Empirical-evidence Equilibria in Stochastic Games

Nicolas Dudebout

Georgia Institute of Technology

Empirical-evidence Equilibria (EEEs)

At Nash equilibrium in a stochastic game, each agent is playing an optimal strategy for a POMDP

Empirical-evidence Equilibria (EEEs)

At Nash equilibrium in a stochastic game, each agent is playing an optimal strategy for a POMDP

EEE approach:

0. Pick arbitrary strategies
1. Formulate simple but consistent models
2. Design strategies optimal w.r.t. models, then, back to 1.

The fixed points are EEEs

Empirical-evidence Equilibria (EEEs)

At Nash equilibrium in a stochastic game, each agent is playing an optimal strategy for a POMDP

EEE approach:

0. Pick arbitrary strategies
1. Formulate simple but consistent models
2. Design strategies optimal w.r.t. models, then, back to 1.

The fixed points are EEEs

Example

Asset management on the stock market

Empirical-evidence Equilibria (EEEs)

At Nash equilibrium in a stochastic game, each agent is playing an optimal strategy for a POMDP

EEE approach:

0. Pick arbitrary strategies
1. Formulate simple but consistent models
2. Design strategies optimal w.r.t. models, then, back to 1.

The fixed points are EEEs

Example

Asset management on the stock market

Single-agent Setup

Single-agent Setup

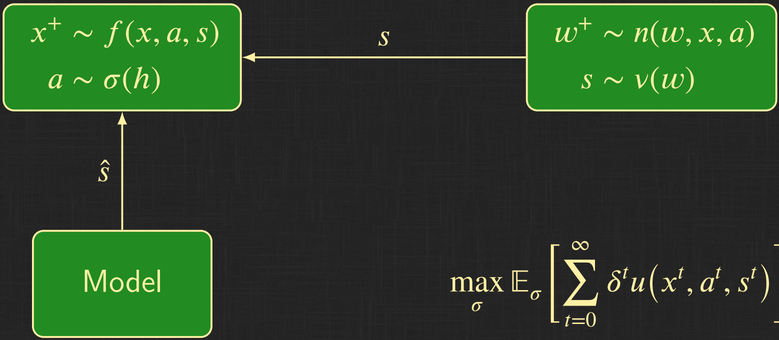
$$\max_{\sigma} \mathbb{E}_{\sigma}\left[\sum_{t=0}^{\infty} \delta^t u\left(x^t, a^t, s^t\right)\right]$$

Single-agent Setup

$$\begin{tikzcd} \boxed{\begin{array}{c} x^+ \sim f(x, a, s) \\ a \sim \sigma(h) \end{array}} & \boxed{\begin{array}{c} w^+ \sim n(w, x, a) \\ s \sim \nu(w) \end{array}} \arrow[l, "s"'] \end{tikzcd}$$

$$\max_{\sigma} \mathbb{E}_{\sigma} \left[\sum_{t=0}^{\infty} \delta^t u\left(x^t, a^t, s^t\right) \right]$$

Single-agent Setup



$$\max_{\sigma} \mathbb{E}_{\sigma}\left[\sum_{t=0}^{\infty} \delta^t u\left(x^t, a^t, s^t\right)\right]$$

Single-agent Setup

$$\begin{tikzcd}
\boxed{\begin{array}{c} x^+ \sim f(x, a, s) \\ a \sim \sigma(x, z) \end{array}} & \boxed{\begin{array}{c} w^+ \sim n(w, x, a) \\ s \sim \nu(w) \end{array}} \arrow[l, "s"'] \\
\boxed{\begin{array}{c} z^+ \sim m(z) \\ \hat{s}^+ \sim \mu(z) \end{array}} \arrow[u, "\hat{s}"] &
\end{tikzcd}$$

$$\max_{\sigma} \mathbb{E}_{\sigma} \left[\sum_{t=0}^{\infty} \delta^t u\left(x^t, a^t, s^t\right) \right]$$

Single-agent Setup

$$\begin{tikzcd}
\begin{array}{c} x^+ \sim f(x,a,s) \\ a \sim \sigma(x,z) \end{array} & \begin{array}{c} w^+ \sim n(w,x,a) \\ s \sim \nu(w) \end{array} \arrow[l, "s"'] \\
\begin{array}{c} z^+ \sim m(z) \\ \hat{s}^+ \sim \mu(z) \end{array} \arrow[u, "\hat{s}"] &
\end{tikzcd}$$

$$\max_{\sigma} \mathbb{E}_{\sigma}\left[\sum_{t=0}^{\infty} \delta^t u\left(x^t, a^t, s^t\right)\right]$$

- μ consistent with σ
- σ optimal w.r.t. μ

Depth-k Consistency

Binary stochastic process s
0100010001001010010110111010000111010101...

- 0 characteristic: $\mathbb{P}[s=0], \mathbb{P}[s=1]$
- 1 characteristic: $\mathbb{P}\left[ss^+=00\right], \mathbb{P}\left[ss^+=10\right],$ $\mathbb{P}\left[ss^+=01\right], \mathbb{P}\left[ss^+=11\right]$
- ...
- k characteristic: probability of strings of length $k+1$

Definition

Two processes s and $\hat{s}$ are depth-k consistent if they have the same k characteristic

Complete Picture

$$
\begin{tikzcd}
\boxed{\begin{array}{c} x^+ \sim f(x, a, s) \\ a \sim \sigma(x, z) \end{array}} & \boxed{\begin{array}{c} w^+ \sim n(w, x, a) \\ s \sim \nu(w) \end{array}} \arrow[l, "s"'] \\
\boxed{\begin{array}{c} z^+ \sim m_k(z) \\ \hat{s} \sim \mu(z) \end{array}} \arrow[u, "\hat{s}"] &
\end{tikzcd}
$$

Fix a depth $k \in \mathbb{N}$

z contains the last k observed signals

$$
\mu\big(z = (s_1, s_2, \dots, s_k)\big)[s_{k+1}] = \mathbb{P}_\sigma\left[s^{t+1} = s_{k+1} \,\middle|\, s^t = s_k, \dots, s^{t-k+1} = s_1\right]
$$

Empirical-evidence Optimality

Definition

(σ, μ) is an empirical-evidence optimum (EEO) for k iff

- σ is optimal w.r.t. μ
- μ is depth-k consistent with σ

Empirical-evidence Optimality

Definition

(σ, μ) is an empirical-evidence optimum (EEO) for k iff

- σ is optimal w.r.t. μ
- μ is depth-k consistent with σ

Definition

(σ, μ) is an ϵ empirical-evidence optimum (ϵ EEO) for k iff

- σ is ϵ optimal w.r.t. μ
- μ is depth-k consistent with σ

Existence Result

Theorem

For all k and ϵ, there exists an ϵ EEO for k

Existence Result

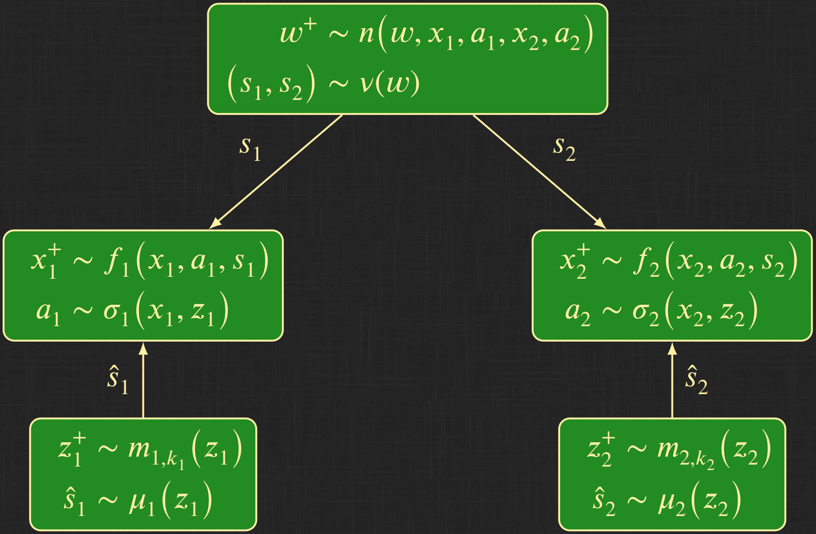
Theorem

For all k and ϵ, there exists an ϵ EEO for k

Proof sketch

- Technical assumption insures ergodicity of s
- $T : \sigma \xmapsto{consistency} \mu \xmapsto{\epsilon\ optimality} \sigma$ is continuous
- $\sigma : \mathscr{X} \times \mathscr{Z} \to \Delta(\mathscr{A})$ is parametrized over a simplex
- Apply Brouwer's fixed point theorem to T

Multiagent Setup



Empirical-evidence Equilibrium

$$\text{Strategies}\ \ \sigma = \left(\sigma_1, \sigma_2, \ldots, \sigma_N\right)$$
$$\text{Models}\ \ \mu = \left(\mu_1, \mu_2, \ldots, \mu_N\right)$$
$$\text{Depths}\ \ k = \left(k_1, k_2, \ldots, k_N\right)$$

Definition

(σ, μ) is an empirical-evidence equilibrium (EEE) for k iff

- for all i, σ_i is optimal w.r.t. μ_i
- for all i, μ_i is depth-k_i consistent with σ

Theorem

For all k and ϵ, there exists an ϵ EEE for k

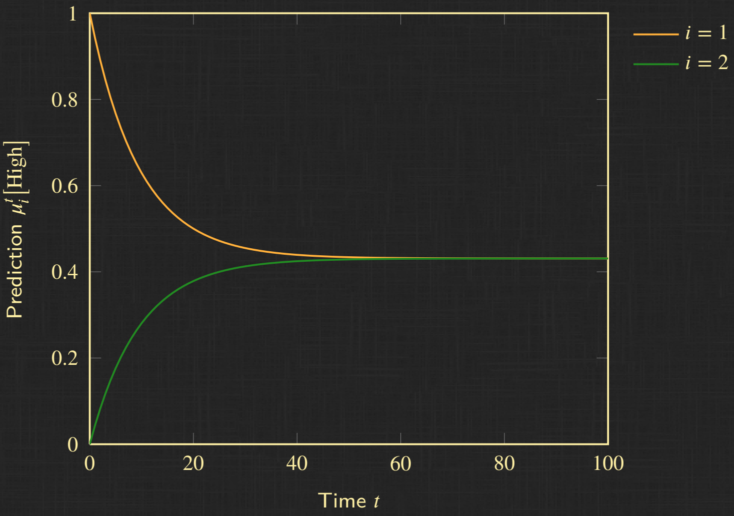
Learning Setup

State	holdings $x_i \in \{0 .. M\}$
Action	sell one, hold, or buy one $a_i \in \{-1, 0, 1\}$
Signal	price $p \in \{\text{Low}, \text{High}\}$
Dynamic	$x_i^+ = x_i + a_i$
Stage cost	$p \cdot a_i$
Nature	market trend $b \in \{\text{Bull}, \text{Bear}\}$
	$w = (b, p)$

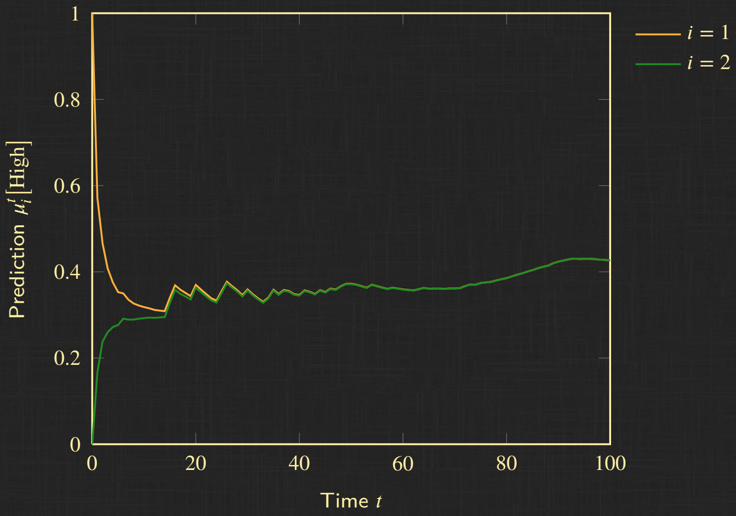
0. Pick arbitrary depth-0 models μ
1. Design strategies σ optimal w.r.t. models μ
2. Formulate consistent models μ_{upd}, then, back to 1.

$$\mu_i^{t+1} = (1-\alpha)\mu_i^t + \alpha\left(\mu_{i,\text{upd}}^t - \mu_i^t\right)$$

Learning Results: Offline



Learning Results: Online



Concluding Remarks

Comparison with mean-field equilibria

- Identical agents with a specific signal
- Depth-0 model
- Large number of agents to recover Nash equilibrium

Concluding Remarks

Comparison with mean-field equilibria

- Identical agents with a specific signal
- Depth-0 model
- Large number of agents to recover Nash equilibrium

Future directions

- Endogenous model $\left(z^+ \sim m(z, x, a)\right)$
- Quality of EEEs
- Learning EEEs