Empirical-evidence Equilibria in Stochastic Games

Nicolas Dudebout

Context

Multiagent problems

- stock market
- group of robots

Context

Multiagent problems

- stock market
- group of robots

- predictive
- prescriptive

Context

Multiagent problems

- stock market
- group of robots

- predictive
- prescriptive

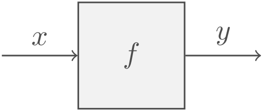
Game-theoretic approach

- selfish agents
- different solution concepts

Empirical-evidence Equilibrium (EEE)

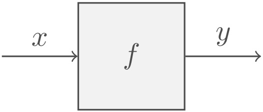
Motivation
Definition
Existence
Comparison
Characterization
Predictive Use
Prescriptive Use

Graphical convention

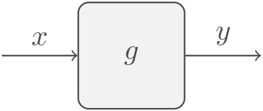


$$y^{+} \sim f(x)$$

Graphical convention

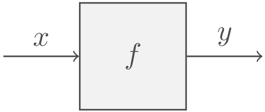


$$y^+ \sim f(x)$$

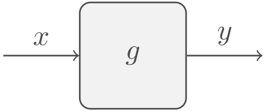


$$y^{t+1} \sim g\big(x^1, \ldots, x^t, y^1, \ldots, y^t\big)$$

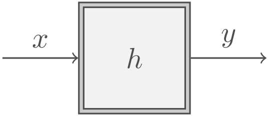
Graphical convention



$$y^+ \sim f(x)$$



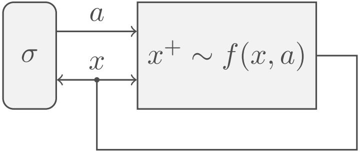
$$y^{t+1} \sim g\big(x^1, \ldots, x^t, y^1, \ldots, y^t\big)$$



$$y^+ \sim h(y, x)$$

Markov Decision Process (MDP)

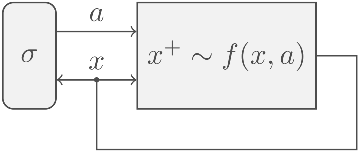
$$\max_{\sigma} \mathbb{E}_{\sigma}\left[\sum_{t=0}^{\infty} \delta^t \cdot u\left(x^t, a^t\right)\right]$$

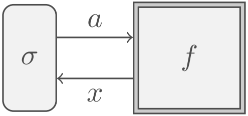


Markov Decision Process (MDP)

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

$$u(x, a), \delta$$

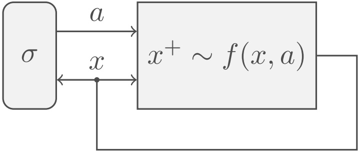


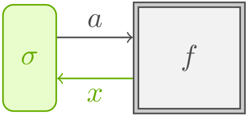


Markov Decision Process (MDP)

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

$$u(x, a), \delta$$

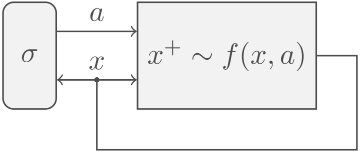


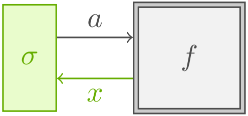


Markov Decision Process (MDP)

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

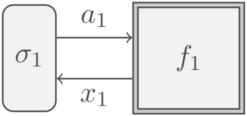
$$u(x, a), \delta$$

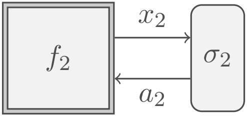




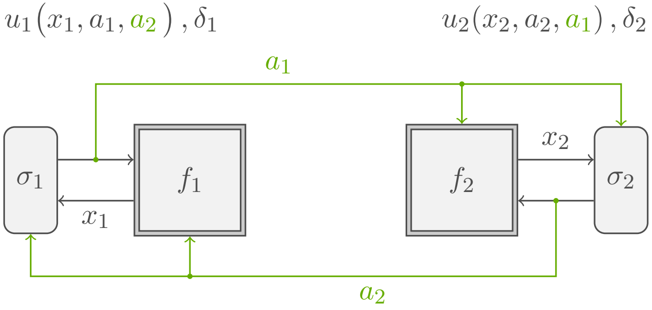
Stochastic Game

$$u_1(x_1, a_1 \qquad), \delta_1 \qquad\qquad u_2(x_2, a_2 \qquad), \delta_2$$



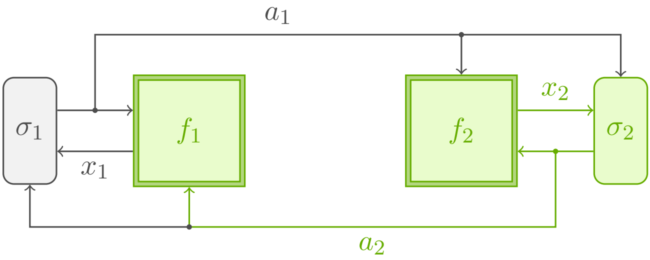


Stochastic Game



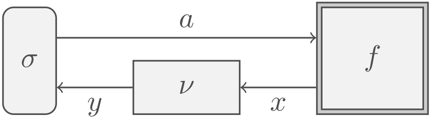
Stochastic Game

$u_1(x_1, a_1, a_2)\ , \delta_1$



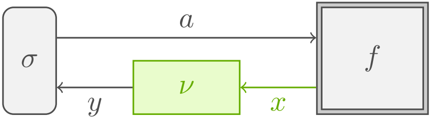
Partially Observable Markov Decision Process (POMDP)

$u(x, a), \delta$

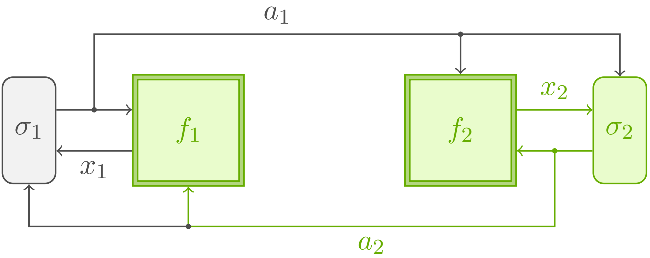


Partially Observable Markov Decision Process (POMDP)

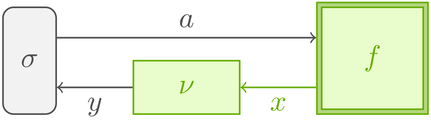
$u(x, a), \delta$



$u_1(x_1, a_1, a_2), \delta_1$



$u(x, a), \delta$



Recap

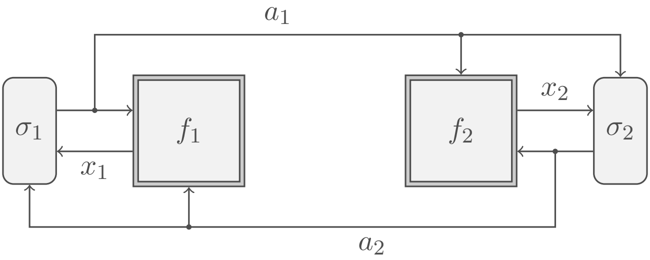
- Multiagent problems
- Game-theoretic approach
- Nash equilibrium in stochastic game $\iff$ unknown POMDPs

- Multiagent problems
- Game-theoretic approach
- Nash equilibrium in stochastic game $\iff$ unknown POMDPs

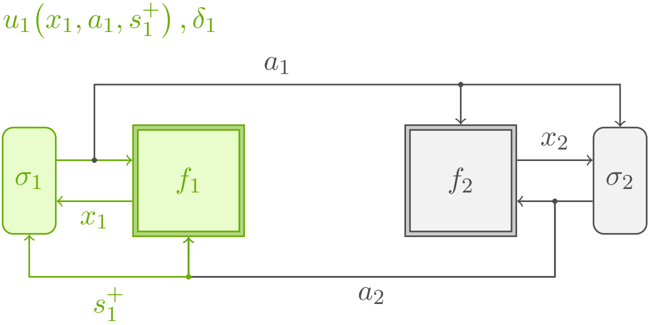
POMDP	intractable
MDP	solved

Stochastic Game

$u_1(x_1, a_1, a_2)\ , \delta_1$

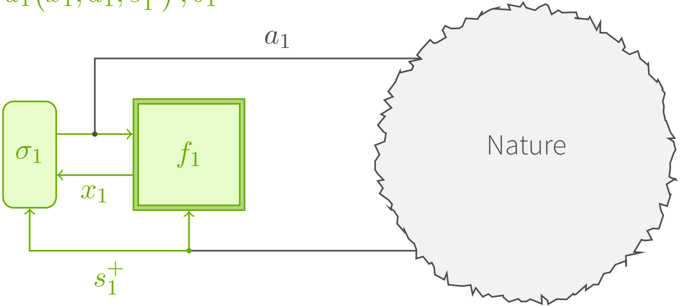


Stochastic Game



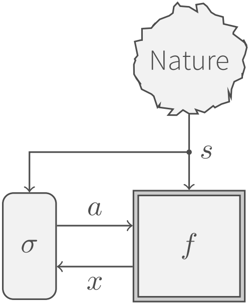
Stochastic Game

$u_1\left(x_1, a_1, s_1^+\right), \delta_1$



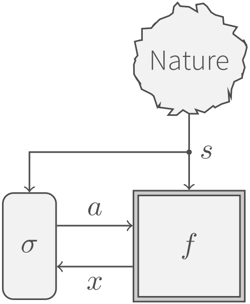
Nature

$u(x, a, s^{+}), \delta$



Nature

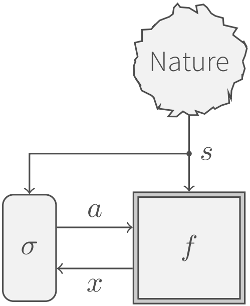
$u(x, a, s^+), \delta$



$$\begin{cases} w^+ \sim n(w, x, a) \\ s^+ \sim \nu(w^+) \end{cases}$$

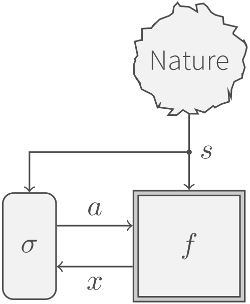
Nature

$u(x, a, s^+), \delta$



Nature

$u(x, a, s^+), \delta$



01000100010010100101101111010...

Simple Consistency

0100010001001010010110111010...

Simple Consistency

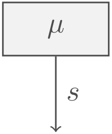
0100010001001010010110111010...

$\mathbb{P}[0], \mathbb{P}[1]$

Simple Consistency

0100010001001010010110111010...

$\mathbb{P}[0], \mathbb{P}[1]$



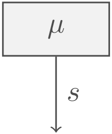
$$\mu[s] = \mathbb{P}[s]$$

Simple Consistency

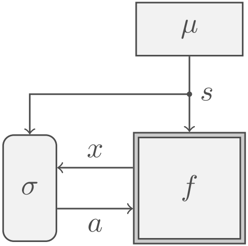
01000100010010100101101110 10...

$u(x, a, s^+), \delta$

$\mathbb{P}[0], \mathbb{P}[1]$



$\mu[s] = \mathbb{P}[s]$

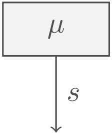


Simple Consistency

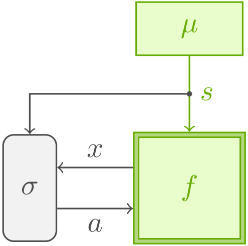
0100010001001010010110111010...

$u(x, a, s^+), \delta$

$\mathbb{P}[0], \mathbb{P}[1]$



$\mu[s] = \mathbb{P}[s]$

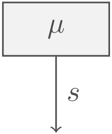


Simple Consistency

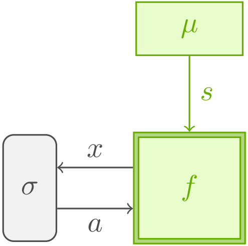
010001000100101001011011101 0...

$u(x, a, s^+), \delta$

$\mathbb{P}[0], \mathbb{P}[1]$



$\mu[s] = \mathbb{P}[s]$

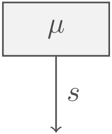


Simple Consistency

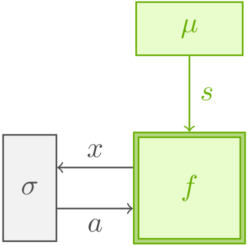
0100010001001010010110111010...

$u(x, a, s^+), \delta$

$\mathbb{P}[0], \mathbb{P}[1]$

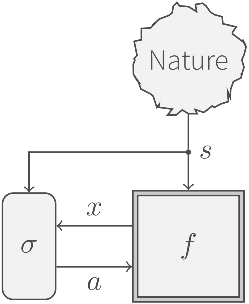


$\mu[s] = \mathbb{P}[s]$



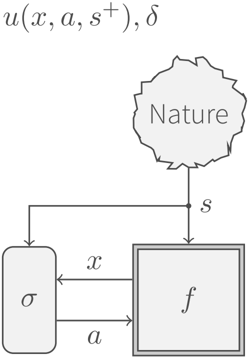
Two Systems

$u(x, a, s^+), \delta$

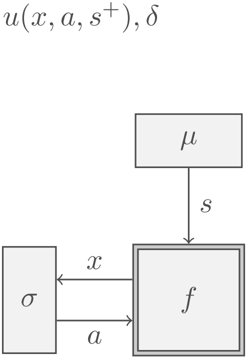


Real System: $\mathbf{R}$

Two Systems

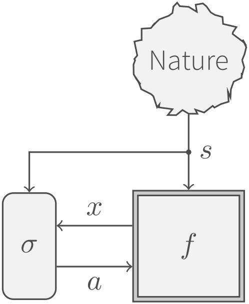


Real System: $\mathbf{R}$



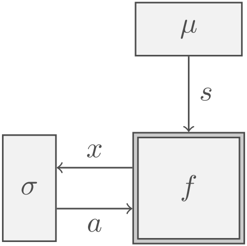
Mockup System: $\mathbf{M}$

Two Systems



Real System: $\mathbf{R}$

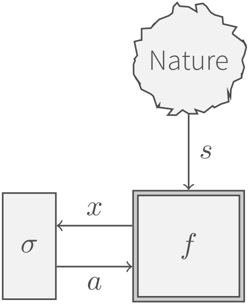
$u(x, a, s^+), \delta$



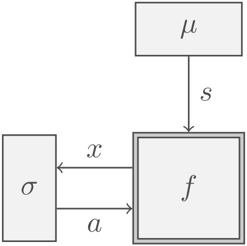
Mockup System: $\mathbf{M}$

Two Systems

$u(x, a, s^+), \delta$



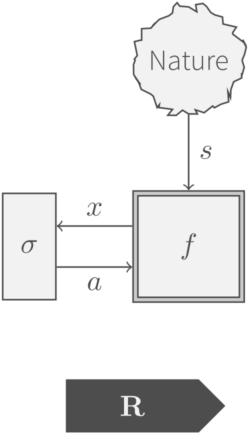
Real System: $\mathbf{R}$

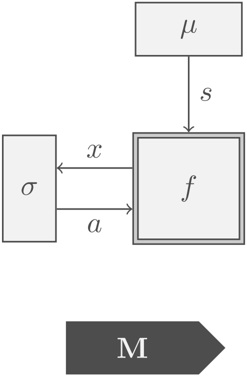


Mockup System: $\mathbf{M}$

Two Systems

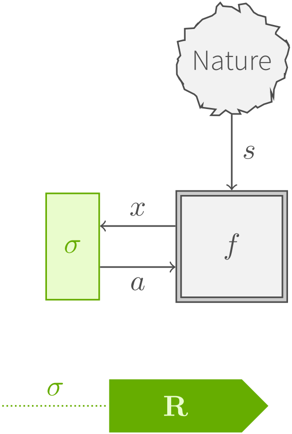
$u(x, a, s^+), \delta$

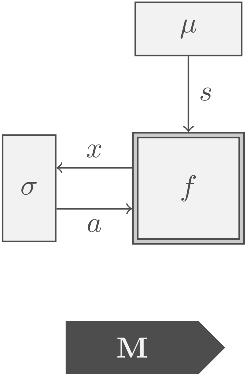




Two Systems

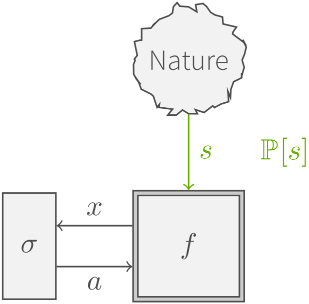
$u(x, a, s^{+}), \delta$

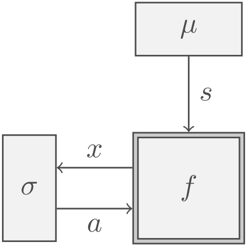




Two Systems

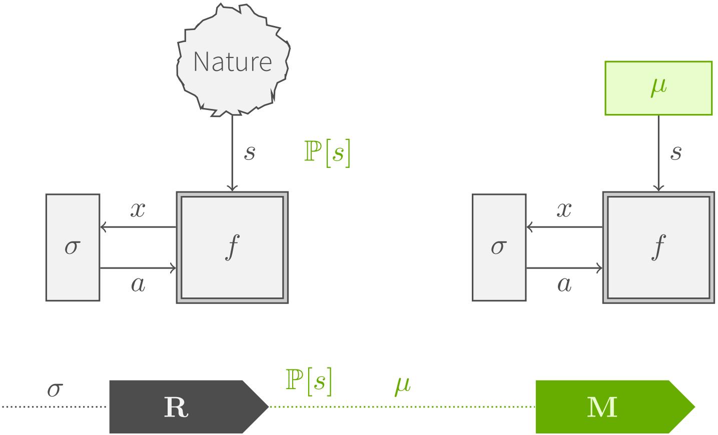
$u(x, a, s^+), \delta$



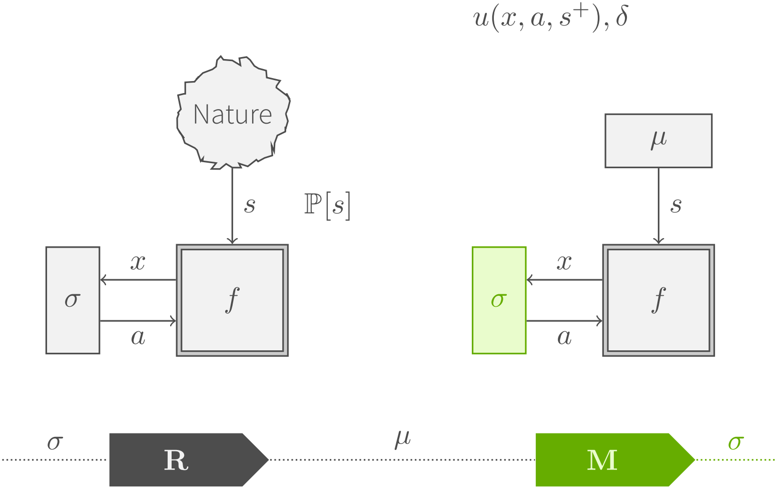


Two Systems

$u(x, a, s^+), \delta$



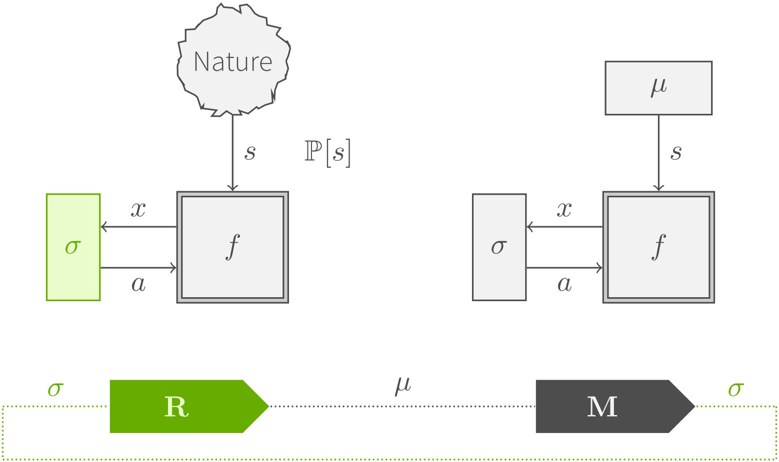
Two Systems



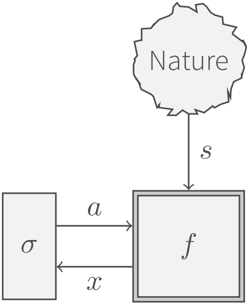
$u(x, a, s^+), \delta$
Nature
s
$\mathbb{P}[s]$
x
σ
f
a
μ
s
x
σ
f
a
σ
R
μ
M
σ

Two Systems

$$u(x, a, s^+), \delta$$

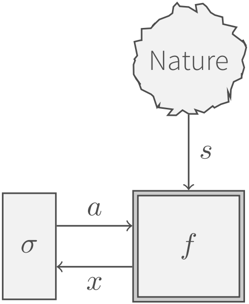


Consistency



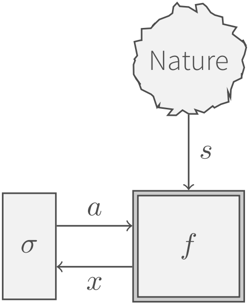
$$\mu\left[s^+\right] = \mathbb{P}\left[s^+\right]$$

Consistency



$$
\begin{aligned}
\mu\left[s^{+}\right] &= \mathbb{P}\left[s^{+}\right] \\
&= \lim_{t\to\infty} \mathbb{P}\left[S^{t+1} = s^{+}\right]
\end{aligned}
$$

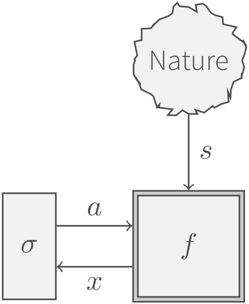
Consistency



$$\begin{cases} w^+ \sim n(w, x, a) \\ s^+ \sim \nu(w^+) \end{cases}$$

$$\begin{aligned} \mu\left[s^+\right] &= \mathbb{P}\left[s^+\right] \\ &= \lim_{t\to\infty} \mathbb{P}\left[S^{t+1} = s^+\right] \end{aligned}$$

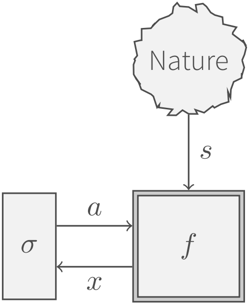
Consistency



$$\begin{cases} w^{+} \sim n(w, x, a) \\ s^{+} \sim \nu\left(w^{+}\right) \end{cases}$$

$$\begin{aligned} \mu\left[s^{+}\right] &= \mathbb{P}\left[s^{+}\right] \\ &= \lim_{t\to\infty} \mathbb{P}\left[S^{t+1} = s^{+}\right] \\ &= \sum_{w^{+},w,x,a} \nu\left(w^{+}\right)\left[s^{+}\right] \cdot \pi[w,x] \cdot \sigma(x)[a] \cdot n(w,x,a)\left[w^{+}\right] \end{aligned}$$

Consistency



$$\begin{cases} w^+ \sim n(w, x, a) \\ s^+ \sim \nu(w^+) \end{cases}$$

$$\begin{aligned} \mu[s^+] &= \mathbb{P}[s^+] \\ &= \lim_{t\to\infty} \mathbb{P}[S^{t+1} = s^+] \\ &= \sum_{w^+, w, x, a} \nu(w^+)[s^+] \cdot \pi[w, x] \cdot \sigma(x)[a] \cdot n(w, x, a)[w^+] \end{aligned}$$

Depth-1 Consistency

01010101010101010101010101…

Depth-1 Consistency

0101010101010101010101010101...

$\mathbb{P}[00], \mathbb{P}[01], \mathbb{P}[11], \mathbb{P}[10]$

Depth-1 Consistency

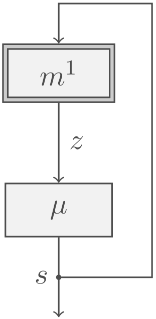
0101010101010101010101010101...

$$\mathbb{P}[00], \mathbb{P}[01], \mathbb{P}[11], \mathbb{P}[10] \iff \mathbb{P}[0 \mid 0], \mathbb{P}[1 \mid 0], \mathbb{P}[0 \mid 1], \mathbb{P}[1 \mid 1]$$

Depth-1 Consistency

01010101010101010101010101010101...

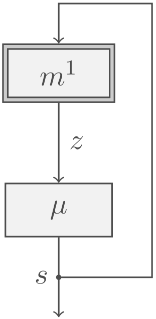
$$\mathbb{P}[00], \mathbb{P}[01], \mathbb{P}[11], \mathbb{P}[10] \iff \mathbb{P}[0 \mid 0], \mathbb{P}[1 \mid 0], \mathbb{P}[0 \mid 1], \mathbb{P}[1 \mid 1]$$

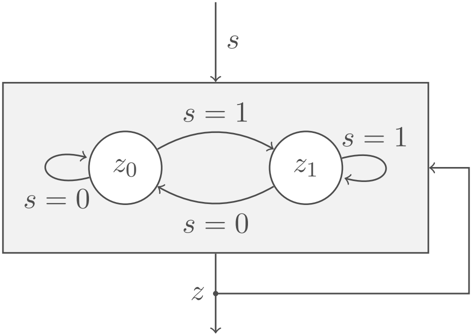


Depth-1 Consistency

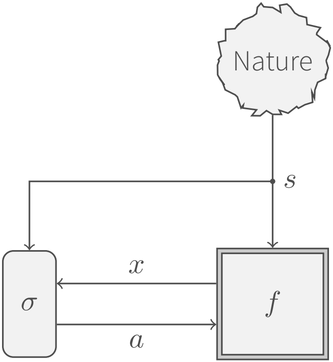
01010101010101010101010101010101…

$$\mathbb{P}[00], \mathbb{P}[01], \mathbb{P}[11], \mathbb{P}[10] \iff \mathbb{P}[0 \mid 0], \mathbb{P}[1 \mid 0], \mathbb{P}[0 \mid 1], \mathbb{P}[1 \mid 1]$$



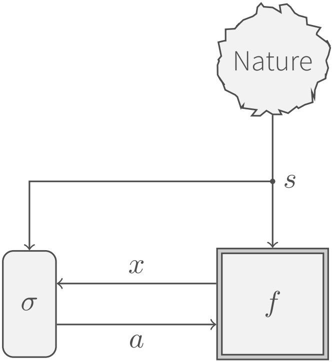


Two Systems

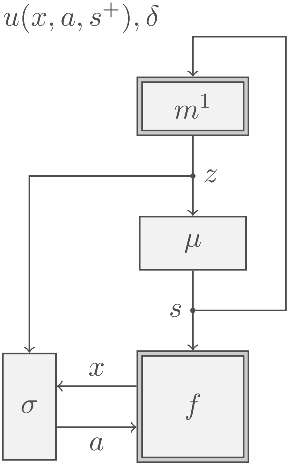


Real System: $\mathbf{R}$

Two Systems

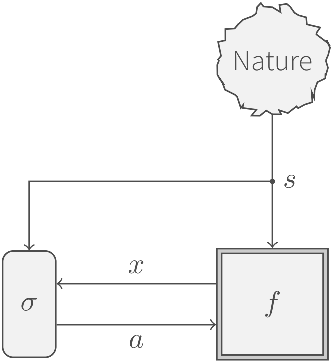


Real System: $\mathbf{R}$

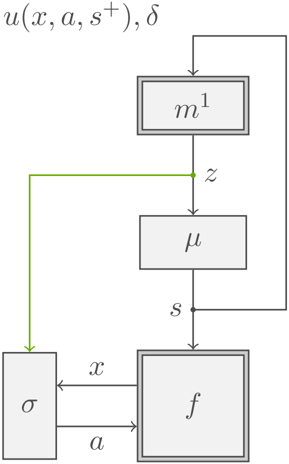


Mockup System: $\mathbf{M}$

Two Systems

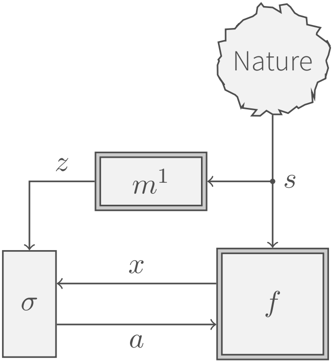


Real System: $\mathbf{R}$

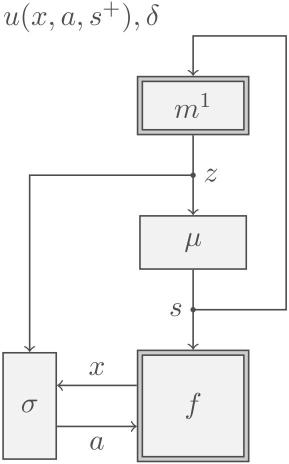


Mockup System: $\mathbf{M}$

Two Systems

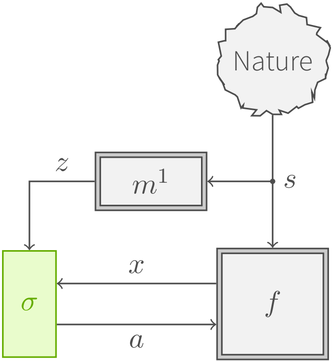


Real System: $\mathbf{R}$

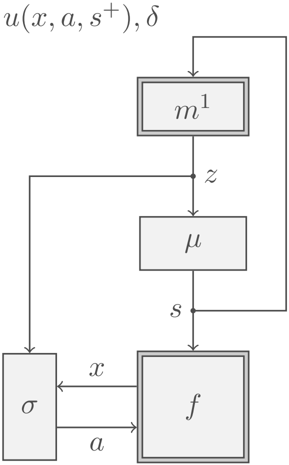


Mockup System: $\mathbf{M}$

Two Systems

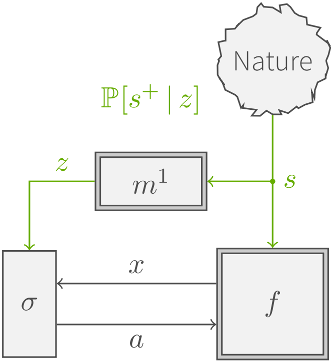


Real System: $\mathbf{R}$

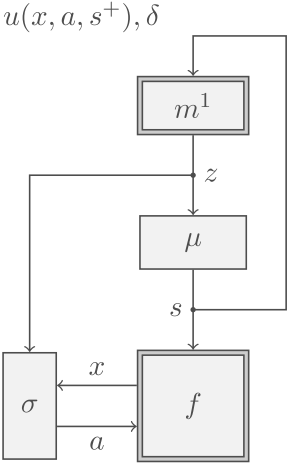


Mockup System: $\mathbf{M}$

Two Systems

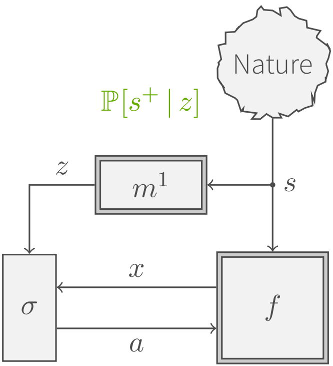


Real System: $\mathbf{R}$

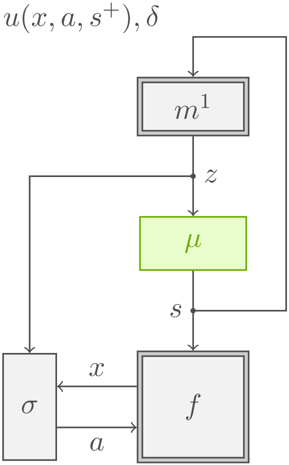


Mockup System: $\mathbf{M}$

Two Systems

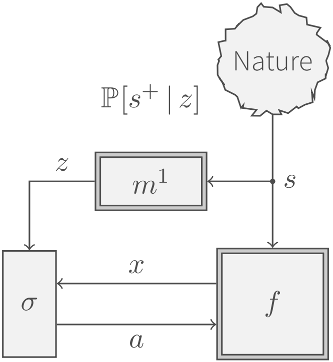


Real System: $\mathbf{R}$

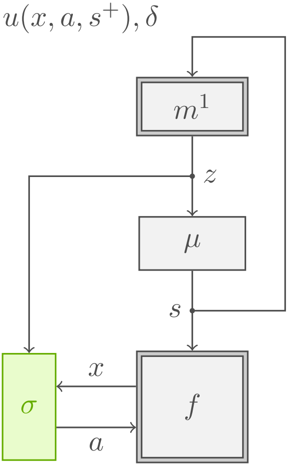


Mockup System: $\mathbf{M}$

Two Systems

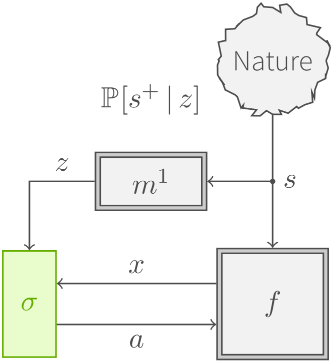


Real System: $\mathbf{R}$

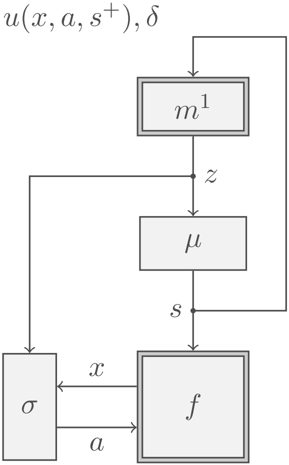


Mockup System: $\mathbf{M}$

Two Systems

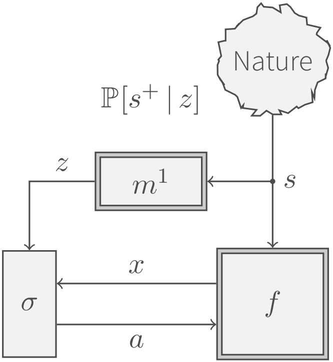


Real System: $\mathbf{R}$

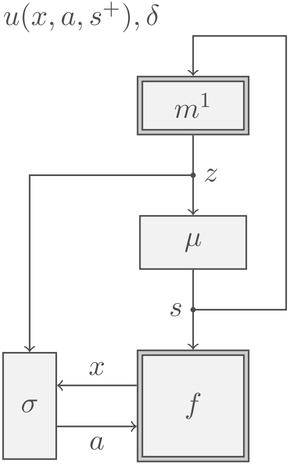


Mockup System: $\mathbf{M}$

Two Systems

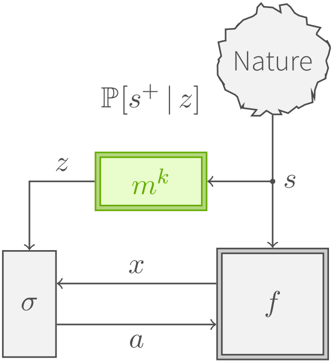


Real System: $\mathbf{R}$

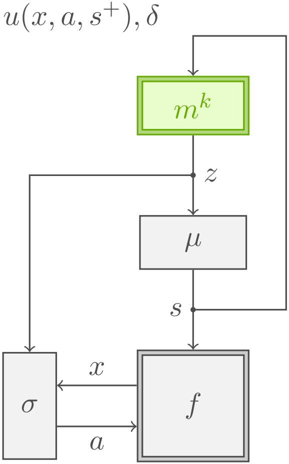


Mockup System: $\mathbf{M}$

Two Systems

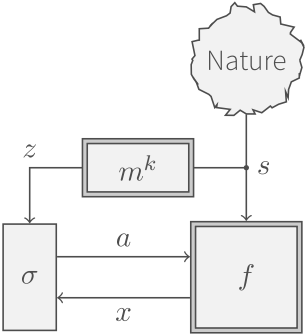


Real System: $\mathbf{R}$



Mockup System: $\mathbf{M}$

Depth-k Consistency



$$\begin{cases} w^+ \sim n(w, x, a) \\ s^+ \sim \nu(w^+) \end{cases}$$

$$\begin{aligned} \mu(z)\left[s^+\right] &= \mathbb{P}\left[s^+ \mid z\right] \\ &= \lim_{t\to\infty} \mathbb{P}\left[S^{t+1} = s^+ \mid Z^t = z\right] \end{aligned}$$

Recap

Start with one agent

Arbitrarily fix a model m^k

Split hard problem:

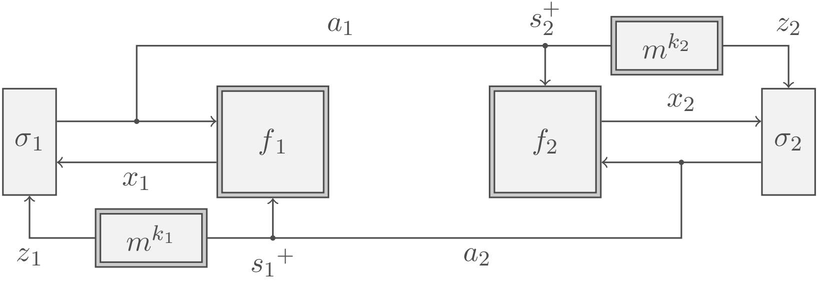
- Markov chain $\mathbf{R}$
  $\implies$ consistent predictor μ
- MDP $\mathbf{M}$
  $\implies$ optimal strategy σ

EEEs are fixed points of:

Stochastic Game

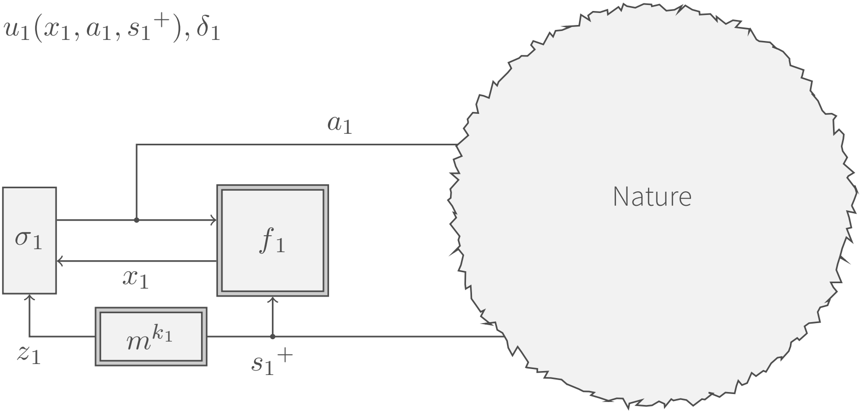
$u_1(x_1, a_1, s_1{}^+), \delta_1$

$u_2(x_2, a_2, s_2^+), \delta_2$



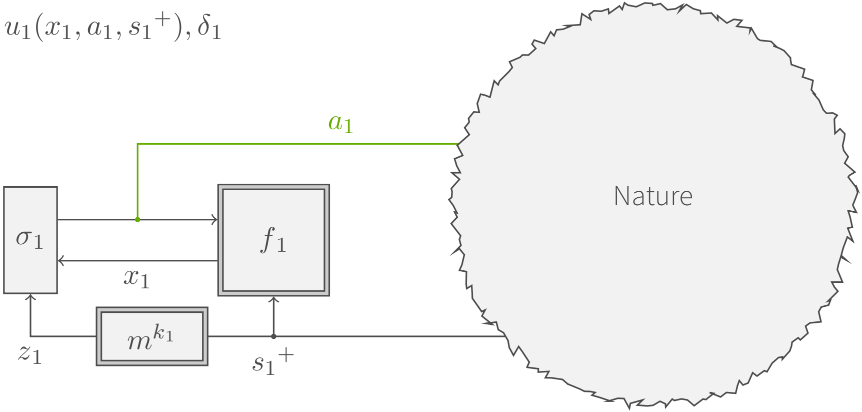
Stochastic Game

$u_1(x_1, a_1, {s_1}^+), \delta_1$



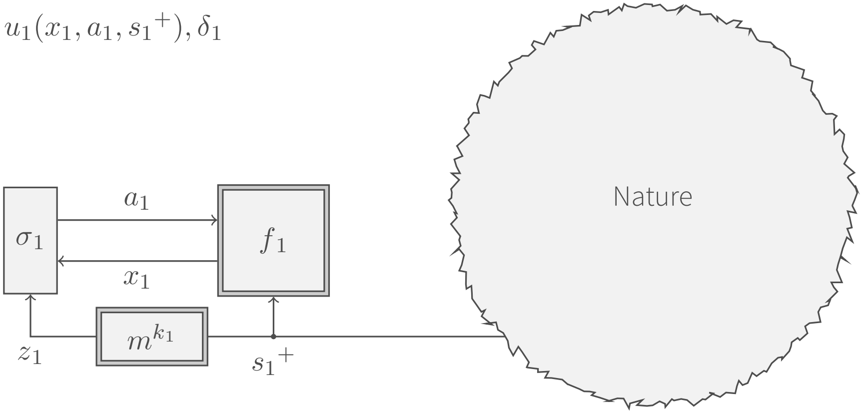
Stochastic Game

$u_1(x_1, a_1, {s_1}^+), \delta_1$



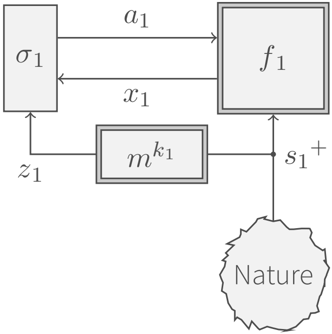
Stochastic Game

$u_1(x_1, a_1, {s_1}^+), \delta_1$

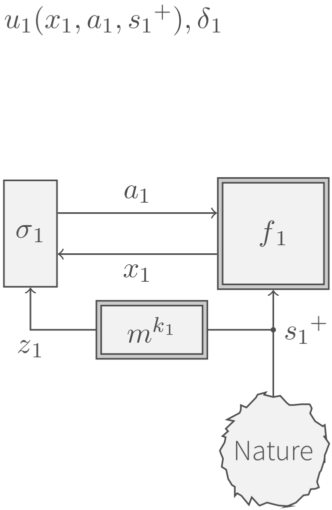


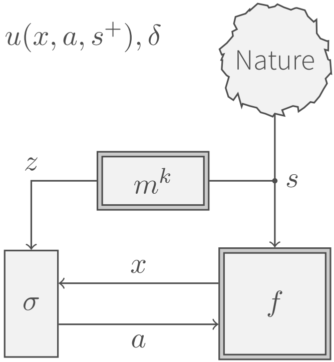
Stochastic Game

$$u_1(x_1, a_1, {s_1}^+), \delta_1$$

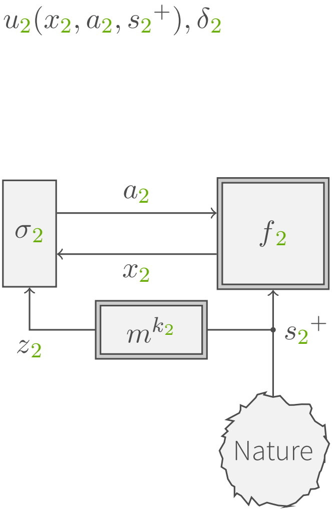


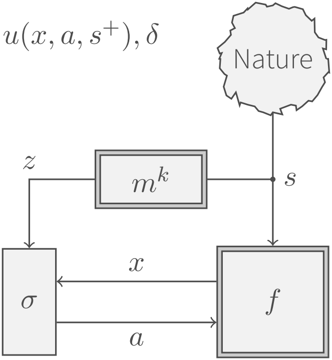
Stochastic Game





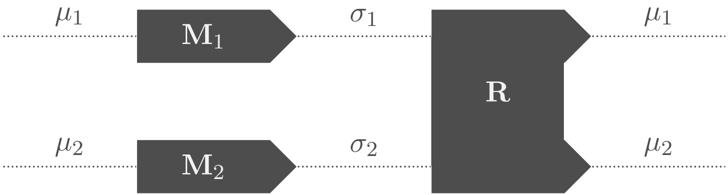
Stochastic Game





Empirical-evidence Equilibrium

m^{k_1} and m^{k_2} fixed

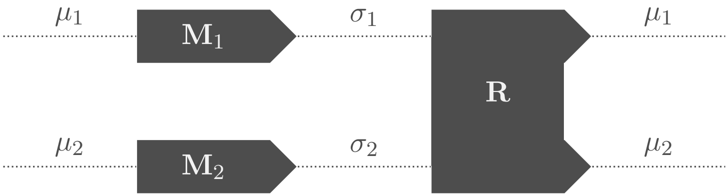


$(\mu_1, \sigma_1, \mu_2, \sigma_2)$ is an empirical-evidence equilibrium (EEE) if:

- μ_1 is consistent with $\mathbf{R}$
- μ_2 is consistent with $\mathbf{R}$
- σ_1 is optimal for $\mathbf{M}_1$
- σ_2 is optimal for $\mathbf{M}_2$

Empirical-evidence Equilibrium

m^{k_1} and m^{k_2} fixed



$(\mu_1, \sigma_1, \mu_2, \sigma_2)$ is an ε empirical-evidence equilibrium (ε EEE)if:

- μ_1 is consistent with $\mathbf{R}$
- μ_2 is consistent with $\mathbf{R}$
- σ_1 is ε optimal for $\mathbf{M}_1$
- σ_2 is ε optimal for $\mathbf{M}_2$

EEE vs Nash

- optimization complexity fixed by agent not opponents
- always implementable
- each agent knows when at equilibrium
- less intrinsic to the problem

Existence of ε EEEs

Theorem

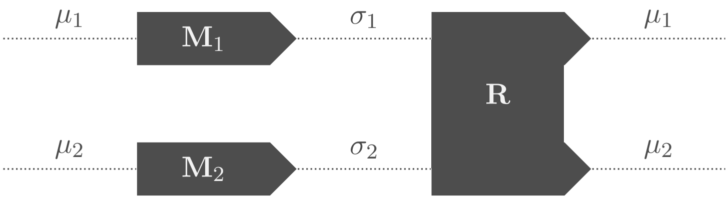
For any m^{k_1} and m^{k_2}, there exists an ε EEE.

Existence of ε EEEs

Theorem

For any m^{k_1} and m^{k_2}, there exists an ε EEE.

Proof.



- ε and Gibbs distribution $\implies \mu_i \mapsto \sigma_i$ is a function
- $\mu \mapsto \mu$ is a continuous function
- set of predictors is compact and convex
- Brouwer's fixed point theorem

$\square$

Existence of EEEs

New

Theorem

For any m^{k_1} and m^{k_2}, there exists a EEE.

Existence of EEEs New

Theorem

For any m^{k_1} and m^{k_2}, there exists a EEE.

Proof.

- $\mu \mapsto \mu$ is a closed-graph correspondence
- set of predictors is compact and convex
- Kakutani's fixed point theorem

$\square$

Characterization of EEEs

New

Theorem

Exogenous EEEs in perfect-monitoring repeated games yield correlated equilibria of the underlying one-shot game.

Repeated game:

Stochastic game without a state

Correlated equilibrium:

Nash equilibrium with common source of randomness

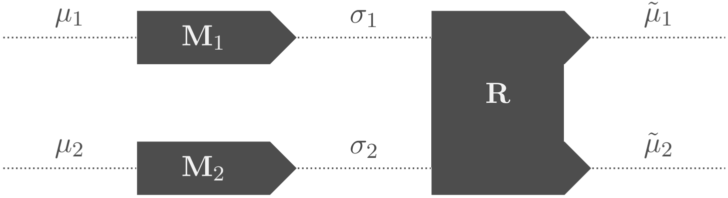
Recap

- multiagent EEE identical to single agent
- each agent arbitrarily picks a model m^k
- EEEs always exist
- EEEs induce correlated equilibria in repeated games

Asset Management Example

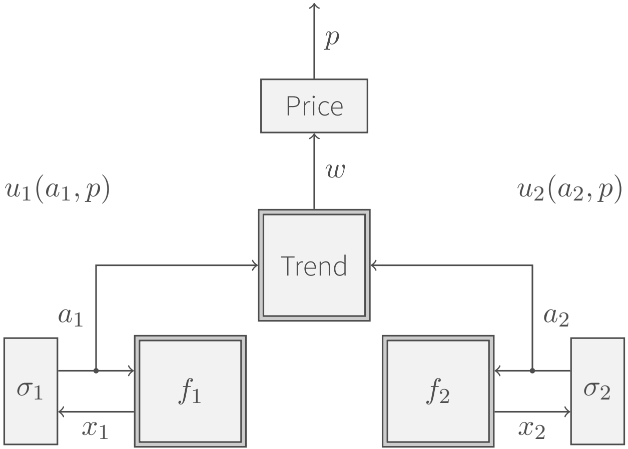
State	holdings $x_i \in [\![0, M]\!]$
Action	sell one, hold, or buy one $a_i \in \{-1, 0, 1\}$
Signal	price $p \in \{\text{Low}, \text{High}\}$
Dynamic	$x_i^+ = x_i + a_i$
Stage cost	$p \cdot a_i$
Nature	market trend $w \in \{\text{Bull}, \text{Bear}\}$
Model	depth 0

Iterative Process

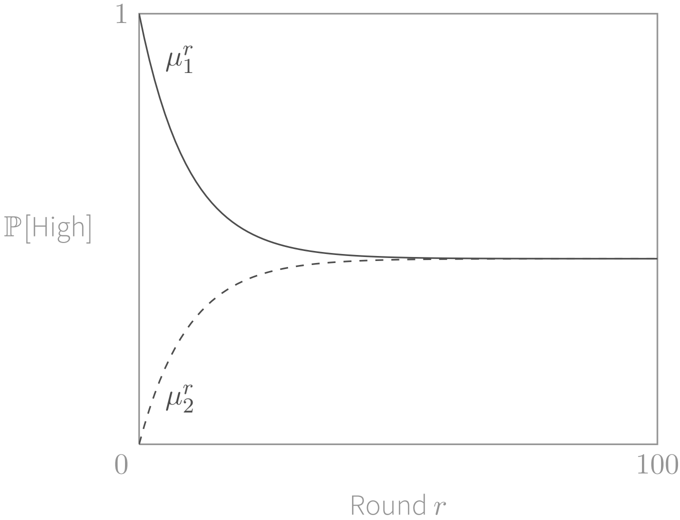


Update Rule $\quad \mu_i^{r+1} = (1 - \alpha^r)\mu_i^r + \alpha^r(\tilde{\mu}_i - \mu_i^r)$

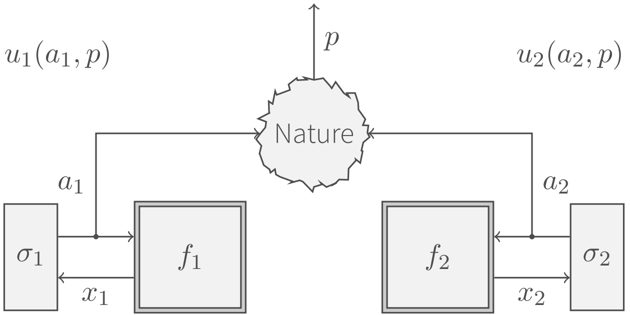
Theoretical Predictor



Update Rule $\quad \mu_i^{r+1} = (1-\alpha)\mu_i^r + \alpha(\tilde{\mu}_i - \mu_i^r)$


1
μ_1^r
$\mathbb{P}$[High]
μ_2^r
0
100
Round r

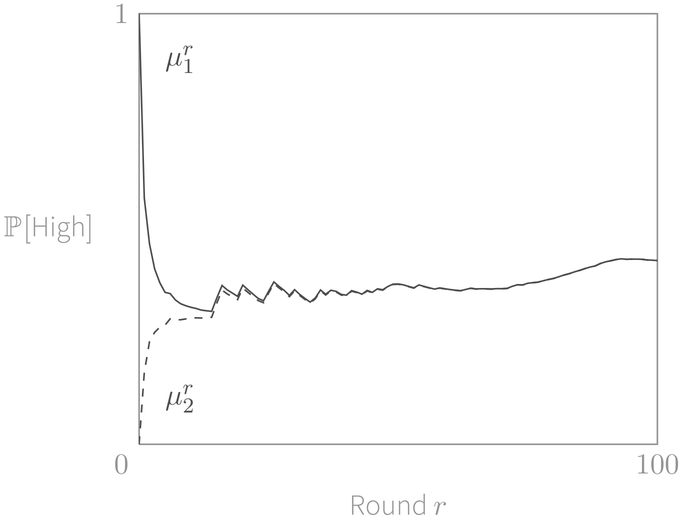
Empirical Predictor



Update Rule

$$\mu_i^{r+1} = (1 - \alpha^r)\mu_i^r + \alpha^r \left(\tilde{\mu}_i^T - \mu_i^r\right)$$

α^r non-summable, square-summable



1
μ_1^r
$\mathbb{P}$[High]
μ_2^r
0
100
Round r

Hawk-dove Game

Repeated game

	h	d
H	$-1, -1$	$6, 0$
D	$0, 6$	$3, 3$

Nash equilibria (H, d) and (D, h)

Want correlated equilibrium alternating between the two

Hawk-dove Game

Depth-2 models

Strategies:

$$\sigma_1(\mathrm{d},\mathrm{h}) = 0.999\,\mathrm{H} + 0.001\,\mathrm{D}$$
$$\sigma_1(\mathrm{h},\mathrm{d}) = 0.999\,\mathrm{D} + 0.001\,\mathrm{H}$$
$$\sigma_1(\mathrm{h},\mathrm{h}) = 0.5\mathrm{H} + 0.5\mathrm{D}$$
$$\sigma_1(\mathrm{d},\mathrm{d}) = 0.5\mathrm{H} + 0.5\mathrm{D}$$

Associated predictors:

$$\mu_1(\mathrm{d},\mathrm{h}) = 0.996\,\mathrm{d} + 0.004\,\mathrm{h}$$
$$\mu_1(\mathrm{h},\mathrm{d}) = 0.996\,\mathrm{h} + 0.004\,\mathrm{d}$$
$$\mu_1(\mathrm{h},\mathrm{h}) = 0.5\,\mathrm{h} + 0.5\,\mathrm{d}$$
$$\mu_1(\mathrm{d},\mathrm{d}) = 0.5\,\mathrm{h} + 0.5\,\mathrm{d}$$

Strategy approximately optimal as δ close enough to one

Generalizes to any convex combination of pure Nash equilibria

Recap

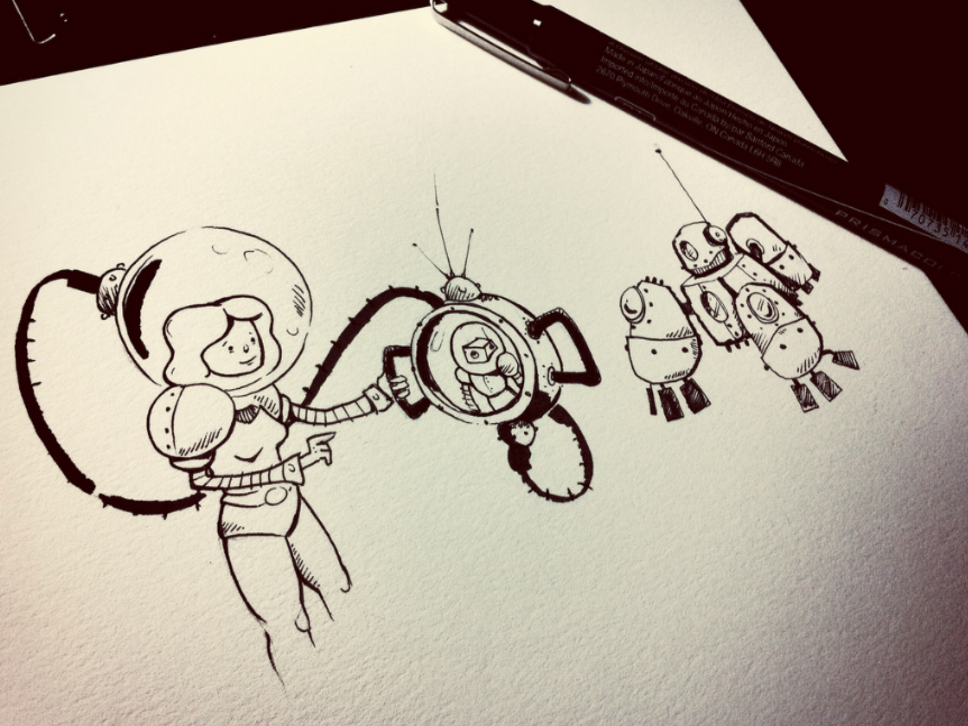
Predictive	given models and adaptation rule a EEE emerges
Prescriptive	implement desired outcome as a EEE

Extensions

- n agents
- endogenous models $z^+ \sim m(z, x, a, s)$
- notions of consistency: approximate, weak, and eventual New
- convergence of empirical iterative process when theoretical one converges New

Empirical-evidence Equilibrium (EEE)

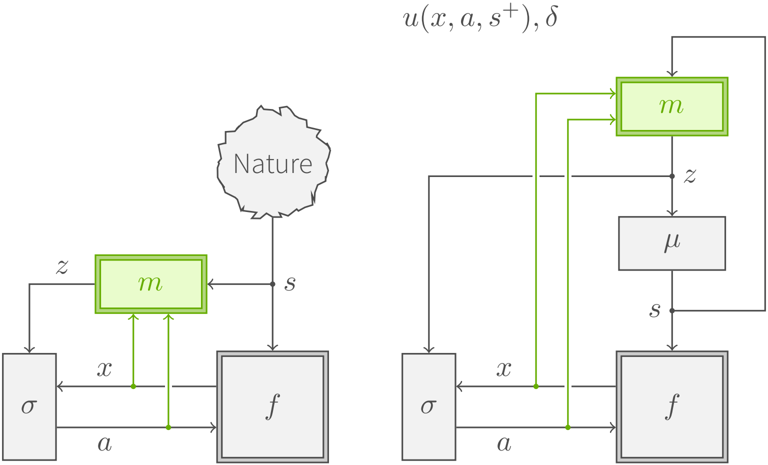
Motivation	intractable problem
Definition	split into Markov chain and consistent MDPs
Existence	fixed-point theorems
Comparison	lower computational requirements
Characterization	correlated equilibrium in repeated game
Predictive Use	model to understand stock price
Prescriptive Use	desired outcome encoded as EEE


Made in Japan/Fabriqué au Japon/Hecho en Japon
Imported into/Importé au Canada by/par Sanford Canada
2670 Plymouth Drive, Oakville, ON Canada L6H 5R6

Publications

- N. Dudebout and J. S. Shamma, “Empirical Evidence Equilibria in Stochastic Games,” in *51st IEEE Conference on Decision and Control*, Dec. 2012, pp. 5780–5785
- N. Dudebout and J. S. Shamma, “Exogenous Empirical-evidence Equilibria in Perfect-monitoring Repeated Games Yield Correlated Equilibria,” in *53rd IEEE Conference on Decision and Control*, Submitted
- N. Dudebout and J. S. Shamma, “Empirical-evidence Equilibrium,” in *Games and Economic Behavior*, In Preparation

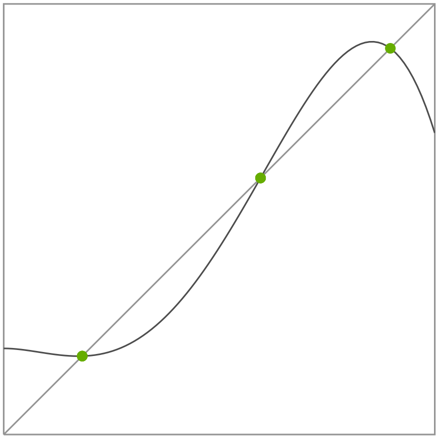
Endogenous Model



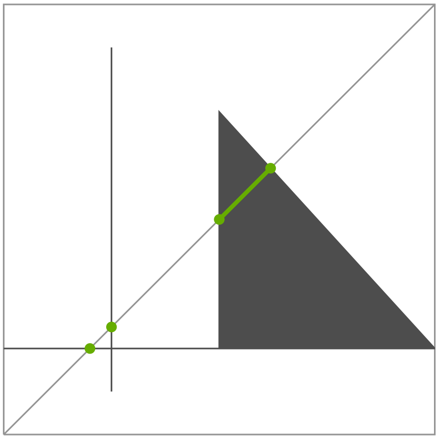
Real System: $\mathbf{R}$

Mockup System: $\mathbf{M}$

Brouwer's Fixed-point Theorem



Kakutani's Fixed-point Theorem



Consistency Formula

$$\mu(z)\left[s^{+}\right] = \sum_{w^{+}} \nu\left(w^{+}\right)\left[s^{+}\right] \frac{\sum_{w,x,a} \pi_{\sigma}[w,x,z] \cdot \sigma(z)[a] \cdot n(w,x,a)[w^{+}]}{\sum_{w,x} \pi_{\sigma}[w,x,z]}$$

Consistency

Strong Consistency

$$\mu(z)\left[s^+\right] = \lim_{t\to\infty} \mathbb{P}\left[S^{t+1} = s^+ \,\middle|\, Z^t = z\right]$$

Weak Consistency New

$$\mu(z)\left[s^+\right] = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{P}\left[S^{t+1} = s^+ \,\middle|\, Z^t = z\right]$$

Eventual Consistency New

$$\lim_{t\to\infty} \mathbb{P}\left[Z^t = z\right] > 0 \implies \mu(z)\left[s^+\right] = \lim_{t\to\infty} \mathbb{P}\left[S^{t+1} = s^+ \,\middle|\, Z^t = z\right]$$

Learning Result

Theorem

Suppose the theoretical learning dynamic has a Lyapunov function. For a large enough observation window, the empirical learning dynamic converges.

Proof.

- ODE method for stochastic approximation
- Lyapunov stability of perturbed systems

□