would be chosen if all the signals were public information... there is no reason why this should be the case"
- Instead simpler models are studied where either each agents acts only once, or agents are not truly rational.
  - "different motivation is simply technical expediency" (Ellison and Fundenberg)
  - "to keep the model mathematically tractable... this possibility [fully Bayesian agents] is precluded in our model... simplifying the belief revision process considerably" (Bala and Goyal)

Agreeing on beliefs implies optimal learning

- Belief Learning Theorem (M., Sly and Tamuz (2012))
  In the revealed beliefs model the limit $X = lim_{t \to \infty} X_i(t)$ satisfies:

  $$X = \mathbb{P}\left[S = 1 \mid \omega_1, \ldots, \omega_n\right],$$

  In particular there exist $c > 0$,

  $$\mathbb{P}\left[S = A\right] \geq 1 - e^{-cn}.$$

Agreeing on beliefs implies optimal learning

- Belief Learning Theorem (M., Sly and Tamuz (2012))
  In the revealed beliefs model the limit $X = lim_{t\to\infty} X_i(t)$ satisfies:

  $$X = \mathbb{P}\left[S = 1 \mid \omega_1, \ldots, \omega_n\right],$$

  In particular there exist $c > 0$,

  $$\mathbb{P}\left[S = A\right] \geq 1 - e^{-cn}.$$

- Agents aggregate their information optimally without necessarily learning all the signals.

Agreeing on beliefs implies optimal learning

- Belief Learning Theorem (M., Sly and Tamuz (2012))
  In the revealed beliefs model the limit $X = lim_{t \to \infty} X_i(t)$ satisfies:

  $$X = \mathbb{P}\left[S = 1 \mid \omega_1, \ldots, \omega_n\right],$$

  In particular there exist $c > 0$,

  $$\mathbb{P}\left[S = A\right] \geq 1 - e^{-cn}.$$

- Agents aggregate their information optimally without necessarily learning all the signals.
- As Aumann noted, the theorem is in fact much more general - it applies to any Bayesian model where all agents have same posterior.

Agreeing on beliefs implies optimal learning

- Belief Learning Theorem (M., Sly and Tamuz (2012))
  In the revealed beliefs model the limit $X = lim_{t \to \infty} X_i(t)$ satisfies:

  $$X = \mathbb{P}\left[S = 1 \mid \omega_1, \ldots, \omega_n\right],$$

  In particular there exist $c > 0$,

  $$\mathbb{P}\left[S = A\right] \geq 1 - e^{-cn}.$$

- Agents aggregate their information optimally without necessarily learning all the signals.
- As Aumann noted, the theorem is in fact much more general - it applies to any Bayesian model where all agents have same posterior.
- Independence of signals is needed. Example: $S = S_1 \oplus S_2$.

The Action Learning Theorem

- Recall that in the revealed actions model: $A_i(t) = \text{argmax}_{a \in \{-1,1\}} \mathbb{P}[a = S | \mathcal{F}_i(t)]$ is announced to the neighbors at time t.

The Action Learning Theorem

- Recall that in the revealed actions model: $A_i(t) = \text{argmax}_{a \in \{-1,1\}} \mathbb{P}[a = S | \mathcal{F}_i(t)]$ is announced to the neighbors at time t.
- Technical Assumption: "non-atomic beliefs": The belief $\mathbb{P}[S = 1 \mid \omega_1]$ is non-atomic which **avoids ties**.

The Action Learning Theorem

- Recall that in the revealed actions model: $A_i(t) = \text{argmax}_{a \in \{-1,1\}} \mathbb{P}[a = S | \mathcal{F}_i(t)]$ is announced to the neighbors at time t.
- Technical Assumption: "non-atomic beliefs": The belief $\mathbb{P}[S = 1 \mid \omega_1]$ is non-atomic which **avoids ties**.
- Action Learning Theorem (M., Sly and Tamuz (2012))

The Action Learning Theorem

- Recall that in the revealed actions model: $A_i(t) = \text{argmax}_{a \in \{-1,1\}} \mathbb{P}[a = S | \mathcal{F}_i(t)]$ is announced to the neighbors at time t.
- Technical Assumption: "non-atomic beliefs": The belief $\mathbb{P}[S = 1 \mid \omega_1]$ is non-atomic which **avoids ties**.
- Action Learning Theorem (M., Sly and Tamuz (2012))
- In the revealed actions model with non-atomic beliefs and $F_+ \neq F_-$ there exists a sequence $q(n, F_+, F_-) \to 1$ such for all connected graphs G of size n

$$\mathbb{P}[\text{Learning}] \geq q(n).$$

The Action Learning Theorem

- Recall that in the revealed actions model: $A_i(t) = \text{argmax}_{a \in \{-1,1\}} \mathbb{P}[a = S | \mathcal{F}_i(t)]$ is announced to the neighbors at time t.
- Technical Assumption: "non-atomic beliefs": The belief $\mathbb{P}[S = 1 \mid \omega_1]$ is non-atomic which **avoids ties**.
- Action Learning Theorem (M., Sly and Tamuz (2012))
- In the revealed actions model with non-atomic beliefs and $F_+ \neq F_-$ there exists a sequence $q(n, F_+, F_-) \to 1$ such for all connected graphs G of size n

$$\mathbb{P}[\text{Learning}] \geq q(n).$$

- I.e., asymptotic learning on general graphs.

The Action Learning Theorem

- Recall that in the revealed actions model: $A_i(t) = \text{argmax}_{a \in \{-1,1\}} \mathbb{P}[a = S | \mathcal{F}_i(t)]$ is announced to the neighbors at time t.
- Technical Assumption: "non-atomic beliefs": The belief $\mathbb{P}[S = 1 \mid \omega_1]$ is non-atomic which **avoids ties**.
- Action Learning Theorem (M., Sly and Tamuz (2012))
- In the revealed actions model with non-atomic beliefs and $F_+ \neq F_-$ there exists a sequence $q(n, F_+, F_-) \to 1$ such for all connected graphs G of size n

$$\mathbb{P}[\text{Learning}] \geq q(n).$$

- I.e., asymptotic learning on general graphs.
- False: without non-atomic assumption, on directed graphs, w.o independence.

Prior work - some tractable models

- In the **revealed actions** model when the social network is the **complete graph** then Bayesian updates are **tractable** (Mossel and Tamuz (2010)).
  - Convergence to the same action.
  - As the number of agents goes to infinity the chance of converging to the correct action goes to 1.

Prior work - some tractable models

- In the **revealed actions** model when the social network is the **complete graph** then Bayesian updates are **tractable** (Mossel and Tamuz (2010)).
  - Convergence to the same action.
  - As the number of agents goes to infinity the chance of converging to the correct action goes to 1.
- In the **revealed beliefs** model when private signal are **Gaussian** in the revealed beliefs model then Bayesian updates are tractable for any social network (DeMarzo, D. Vayanos, and J. Zwiebel. 2003 Mossel and Tamuz 2009).
  - Perfectly efficient learning is achieved.
  - Converge in time $\leq$ size $\times$ diameter.

Prior work - some tractable models

- In the **revealed actions** model when the social network is the **complete graph** then Bayesian updates are **tractable** (Mossel and Tamuz (2010)).
  - Convergence to the same action.
  - As the number of agents goes to infinity the chance of converging to the correct action goes to 1.
- In the **revealed beliefs** model when private signal are **Gaussian** in the revealed beliefs model then Bayesian updates are tractable for any social network (DeMarzo, D. Vayanos, and J. Zwiebel. 2003 Mossel and Tamuz 2009).
  - Perfectly efficient learning is achieved.
  - Converge in time $\leq$ size $\times$ diameter.
- When calculations are tractable then they are also easier to analyze.

Learning with the revealed actions model

Action Learning Theorem (M., Sly and Tamuz (2012))
In the revealed actions model with non-atomic beliefs there exists a sequence $q(n) \to 1$ such for all connected graphs G of size n

$$\mathbb{P}\left[\text{Learning}\right] \geq q(n).$$

I.e., asymptotic learning on general graphs.

Learning with the revealed actions model

Action Learning Theorem (M., Sly and Tamuz (2012))

In the revealed actions model with non-atomic beliefs there exists a sequence $q(n) \to 1$ such for all connected graphs G of size n

$$\mathbb{P}\left[\text{Learning}\right] \geq q(n).$$

I.e., asymptotic learning on general graphs.

False in general without non-atomic assumption and on directed graphs, independence is also needed.

Abstract Proof Approach

- Dynamics are very complicated. Abstract approach needed: Assume by contradiction there is a sequence $G_n = (V_n, E_n)$ of graphs with $|V_n| \to \infty$ and $\limsup \mathbb{P}[\text{ Learning in } G_n] < 1$.

Abstract Proof Approach

- Dynamics are very complicated. Abstract approach needed: Assume by contradiction there is a sequence $G_n = (V_n, E_n)$ of graphs with $|V_n| \to \infty$ and $\limsup \mathbb{P}[\text{ Learning in } G_n] < 1$.
- Let $p_i(t) := \mathbb{P}[A_i(t) = S]$ be the "correctness probability".

Abstract Proof Approach

- Dynamics are very complicated. Abstract approach needed: Assume by contradiction there is a sequence $G_n = (V_n, E_n)$ of graphs with $|V_n| \to \infty$ and $\limsup \mathbb{P}[\text{ Learning in } G_n] < 1$.
- Let $p_i(t) := \mathbb{P}[A_i(t) = S]$ be the "correctness probability".
- Increasing: $p_i(t+1) \geq p_i(t)$.

Abstract Proof Approach

- Dynamics are very complicated. Abstract approach needed: Assume by contradiction there is a sequence $G_n = (V_n, E_n)$ of graphs with $|V_n| \to \infty$ and $\limsup \mathbb{P}[\text{ Learning in } G_n] < 1$.
- Let $p_i(t) := \mathbb{P}[A_i(t) = S]$ be the "correctness probability".
- Increasing: $p_i(t+1) \geq p_i(t)$.
- Locally defined: $p_i(t)$ depends only on $B_i(t)$.

Abstract Proof Approach

- Dynamics are very complicated. Abstract approach needed: Assume by contradiction there is a sequence $G_n = (V_n, E_n)$ of graphs with $|V_n| \to \infty$ and $\limsup \mathbb{P}[\text{ Learning in } G_n] < 1$.
- Let $p_i(t) := \mathbb{P}[A_i(t) = S]$ be the "correctness probability".
- Increasing: $p_i(t+1) \geq p_i(t)$.
- Locally defined: $p_i(t)$ depends only on $B_i(t)$.
- Imitation principle: If $i \sim j$ then $p_i(t+1) \geq p_j(t)$.

Abstract Proof Approach

- Dynamics are very complicated. Abstract approach needed: Assume by contradiction there is a sequence $G_n = (V_n, E_n)$ of graphs with $|V_n| \to \infty$ and $\limsup \mathbb{P}[\text{ Learning in } G_n] < 1$.
- Let $p_i(t) := \mathbb{P}[A_i(t) = S]$ be the "correctness probability".
- Increasing: $p_i(t+1) \geq p_i(t)$.
- Locally defined: $p_i(t)$ depends only on $B_i(t)$.
- Imitation principle: If $i \sim j$ then $p_i(t+1) \geq p_j(t)$.
- $\implies$ For G connected there exists $p(G) = sup_t p_i(t)$ which does not depend on i.

Abstract Proof Approach

- Dynamics are very complicated. Abstract approach needed: Assume by contradiction there is a sequence $G_n = (V_n, E_n)$ of graphs with $|V_n| \to \infty$ and $\limsup \mathbb{P}[\text{ Learning in } G_n] < 1$.
- Let $p_i(t) := \mathbb{P}[A_i(t) = S]$ be the "correctness probability".
- Increasing: $p_i(t+1) \geq p_i(t)$.
- Locally defined: $p_i(t)$ depends only on $B_i(t)$.
- Imitation principle: If $i \sim j$ then $p_i(t+1) \geq p_j(t)$.
- $\implies$ For G connected there exists $p(G) = sup_t p_i(t)$ which does not depend on i.
- First round: For some $c > 0$, $p_i(2) \geq 1 - e^{-c\deg(i)}$

Abstract Proof Approach

- Dynamics are very complicated. Abstract approach needed: Assume by contradiction there is a sequence $G_n = (V_n, E_n)$ of graphs with $|V_n| \to \infty$ and $\limsup \mathbb{P}[\text{ Learning in } G_n] < 1$.
- Let $p_i(t) := \mathbb{P}[A_i(t) = S]$ be the "correctness probability".
- Increasing: $p_i(t+1) \geq p_i(t)$.
- Locally defined: $p_i(t)$ depends only on $B_i(t)$.
- Imitation principle: If $i \sim j$ then $p_i(t+1) \geq p_j(t)$.
- $\implies$ For G connected there exists $p(G) = sup_t p_i(t)$ which does not depend on i.
- First round: For some $c > 0$, $p_i(2) \geq 1 - e^{-c\deg(i)}$
- $\implies$ G_n are bounded degree.

Local limits of graphs and Agreement Probabilities

- Definition (local convergence): $(G_n, i_n) \xrightarrow{L} (G, i)$ if for each $t > 0$, for large enough n the neighbourhoods $B_{i_n}^{G_n}(t)$ and $B_i^G(t)$ are isomorphic.

Local limits of graphs and Agreement Probabilities

- Definition (local convergence): $(G_n, i_n) \xrightarrow{L} (G, i)$ if for each $t > 0$, for large enough n the neighbourhoods $B_{i_n}^{G_n}(t)$ and $B_i^G(t)$ are isomorphic.
- Claim: If $(G_n, i_n) \xrightarrow{L} (G, i)$ then

$$p(G) \leq \liminf p(G_n).$$

Local limits of graphs and Agreement Probabilities

- Definition (local convergence): $(G_n, i_n) \xrightarrow{L} (G, i)$ if for each $t > 0$, for large enough n the neighbourhoods $B_{i_n}^{G_n}(t)$ and $B_i^G(t)$ are isomorphic.
- Claim: If $(G_n, i_n) \xrightarrow{L} (G, i)$ then

$$p(G) \leq \liminf p(G_n).$$

- Proof: $p_t(i) = lim_{n\to\infty} p_t(i_n) \leq \liminf p(G_n)$

Local limits of graphs and Agreement Probabilities

- Definition (local convergence): $(G_n, i_n) \xrightarrow{L} (G, i)$ if for each $t > 0$, for large enough n the neighbourhoods $B_{i_n}^{G_n}(t)$ and $B_i^G(t)$ are isomorphic.
- Claim: If $(G_n, i_n) \xrightarrow{L} (G, i)$ then

$$p(G) \leq \liminf p(G_n).$$

- Proof: $p_t(i) = lim_{n\to\infty} p_t(i_n) \leq \liminf p(G_n)$
- Fact: Bounded degree graphs are sequentially compact under local limits.

Local limits of graphs and Agreement Probabilities

- Definition (local convergence): $(G_n, i_n) \xrightarrow{L} (G, i)$ if for each $t > 0$, for large enough n the neighbourhoods $B_{i_n}^{G_n}(t)$ and $B_i^G(t)$ are isomorphic.
- Claim: If $(G_n, i_n) \xrightarrow{L} (G, i)$ then

$$p(G) \leq \liminf p(G_n).$$

- Proof: $p_t(i) = lim_{n \to \infty} p_t(i_n) \leq \liminf p(G_n)$
- Fact: Bounded degree graphs are sequentially compact under local limits.
- Conclusion: To prove the theorem it suffices to show that $p(G) = 1$ for all infinite connected graphs G.

Local limits of graphs and Agreement Probabilities

- Definition (local convergence): $(G_n, i_n) \xrightarrow{L} (G, i)$ if for each $t > 0$, for large enough n the neighbourhoods $B_{i_n}^{G_n}(t)$ and $B_i^G(t)$ are isomorphic.
- Claim: If $(G_n, i_n) \xrightarrow{L} (G, i)$ then

$$p(G) \leq \liminf p(G_n).$$

- Proof: $p_t(i) = lim_{n\to\infty} p_t(i_n) \leq \liminf p(G_n)$
- Fact: Bounded degree graphs are sequentially compact under local limits.
- Conclusion: To prove the theorem it suffices to show that $p(G) = 1$ for all infinite connected graphs G.
- Let

$$p^* = \inf_{G \text{ infinite}} p(G).$$

Limiting Actions and Transitive Graphs

- Recall Claim: Under the non-atomic beliefs three possible limiting actions are possible
  1. For all i, $A_i(t) \to 1$ and $X_i(\infty) > \frac{1}{2}$.
  2. For all i, $A_i(t) \to -1$ and $X_i(\infty) < \frac{1}{2}$.
  3. For all i, $A_i(t)$ does not converge and $X_i(\infty) = \frac{1}{2}$.

Limiting Actions and Transitive Graphs

- Recall Claim: Under the non-atomic beliefs three possible limiting actions are possible
  1. For all i, $A_i(t) \to 1$ and $X_i(\infty) > \frac{1}{2}$.
  2. For all i, $A_i(t) \to -1$ and $X_i(\infty) < \frac{1}{2}$.
  3. For all i, $A_i(t)$ does not converge and $X_i(\infty) = \frac{1}{2}$.
- Let $A^* \in \{1, -1, *\}$ denote the limiting action.

Limiting Actions and Transitive Graphs

- Recall Claim: Under the non-atomic beliefs three possible limiting actions are possible
  1. For all i, $A_i(t) \to 1$ and $X_i(\infty) > \frac{1}{2}$.
  2. For all i, $A_i(t) \to -1$ and $X_i(\infty) < \frac{1}{2}$.
  3. For all i, $A_i(t)$ does not converge and $X_i(\infty) = \frac{1}{2}$.
- Let $A^* \in \{1, -1, *\}$ denote the limiting action.
- Note: On transitive graphs a 0–1 law implies $\mathbb{P}[A^* = S | S = \pm] \in \{0, 1\}$.

Limiting Actions and Transitive Graphs

- Recall Claim: Under the non-atomic beliefs three possible limiting actions are possible
  1. For all i, $A_i(t) \to 1$ and $X_i(\infty) > \frac{1}{2}$.
  2. For all i, $A_i(t) \to -1$ and $X_i(\infty) < \frac{1}{2}$.
  3. For all i, $A_i(t)$ does not converge and $X_i(\infty) = \frac{1}{2}$.
- Let $A^* \in \{1, -1, *\}$ denote the limiting action.
- Note: On transitive graphs a 0–1 law implies $\mathbb{P}[A^* = S | S = \pm] \in \{0, 1\}$.
- But $\mathbb{P}[A^* = S] \geq \mathbb{P}[A_1(t) = S] > 0.5$ so $\mathbb{P}[A^* = S] = 1$.

Limiting Actions and Transitive Graphs

- Recall Claim: Under the non-atomic beliefs three possible limiting actions are possible
  1. For all i, $A_i(t) \to 1$ and $X_i(\infty) > \frac{1}{2}$.
  2. For all i, $A_i(t) \to -1$ and $X_i(\infty) < \frac{1}{2}$.
  3. For all i, $A_i(t)$ does not converge and $X_i(\infty) = \frac{1}{2}$.
- Let $A^* \in \{1, -1, *\}$ denote the limiting action.
- Note: On transitive graphs a 0–1 law implies $\mathbb{P}[A^* = S | S = \pm] \in \{0, 1\}$.
- But $\mathbb{P}[A^* = S] \geq \mathbb{P}[A_1(t) = S] > 0.5$ so $\mathbb{P}[A^* = S] = 1$.
- More work needed for general graphs.

Informative times

- The proof proceeds by using induction to find a vertex i and times $t_1 < \ldots < t_k$ such that $\mathbb{P}[A_i(t_\ell) = S] \approx p^*$ and the $A_i(t_\ell)$ are almost independent.

Informative times

- The proof proceeds by using induction to find a vertex i and times $t_1 < \ldots < t_k$ such that $\mathbb{P}[A_i(t_\ell) = S] \approx p^*$ and the $A_i(t_\ell)$ are almost independent.
- Inductive Hypothesis: On any infinite graph G for any $k, \epsilon > 0$ there exists a vertex i such by some time t there exist $\mathcal{F}_i(t)$ measureable random variables $Y_1, \ldots, Y_k$ taking values in $\{-1, 1\}$ such that
  - For each ℓ, $\mathbb{P}[Y_\ell = S] \geq p^* - \epsilon$.
  - The variables $Y_1, \ldots, Y_k$ are ϵ close to conditionally independent random variables, given S, in total variation distance.

Informative times

- The proof proceeds by using induction to find a vertex i and times $t_1 < \ldots < t_k$ such that $\mathbb{P}[A_i(t_\ell) = S] \approx p^*$ and the $A_i(t_\ell)$ are almost independent.
- Inductive Hypothesis: On any infinite graph G for any $k, \epsilon > 0$ there exists a vertex i such by some time t there exist $\mathcal{F}_i(t)$ measureable random variables $Y_1, \ldots, Y_k$ taking values in $\{-1, 1\}$ such that
  - For each ℓ, $\mathbb{P}[Y_\ell = S] \geq p^* - \epsilon$.
  - The variables $Y_1, \ldots, Y_k$ are ϵ close to conditionally independent random variables, given S, in total variation distance.
- Note that the case $k = 1$ follows from the definition of p^*.

Informative times

- The proof proceeds by using induction to find a vertex i and times $t_1 < \ldots < t_k$ such that $\mathbb{P}[A_i(t_\ell) = S] \approx p^*$ and the $A_i(t_\ell)$ are almost independent.
- Inductive Hypothesis: On any infinite graph G for any $k, \epsilon > 0$ there exists a vertex i such by some time t there exist $\mathcal{F}_i(t)$ measureable random variables $Y_1, \ldots, Y_k$ taking values in $\{-1, 1\}$ such that
  - For each ℓ, $\mathbb{P}[Y_\ell = S] \geq p^* - \epsilon$.
  - The variables $Y_1, \ldots, Y_k$ are ϵ close to conditionally independent random variables, given S, in total variation distance.
- Note that the case $k = 1$ follows from the definition of p^*.
- The induction claim implies the theorem by taking the majority of the Y_ℓ. This identifies S with probability better than p^* unless $p^* = 1$.

Proof Sketch of Main Induction Step

- Note for any i and any $\epsilon' > 0$, there exists t' and an $\mathcal{F}_{t'}$-measurable $\tilde{A}^*$ such that $\mathbb{P}\left[A^* = \tilde{A}^*\right] \geq 1 - \epsilon'$.

Proof Sketch of Main Induction Step

- Note for any i and any $\epsilon' > 0$, there exists t' and an $\mathcal{F}_{t'}$-measurable $\tilde{A}^*$ such that $\mathbb{P}\left[A^* = \tilde{A}^*\right] \geq 1 - \epsilon'$.
- By passing to a subsequence choose $i_1, i_2, \ldots$ so that $d(i, i_\ell) \to \infty$ and $(G, i_\ell) \xrightarrow{L} (H, i^*)$. By induction we can find $j^* \in H$ with k informative times by time t^*.

Proof Sketch of Main Induction Step

- Note for any i and any $\epsilon' > 0$, there exists t' and an $\mathcal{F}_{t'}$-measurable $\tilde{A}^*$ such that $\mathbb{P}\left[A^* = \tilde{A}^*\right] \geq 1 - \epsilon'$.
- By passing to a subsequence choose $i_1, i_2, \ldots$ so that $d(i, i_\ell) \to \infty$ and $(G, i_\ell) \xrightarrow{L} (H, i^*)$. By induction we can find $j^* \in H$ with k informative times by time t^*.
- $\implies \exists j \in G$ with
  - $d(i, j) > t' + t^*$
  - j has k informative times by time t^*

  These are conditionally independent of $\tilde{A}^*$ given S and hence approximately conditionally independent of A^*.

Proof Sketch of Main Induction Step

- Note for any i and any $\epsilon' > 0$, there exists t' and an $\mathcal{F}_{t'}$-measurable $\tilde{A}^*$ such that $\mathbb{P}\left[A^* = \tilde{A}^*\right] \geq 1 - \epsilon'$.
- By passing to a subsequence choose $i_1, i_2, \ldots$ so that $d(i, i_\ell) \to \infty$ and $(G, i_\ell) \xrightarrow{L} (H, i^*)$. By induction we can find $j^* \in H$ with k informative times by time t^*.
- $\implies \exists j \in G$ with
  - $d(i,j) > t' + t^*$
  - j has k informative times by time t^*

  These are conditionally independent of $\tilde{A}^*$ given S and hence approximately conditionally independent of A^*.
- But eventually agent j will learn A^* too giving j another informative time. This completes the induction.

Remarks about the proof

- Finite to infinite principle.

Remarks about the proof

- Finite to infinite principle.
- Like many finite to infinite proofs, no rate.

Agreeing on beliefs implies learning - proof

Belief Learning Theorem (M. Sly and Tamuz (2012))
If there exists a random variable X such that
$X = X_i := \mathbb{E}[S \mid \mathcal{F}_i(\infty)]$ for all i then all agents learned optimally:

$$X = \mathbb{P}[S = 1 \mid \omega_1, \ldots, \omega_n].$$

Proof Sketch

- $$Z_i := \log \frac{\mathbb{P}[S = 1 \mid \omega_i]}{\mathbb{P}[S = 0 \mid \omega_i]} = \log \frac{\mathbb{P}[\omega_i \mid S = 1]}{\mathbb{P}[\omega_i \mid S = 0]}, \quad Z = \sum_i Z_i$$

  so

  $$\mathbb{P}[S = 1 \mid \omega_1, \ldots, \omega_n] = L(Z)$$

  where $L(x) = e^x/(e^x + e^{-x})$.

Proof Sketch

- $$Z_i := \log \frac{\mathbb{P}[S = 1 \mid \omega_i]}{\mathbb{P}[S = 0 \mid \omega_i]} = \log \frac{\mathbb{P}[\omega_i \mid S = 1]}{\mathbb{P}[\omega_i \mid S = 0]}, \quad Z = \sum_i Z_i$$

  so

  $$\mathbb{P}[S = 1 \mid \omega_1, \ldots, \omega_n] = L(Z)$$

  where $L(x) = e^x/(e^x + e^{-x})$.
- Since X is $\mathcal{F}_i$ measurable

  $$X = \mathbb{E}[L(Z) \mid \mathcal{F}_i] = \mathbb{E}[L(Z) \mid X].$$

Proof Sketch

- $$Z_i := \log \frac{\mathbb{P}[S=1 \mid \omega_i]}{\mathbb{P}[S=0 \mid \omega_i]} = \log \frac{\mathbb{P}[\omega_i \mid S=1]}{\mathbb{P}[\omega_i \mid S=0]}, \quad Z = \sum_i Z_i$$

  so

  $$\mathbb{P}[S=1 \mid \omega_1, \ldots, \omega_n] = L(Z)$$

  where $L(x) = e^x/(e^x + e^{-x})$.
- Since X is $\mathcal{F}_i$ measurable

  $$X = \mathbb{E}[L(Z) \mid \mathcal{F}_i] = \mathbb{E}[L(Z) \mid X].$$

- Hence since Z_i is $\mathcal{F}_i$ measurable

  $$\begin{aligned} \mathbb{E}[Z_i \cdot L(Z) \mid X] &= \mathbb{E}[\mathbb{E}[Z_i \cdot L(Z) \mid \mathcal{F}_i] \mid X] \\ &= \mathbb{E}[Z_i \cdot X \mid X] \\ &= \mathbb{E}[Z_i \mid X]\,\mathbb{E}[L(Z) \mid X] \end{aligned}$$

Proof concluded

$$\mathbb{E}\left[Z_i \cdot L(Z) \mid X\right] = \mathbb{E}\left[Z_i \mid X\right] \mathbb{E}\left[L(Z) \mid X\right]$$

- Summing over i we get that

$$\mathbb{E}\left[Z \cdot L(Z) \mid X\right] = \mathbb{E}\left[Z \mid X\right] \mathbb{E}\left[L(Z) \mid X\right]$$

Proof concluded

$$\mathbb{E}\left[Z_i \cdot L(Z) \mid X\right] = \mathbb{E}\left[Z_i \mid X\right]\mathbb{E}\left[L(Z) \mid X\right]$$

- Summing over i we get that

$$\mathbb{E}\left[Z \cdot L(Z) \mid X\right] = \mathbb{E}\left[Z \mid X\right]\mathbb{E}\left[L(Z) \mid X\right]$$

- Since $L(x)$ is strictly increasing this implies that Z is constant conditional on X, i.e. Z is X measurable so

$$X = \mathbb{E}\left[L(Z) \mid X\right] = \mathbb{E}\left[L(Z) \mid Z\right] = L(Z)$$

Proof concluded

$$\mathbb{E}\left[Z_i \cdot L(Z) \mid X\right] = \mathbb{E}\left[Z_i \mid X\right] \mathbb{E}\left[L(Z) \mid X\right]$$

- Summing over i we get that

$$\mathbb{E}\left[Z \cdot L(Z) \mid X\right] = \mathbb{E}\left[Z \mid X\right] \mathbb{E}\left[L(Z) \mid X\right]$$

- Since $L(x)$ is strictly increasing this implies that Z is constant conditional on X, i.e. Z is X measurable so

$$X = \mathbb{E}\left[L(Z) \mid X\right] = \mathbb{E}\left[L(Z) \mid Z\right] = L(Z)$$

- So the agreed value X equal to the optimal estimator $L(Z)$ as needed.

Open Problems

- Both models: Rate of convergence? Dependence on graph?

Open Problems

- Both models: Rate of convergence? Dependence on graph?
- Actions Learning: how does the probability of learning change with the graph?

Open Problems

- Both models: Rate of convergence? Dependence on graph?
- Actions Learning: how does the probability of learning change with the graph?
- What if the agents aren't truthful but act strategically (in a game theoretic manner)?

Thanks for listening

Questions?