Common Learning

Martin W. Cripps, Jeffrey C. Ely, George J. Mailath, Larry Samuelson

Econometrica 2008

September 23, 2008

A Model of Multi-Agent Learning

- 2 agents $\ell = 1, 2$, discrete time $t = 0, 1, 2 \ldots$.
- Before period zero, nature selects a parameter θ from the finite set Θ according to the prior distribution p.
- Conditional on θ, a stochastic process $\zeta^\theta \equiv \{\zeta_t^\theta\}_{t=0}^\infty \equiv \{\zeta_{1t}^\theta, \zeta_{2t}^\theta\}_{t=0}^\infty$ generates a signal profile $z_t \equiv (z_{1t}, z_{2t}) \in Z_1 \times Z_2$ for each period t.
- A state $\omega \in \Omega \equiv \Theta \times Z^\infty$ consists of a parameter and a sequence of signal profiles.
- Let P denote the measure defined on Ω and induced by the prior p and the signal processes $\left(\zeta^\theta\right)_{\theta \in \Theta}$.
- A period-t history of agent ℓ is denoted by $h_{\ell t} \equiv (z_{\ell 0}, z_{\ell 1}, \ldots, z_{\ell t-1})$. Let $H_{\ell t} \equiv (Z_\ell)^t$ be the space of period-t histories for agent ℓ, and $\{\mathcal{H}_{\ell t}\}_{t=0}^\infty$ denote the filtration induced on Ω by agent ℓ's histories.
- Finally, the random variables $\{P(\theta | \mathcal{H}_{\ell t})\}_{t=0}^\infty$ give agent ℓ's beliefs about the parameter θ at the start of each period.

Individual Learning

Let P^{θ} denote the measure conditional on a given parameter and define

$$B^{q}_{\ell t}(F) \equiv \{\omega \in \Omega : E[1_F | \mathcal{H}_{\ell t}](\omega) \geq q\}$$

Definition 1. Agent ℓ learns parameter θ if conditional on parameter θ, agent ℓ's posterior on θ converges in probability to 1, i.e.

$$\forall q \in (0,1) \quad \exists T \quad \text{s.t.} \quad \forall t > T : \quad P^{\theta}\left(B^{q}_{\ell t}(\theta)\right) > q$$

which is equivalent to requiring that

$$\forall q \in (0,1) : \quad \lim_{t \to \infty} P^{\theta}\left(B^{q}_{\ell t}(\theta)\right) = 1$$

Note: knowledge versus approximate knowledge.

Example of Individual Learning

$\Theta = \{0,1\}$, $Z_\ell = \{0,1,2,3,\ldots\}$ for $\ell = 1,2$.
Assume that $\theta = 1$ occurs w/prob p, while $\theta = 0$ w/prob $1-p$.

P	z_1	z_2
θ	0	0
$\varepsilon(1-\theta)$	1	0
$(1-\varepsilon)\varepsilon(1-\theta)$	1	1
$(1-\varepsilon)^2\varepsilon(1-\theta)$	2	1
$(1-\varepsilon)^3\varepsilon(1-\theta)$	2	2
$(1-\varepsilon)^4\varepsilon(1-\theta)$	3	2
$(1-\varepsilon)^5\varepsilon(1-\theta)$	3	3
.	.	.

Example of Individual Learning

$\Theta = \{0, 1\}$, $Z_\ell = \{0, 1, 2, 3, ...\}$ for $\ell = 1, 2$.
Assume that $\theta = 1$ occurs w/prob p, while $\theta = 0$ w/prob $1 - p$.

P	z_1	z_2
θ	0	0
$\varepsilon(1-\theta)$	1	0
$(1-\varepsilon)\varepsilon(1-\theta)$	1	1
$(1-\varepsilon)^2\varepsilon(1-\theta)$	2	1
$(1-\varepsilon)^3\varepsilon(1-\theta)$	2	2
$(1-\varepsilon)^4\varepsilon(1-\theta)$	3	2
$(1-\varepsilon)^5\varepsilon(1-\theta)$	3	3
.	.	.

$\theta=0$

t	z_{1t}	z_{2t}
1	57	0
2	14	53
3	3	0
4	1	2
5	172	172

$\theta=1$

t	z_{1t}	z_{2t}
1	0	0
2	0	0
3	0	0
4	0	0
5	0	0

P	z_1	z_2
θ	0	0
$\varepsilon(1-\theta)$	1	0
$(1-\varepsilon)\varepsilon(1-\theta)$	1	1
$(1-\varepsilon)^2\varepsilon(1-\theta)$	2	1
$(1-\varepsilon)^3\varepsilon(1-\theta)$	2	2
$(1-\varepsilon)^4\varepsilon(1-\theta)$	3	2
$(1-\varepsilon)^5\varepsilon(1-\theta)$	3	3
.	.	.

$B_{2t}^{q}(\theta = 0)$=?

P	z_1	z_2
θ	0	0
$\varepsilon(1-\theta)$	1	0
$(1-\varepsilon)\varepsilon(1-\theta)$	1	1
$(1-\varepsilon)^2\varepsilon(1-\theta)$	2	1
$(1-\varepsilon)^3\varepsilon(1-\theta)$	2	2
$(1-\varepsilon)^4\varepsilon(1-\theta)$	3	2
$(1-\varepsilon)^5\varepsilon(1-\theta)$	3	3
.	.	.

$B^q_{2t}(\theta = 0)$=?

Let $h_{21} = z_{21} = 0$. Then, $P(\theta = 0|z_{21} = 0) = \frac{\varepsilon(1-p)}{p+\varepsilon(1-p)}$ and $\{\omega : \theta = 0, z_{21} = 0\} \subset B^q_{21}(\theta = 0)$ iff $\frac{\varepsilon(1-p)}{p+\varepsilon(1-p)} \geq q$.

P	z_1	z_2
θ	0	0
$\varepsilon(1-\theta)$	1	0
$(1-\varepsilon)\varepsilon(1-\theta)$	1	1
$(1-\varepsilon)^2\varepsilon(1-\theta)$	2	1
$(1-\varepsilon)^3\varepsilon(1-\theta)$	2	2
$(1-\varepsilon)^4\varepsilon(1-\theta)$	3	2
$(1-\varepsilon)^5\varepsilon(1-\theta)$	3	3
.	.	.

$B_{2t}^q(\theta=0)$=?

Let $h_{21}=z_{21}=0$. Then, $P(\theta=0|z_{21}=0)=\frac{\varepsilon(1-p)}{p+\varepsilon(1-p)}$ and $\{\omega:\theta=0,z_{21}=0\}\subset B_{21}^q(\theta=0)$ iff $\frac{\varepsilon(1-p)}{p+\varepsilon(1-p)}\geq q$.

Let $h_{21}=z_{21}=1$. Then, trivially, $P(\theta=0|z_{21}=1)=1$ and $\{\omega:\theta=0,z_{21}=1\}\subset B_{21}^q(\theta=0)$ for $\forall q\in(0,1)$. Same result applies for all $z_{21}\geq 1$.

Then, $B_{21}^q(\theta=0)=\Omega^{\theta=0}\setminus\{\omega:\theta=0,z_{21}=0\}$ for $q>\frac{\varepsilon(1-p)}{p+\varepsilon(1-p)}$ and $P^{\theta=0}\left(B_{21}^q(\theta=0)\right)=1-\varepsilon$.

P	z_1	z_2
θ	0	0
$\varepsilon(1-\theta)$	1	0
$(1-\varepsilon)\varepsilon(1-\theta)$	1	1
$(1-\varepsilon)^2\varepsilon(1-\theta)$	2	1
$(1-\varepsilon)^3\varepsilon(1-\theta)$	2	2
$(1-\varepsilon)^4\varepsilon(1-\theta)$	3	2
$(1-\varepsilon)^5\varepsilon(1-\theta)$	3	3
.	.	.

$B_{2t}^q(\theta = 0)$=?
Let $h_{21} = z_{21} = 0$. Then, $P(\theta = 0|z_{21} = 0) = \frac{\varepsilon(1-p)}{p+\varepsilon(1-p)}$ and
$\{\omega : \theta = 0, z_{21} = 0\} \subset B_{21}^q(\theta = 0)$ iff $\frac{\varepsilon(1-p)}{p+\varepsilon(1-p)} \geq q$.
Let $h_{21} = z_{21} = 1$. Then, trivially, $P(\theta = 0|z_{21} = 1) = 1$ and
$\{\omega : \theta = 0, z_{21} = 1\} \subset B_{21}^q(\theta = 0)$ for $\forall q \in (0, 1)$. Same result applies for all $z_{21} \geq 1$.
Then, $B_{21}^q(\theta = 0) = \Omega^{\theta=0} \setminus \{\omega : \theta = 0, z_{21} = 0\}$ for $q > \frac{\varepsilon(1-p)}{p+\varepsilon(1-p)}$ and
$P^{\theta=0}(B_{21}^q(\theta = 0)) = 1 - \varepsilon$.
What about $B_{22}^q(\theta = 0)$? $P(\theta = 0|z_{21} = 0, z_{22} = 1) = 1$ while
$P(\theta = 0|z_{21} = 0, z_{22} = 0) = \frac{\varepsilon^2(1-p)}{p+\varepsilon^2(1-p)}$. So
$B_{22}^q(\theta = 0) = \Omega^{\theta=0} \setminus \{\omega : \theta = 0, z_{21} = 0, z_{22} = 0\}$ and $P^{\theta=0}(B_{22}^q(\theta = 0)) = 1 - \varepsilon^2$.

P	z_1	z_2
θ	0	0
$\varepsilon(1-\theta)$	1	0
$(1-\varepsilon)\varepsilon(1-\theta)$	1	1
$(1-\varepsilon)^2\varepsilon(1-\theta)$	2	1
$(1-\varepsilon)^3\varepsilon(1-\theta)$	2	2
$(1-\varepsilon)^4\varepsilon(1-\theta)$	3	2
$(1-\varepsilon)^5\varepsilon(1-\theta)$	3	3
.	.	.

$B_{2t}^{q}(\theta=0)$=?
Let $h_{21}=z_{21}=0$. Then, $P(\theta=0|z_{21}=0)=\frac{\varepsilon(1-p)}{p+\varepsilon(1-p)}$ and $\{\omega:\theta=0,z_{21}=0\}\subset B_{21}^{q}(\theta=0)$ iff $\frac{\varepsilon(1-p)}{p+\varepsilon(1-p)}\geq q$.
Let $h_{21}=z_{21}=1$. Then, trivially, $P(\theta=0|z_{21}=1)=1$ and $\{\omega:\theta=0,z_{21}=1\}\subset B_{21}^{q}(\theta=0)$ for $\forall q\in(0,1)$. Same result applies for all $z_{21}\geq 1$.
Then, $B_{21}^{q}(\theta=0)=\Omega^{\theta=0}\backslash\{\omega:\theta=0,z_{21}=0\}$ for $q>\frac{\varepsilon(1-p)}{p+\varepsilon(1-p)}$ and $P^{\theta=0}\left(B_{21}^{q}(\theta=0)\right)=1-\varepsilon$.
What about $B_{22}^{q}(\theta=0)$? $P(\theta=0|z_{21}=0,z_{22}=1)=1$ while $P(\theta=0|z_{21}=0,z_{22}=0)=\frac{\varepsilon^2(1-p)}{p+\varepsilon^2(1-p)}$. So $B_{22}^{q}(\theta=0)=\Omega^{\theta=0}\backslash\{\omega:\theta=0,z_{21}=0,z_{22}=0\}$ and $P^{\theta=0}\left(B_{22}^{q}(\theta=0)\right)=1-\varepsilon^2$.
It turns out that $\lim_{t\to\infty}P^{\theta=0}\left(B_{2t}^{q}(\theta=0)\right)=\lim_{t\to\infty}\left(1-\varepsilon^t\right)=1$.

Common Learning

The event that $F \subset \Omega$ is $q-$**believed** at time t occurs if each agents attaches at least probability q to F, that is

$$B_t^q(F) \equiv B_{1t}^q(F) \cap B_{2t}^q(F)$$

The event that F is common $q-$**belief** at time t is defined as

$$C_t^q \equiv \bigcap_{n \geq 1} \left[B_t^q\right]^n(F) = B_t^q(F) \cap B_t^q B_t^q(F) \cap \ldots$$

Definition 2. The agents **commonly learn** parameter θ if

$$\forall q \in (0,1) \quad \exists T \quad \text{s.t.} \quad \forall t > T: \quad P^\theta\left(C_t^q(\theta)\right) > q$$

which is equivalent to requiring that

$$\forall q \in (0,1): \quad \lim_{t \to \infty} P^\theta\left(C_t^q(\theta)\right) = 1$$

Since $C_t^q(F) \subset B_{\ell t}^q(F)$ common learning implies the individual learning.

$\Rightarrow$**Question**: When does individual learning imply common learning?

The Importance of Common Learning

	A	B	W
A	$1, 1$	$-c, -c$	$-c, 0$
B	$-c, -c$	$-c, -c$	$-c, 0$
W	$0, -c$	$0, -c$	$0, 0$

Parameter θ_A

	A	B	W
A	$-c, -c$	$-c, -c$	$-c, 0$
B	$-c, -c$	$1, 1$	$-c, 0$
W	$0, -c$	$0, -c$	$0, 0$

Parameter θ_B

Choosing action A is optimal for an agent in some period t only if the agent attaches probability at least $\frac{c}{c+1} \equiv q$ to the joint event that the parameter is θ_A and the other agent beliefs it is A, and he thinks that the other thinks...

Special Cases: Perfect Correlation and Independence

- If signals are **public** then $P(\theta|\mathcal{H}_{1t}) = P(\theta|\mathcal{H}_{2t})$ for all θ and t.
  $\Rightarrow$individual learning with public signals implies common learning.

- If signals are **independent** then

  $$P^{\theta}\left(B_t^q(\theta)\right) \equiv P^{\theta}\left(B_{1t}^q(\theta) \cap B_{2t}^q(\theta)\right) = P^{\theta}\left(B_{1t}^q(\theta)\right) P^{\theta}\left(B_{2t}^q(\theta)\right)$$

  and if each agent is learning individually then we can find T large enough so that $P^{\theta}\left(B_{1t}^{q'}(\theta)\right) > q'$ and $P^{\theta}\left(B_{2t}^{q'}(\theta)\right) > q'$). Set $q' = \sqrt{q}$. Then $B_t^{\sqrt{q}}(\theta) \subset B_t^q(\theta)$ since $q \in (0,1)$ and we have

  $$\begin{aligned} P^{\theta}\left(B_t^q(\theta)\right) &= P^{\theta}\left(B_{1t}^q(\theta)\right) P^{\theta}\left(B_{2t}^q(\theta)\right) \\ &> P^{\theta}\left(B_{1t}^{\sqrt{q}}(\theta)\right) P^{\theta}\left(B_{2t}^{\sqrt{q}}(\theta)\right) > q \quad \forall q \in (0,1) \end{aligned}$$

  $\Rightarrow$individual learning with independent signals implies common learning.

Main Result

Proposition When each agents' signal space is finite agents commonly learn the value of the parameter. That is, the value of the parameter becomes approximate common knowledge.

Example: Failure of Common Learning

P	z_1	z_2
θ	0	0
$\varepsilon(1-\theta)$	1	0
$(1-\varepsilon)\varepsilon(1-\theta)$	1	1
$(1-\varepsilon)^2\varepsilon(1-\theta)$	2	1
$(1-\varepsilon)^3\varepsilon(1-\theta)$	2	2
$(1-\varepsilon)^4\varepsilon(1-\theta)$	3	2
$(1-\varepsilon)^5\varepsilon(1-\theta)$	3	3
.	.	.

Note that $P^{\theta=0}(z_{1t}=k+1|z_{2t}=k)=\frac{1-\varepsilon}{2-\varepsilon}$ for $k>0$ and that $P^{\theta=0}(z_{2t}=k|z_{1t}=k)=\frac{1-\varepsilon}{2-\varepsilon}$ for $k\geq 0$.
Take $k=1$ and apply it to the second expression.

Robustness of the Result with Respect to the Common Prior Assumption

- The result still holds if the two agents have different priors provided that there is a commonly known bound on the minimum probability any parameter receives in each agent's prior.
- The model also captures settings in which agents have different beliefs about the conditional signal-generating distributions. In this case the analysis can be applied to a reformulated model in which agents are uncertain about the join parameter (θ, ρ^{θ}).