Mean field equilibria in large scale dynamic games

Ramesh Johari

Stanford University

Overview

Dynamic systems with many interacting agents are everywhere:

- Resource sharing (e.g., spectrum sharing)
- Financial markets
- Social networks
- Online auctions

What tools are useful for engineers studying dynamic systems with many interacting agents?

Overview

Dynamic game theory is the primarily tool economics offers us to study such systems.

But: traditional dynamic game theory is impractical and implausible at large scale.

As a result, we lose the ability to provide design guidance for these systems.

This talk is about an approximate approach: *Mean field equilibrium*

This talk

(1) Stochastic dynamic games:
a short introduction

(2) Mean field equilibrium

(3) Existence of MFE

(4) MFE as a approximation to finite systems

(5) Open questions and directions

Stochastic dynamic games: a short introduction

A single agent problem

Discrete time, infinite horizon

x_t : State at time t

a_t : Action at time t

$x_{t+1} \sim P(\cdot \mid x_t, a_t)$: State transition kernel

$\pi(x_t, a_t)$: Per period payoff

β : Discount factor ($0 < \beta < 1$)

A single agent problem

Stochastic control problem:

Maximize $E[\sum_{t \geq 0} \pi(x_t, a_t) \mid x_0 = x]$

over all policies

Under "reasonable" assumptions:

An optimal stationary Markov strategy exists

Optimal value: $V(x) = \max_a[\pi(x, a) + \sum_{x'} V(x') P(x' \mid x, a)]$

Any optimal strategy maximizes RHS in each state x

Ex: Linear-quadratic control

Assume state and action are scalar.

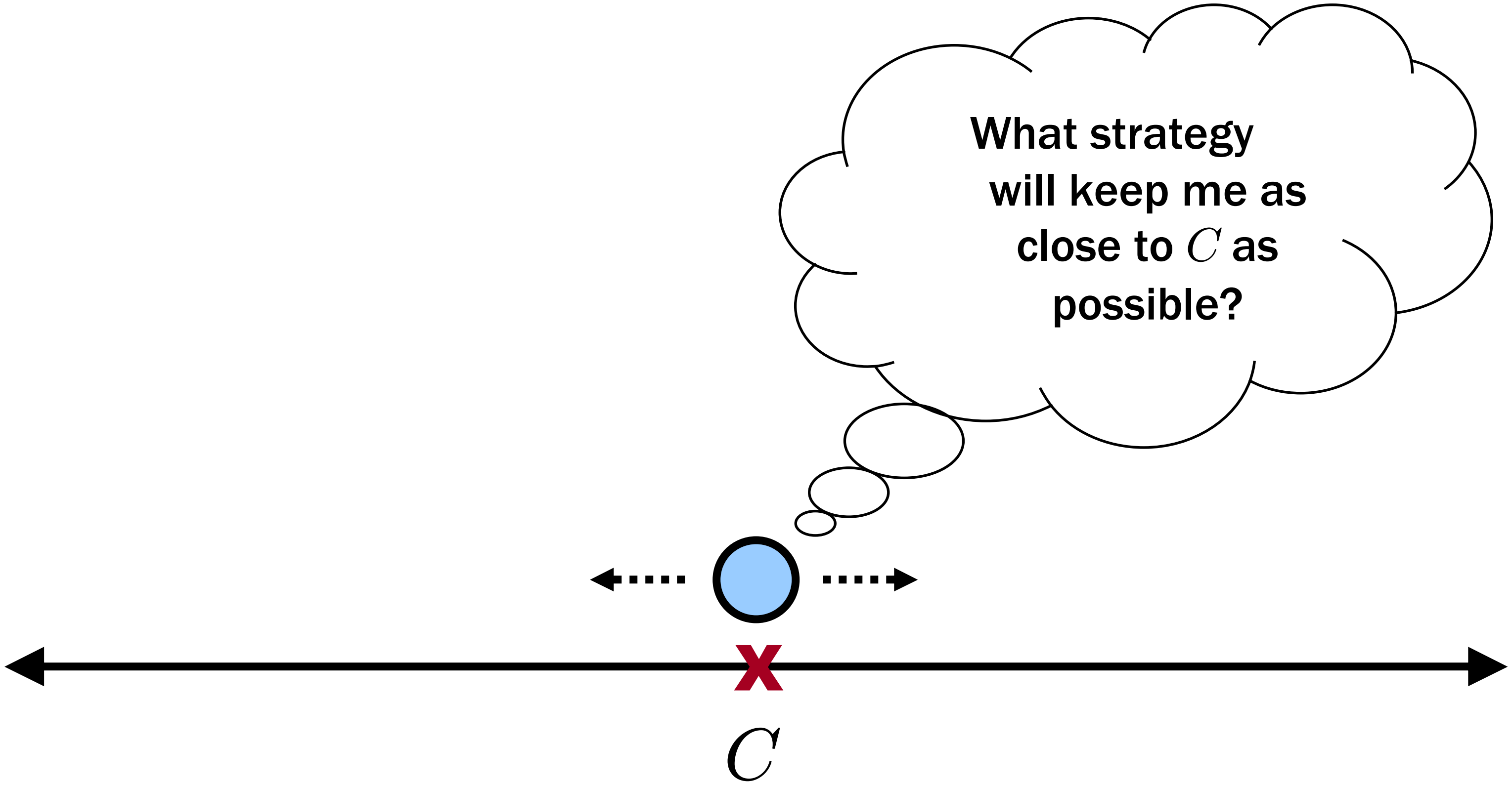
Consider the following model:

Payoff: $\pi(x, a) = -(x - C)^2 - a^2$

Dynamics: $x_{t+1} = A\, x_t + B\, a_t + w_t$,

where w_t is i.i.d. noise

Ex: Linear-quadratic control



Solution: Kalman filtering.

Stochastic dynamic games

A stochastic dynamic game is the *multiple agent* generalization of single agent stochastic control.

Key change:

Agents' payoffs and state transitions can now depend on *other* agents as well.

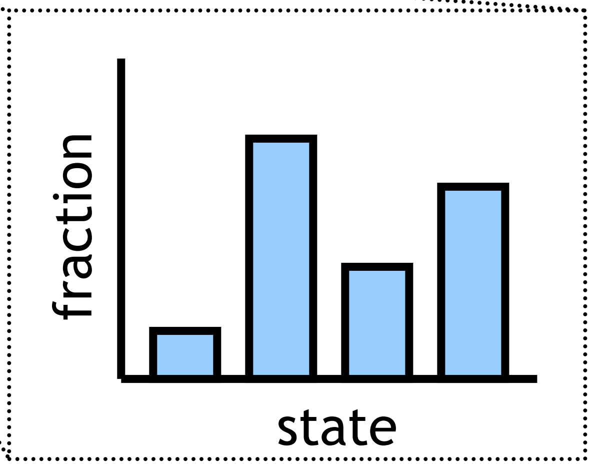
A stochastic dynamic game

Discrete time, infinite horizon

x_t : State of an agent at time t

a_t : Action of an agent at time t

$\mathbf{f}_t$: Empirical distribution of others' states at time t
(the *population state*)



$x_{t+1} \sim \mathbf{P}(\cdot \mid x_t, a_t, \mathbf{f}_t)$:
State transition kernel

$\pi(x_t, a_t, \mathbf{f}_t)$: Per period payoff

β : Discount factor ($0 < \beta < 1$)

Ex 2: Distributed coordination

Again assume state and action are scalar.

Consider the following model:

Payoff: $\pi(x, a) = -(x - \text{mean}(f))^2 - a^2$

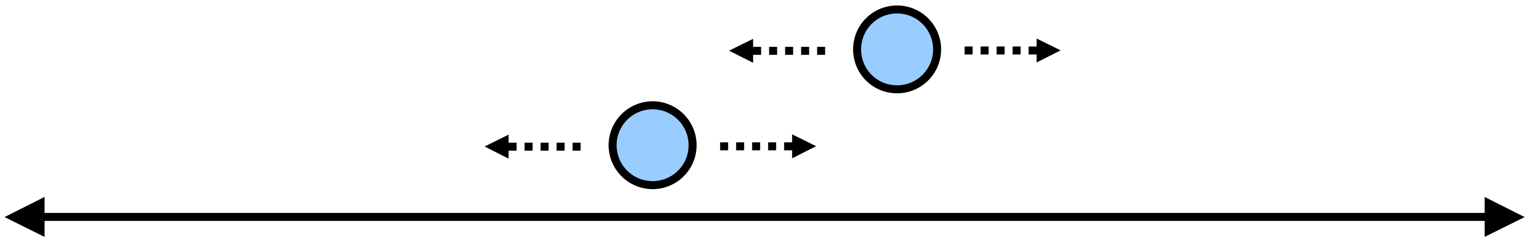
Dynamics: $x_{t+1} = A x_t + B a_t + w_t$,

where w_t is i.i.d. noise

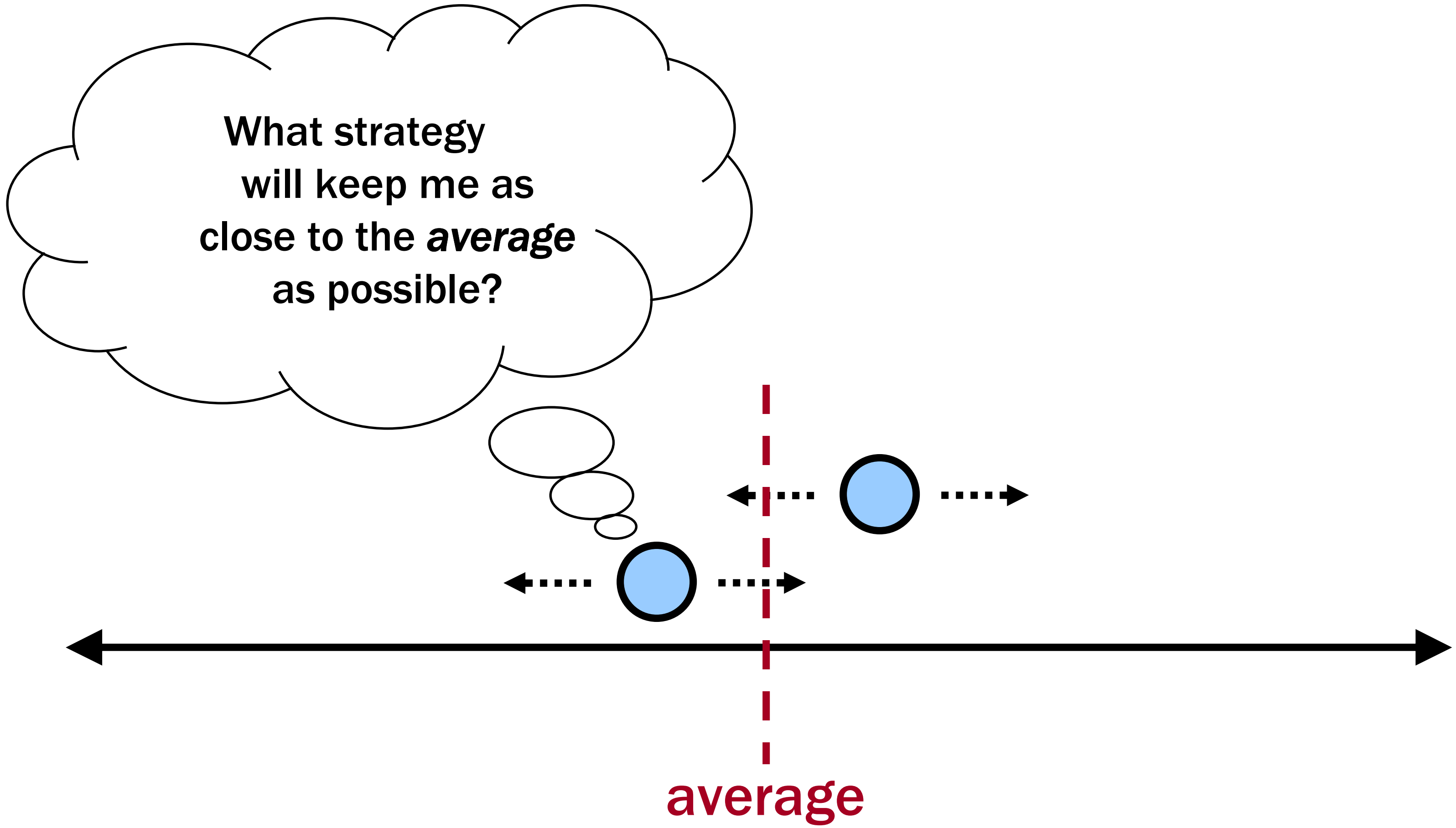
Now the goal is to maintain state close to the mean of the population state, subject to disturbances w.

Ex 2: Distributed coordination

Consider the coordination game with two agents.



Ex 2: Distributed coordination



Reasoning about the game

How should an agent think about the game?

Still want to choose a strategy to maximize expected discounted payoff.

But this depends on how the agent believes her partner's state will evolve!

Reasoning about the game

How does agent 1 reason about agent 2's state?

Answer 1: Assume constant.

Clearly unsatisfactory.

Answer 2: Assume fixed stochastic process.

Better, but assumes agent 2 *does not react to agent 1's state evolution.*

Answer 3: Assume fixed *policy* for agent 2.

Ideal: can perfectly simulate behavior of agent 2… if the strategy is what agent 2 actually plays.

Markov perfect equilibrium

So assume an agent aims to maximize expected discounted payoff:

$$V(x, \mathbf{f} | \mu', \mu) = \mathrm{E}\left[\sum_{t \geq 0} \beta^t \pi(x_t, a_t, \mathbf{f}_t) \middle| x_0 = x, f_0 = \mathbf{f}, \mu', \mu\right]$$

μ' = this agent's strategy

μ = strategy followed by all others (symmetric)

Here μ' is a **cognizant (Markov) strategy**: it can depend on *both* x_t and $\mathbf{f}_t$

Markov perfect equilibrium

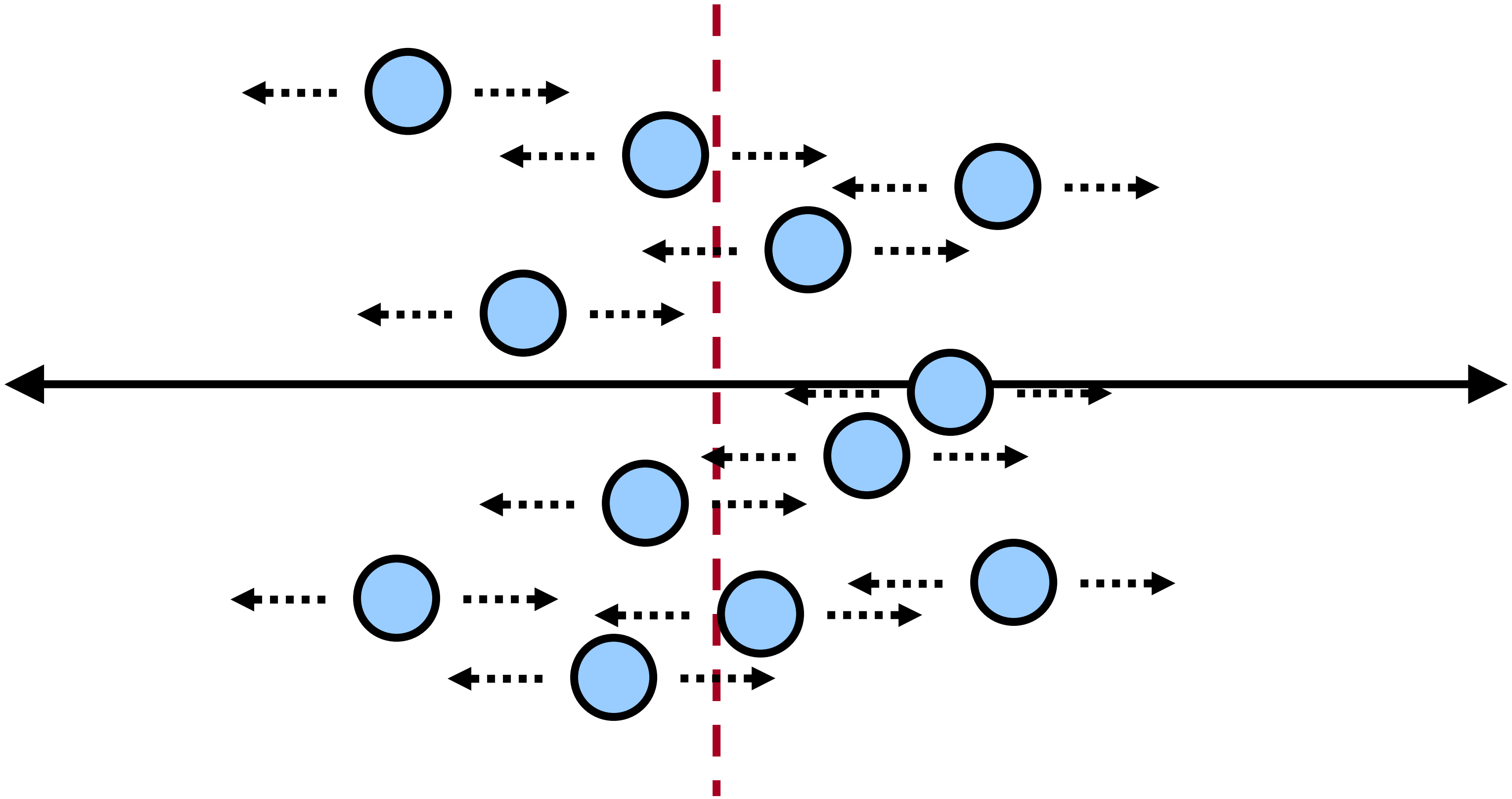
A (symmetric) Markov perfect equilibrium (MPE) is a strategy μ such that:

μ solves each agent's optimal control problem, *when they assume that all other agents will also be using* μ*.*

[Note that the "state" of the new system is larger: for each agent, the state is (x_t, $\mathbf{f}_t$).]

Large scale

What happens with many agents?



Cognizant vs. oblivious

If all players are *cognizant*, each player accounts for:

(1) the precise evolution of other players' states; and

(2) the impact their actions have on other players' state evolution.

Such equilibria are *difficult to compute* and often *implausible in practice.*

[MPE suffers from the "the curse of dimensionality" as the number of players grows.]

By contrast, *oblivious* players ignore (1) and (2).

The mean field model

When the number of agents is large, suppose:

An agent reacts only to the long run average state distribution $\underline{\mathbf{f}}$ of other players.

Mean field expected discounted payoff:

$$\underline{V}(x|\underline{\mu}, \underline{\mathbf{f}}) = \mathrm{E}\left[\sum_{t \geq 0} \beta^t \pi(x_t, a_t, \underline{\mathbf{f}}) \middle| x_0 = x, \underline{\mu}, \underline{\mathbf{f}}\right]$$

Here $\underline{\mu}$ is a **oblivious strategy**: it depends *only* on x_t

The mean field model

Mean field expected discounted payoff:

$$\underline{V}(x|\underline{\mu}, \underline{\mathbf{f}}) = \mathrm{E}\left[\sum_{t \geq 0} \beta^t \pi(x_t, a_t, \underline{\mathbf{f}}) \middle| x_0 = x, \underline{\mu}, \underline{\mathbf{f}}\right]$$

Notice that since f is constant in this problem, this again looks as hard or easy as the original single agent control problem.

Example: In the mean field model of the distributed coordination game, Kalman filtering would again be optimal for an agent.

Mean field equilibrium

Mean field equilibrium

A strategy $\underline{\mu}$ and a population state $\underline{\mathbf{f}}$ constitute a *mean field equilibrium (MFE)* if:

(1) $\underline{\mu}$ is an optimal oblivious strategy given $\underline{\mathbf{f}}$ and

(2) $\underline{\mathbf{f}}$ is a steady state distribution of $\underline{\mu}$

Distribution $\underline{\mathbf{f}}$ → Optimal strategy $\underline{\mu}$ for DP given $\underline{\mathbf{f}}$ → New steady state distribution $\Phi(\underline{\mathbf{f}})$

A MFE population state $\underline{\mathbf{f}}$ is a fixed point of Φ: $\underline{\mathbf{f}} = \Phi(\underline{\mathbf{f}})$

MFE: Related work

- Mean field models arise in a variety of fields: physics, applied math, engineering, economics…
- Mean field models in dynamic games:

Economics

[Jovanovic & Rosenthal '88, Stokey et al. '89, Hopenhayn '92, Sleet '02, Weintraub et al. '08-'10, Acemoglu & Jepsen '10, Bodoh-Creed '11]

Control

[Glynn et al. '04, Lasry & Lions '07, Huang et al. '07-'10, Guéant '09, Tembine et al. '09, Yin et al. '09, Adlakha et al. '09-'11]

Finance [Duffie '09, Duffie '10], Transportation [Friesz et al. '93, etc.]

"Econophysics"

Existence of MFE

The existence problem

The first question to ask about an equilibrium concept is: does it exist?

The usual approach:

Show that the map Φ has a fixed point.

Kakutani-Fan-Glicksberg Theorem:

If Y is compact, and a point-to-set correspondence F : $Y \twoheadrightarrow Y$ has compact, convex, nonempty values, and has a closed graph, then F has a fixed point.

Existence of MFE

Assume for now the state space X and action space A are finite

Existence requires *randomized* (mixed) strategies.

$\Delta(A)$: All distributions over A

$\Delta(X)$: All population states over X

For $\sigma \in \Delta(A)$, define:

$\pi(x, \sigma, \mathbf{f}) = \sum_a \sigma(a)\, \pi(x, a, \mathbf{f})$

$P(x' \mid x, \sigma, \mathbf{f}) = \sum_a \sigma(a)\, P(x' \mid x, a, \mathbf{f})$

Key assumption for existence:

π and P are **continuous in $\underline{\mathbf{f}} \in \Delta(X)$**

Existence of MFE

Let S(f) be all the optimal (randomized) strategies given fixed f.

Let D(μ, f) be all the invariant distributions of P(· | x, μ(x), f).

Note Φ(f) = D(S(f), f).

Bellman optimality principle ⇒ S nonempty.

Finite state space ⇒ D nonempty.

Continuity assumption ⇒ S, D have closed graphs.

Compactness requirements easy to show.

Convexity: requires randomized strategies (exercise).

Existence of MFE

Theorem:

Suppose X and A are finite, and π(x, a, f), P(x' | x, a, f) are continuous in f $\in \Delta(X)$ for all a, x, x'.

Then there exists a MFE.

Approximation

Asymptotic equilibrium

Is MFE a good approximation to equilibrium behavior of cognizant players?

A MFE ($\underline{\mu}$, $\underline{f}$) has the *AE property* if for all x, as number of players $\to \infty$,

$$\sup_{\mu} V(x, \mathbf{f} | \mu, \underline{\mu}) - V(x, \mathbf{f} | \underline{\mu}, \underline{\mu}) \to 0,$$

where the sup is over all *cognizant strategies,* with initial states of others sampled from f.

[V = value function]

[Generalizes a definition of Weintraub et al. (2008)]

Asymptotic equilibrium

Is MFE a good approximation to equilibrium behavior of cognizant players?

A MFE ($\underline{\mu}$, $\underline{f}$) has the ***AE property*** if for all x, as number of players $\to \infty$,

EDP under best cognizant policy, given others play $\underline{\mu}$ **–** EDP under $\underline{\mu}$, given others play $\underline{\mu}$

$\longrightarrow$ 0, almost surely

Asymptotic equilibrium

Suppose:

(1) State and action spaces are **compact**

(2) Payoff is **continuous** in f

Theorem:

The AE property holds for any MFE ($\underline{\mu}$, $\underline{f}$).

[Adlakha, Johari, Weintraub]

Asymptotic equilibrium

Proof technique:

First use *compactness* to show that at every fixed time t, the finite system population state $\mathbf{f}_t$ approaches $\underline{\mathbf{f}}$ as number of agents $\to \infty$.

Then use *continuity* to translate population state convergence to a statement about a difference of payoffs.

Extensions and open questions

(1) Generalizations

The existence and approximation results can be generalized significantly:

- Nonstationarity
  [Weintraub et al.]
- Infinite state spaces, action spaces
  [Adlakha et al.]
- Continuous time
  [Huang et al., Tembine et al., Lasry and Lions, etc.]

Some results use constructive techniques. This is worth exploring further.

(2) Intractability

What does it mean to say MFE is "simpler" than MPE?

Typical argument: "curse of dimensionality".

But in the end, both rely on fixed point arguments to establish existence.

Can we establish in a computational complexity theoretic framework, that MFE is simpler?

(3) Finding MFE

How does one obtain MFE? How do agents "learn" to play an MFE?

Generally quite difficult to find strategies that converge to equilibria, especially in dynamic games.

Common approach: "best response" dynamics.

Some recent work suggests algorithms based on *model predictive control* as a valuable approach to finding MFE.

[Adlakha and Johari, Weintraub et al.]

(4) Uniqueness of MFE

May have many optimal strategies for a given population state.

May have many optimal invariant distributions for a given strategy.

May have many MFE, unrelated to each other (e.g., in distributed coordination game).

What conditions guarantee uniqueness of MFE?

(5) Average reward/infinite time

Our approximation result only holds over finite time intervals.

To obtain *infinite* time convergence, the limit dynamical system must have a unique fixed point:

f(x) = $\sum_{x'}$ f(x') P(x' | x, μ(x), f)

[Similar to *Kurtz' theorem* in stochastic analysis]

Under what conditions is this guaranteed?

Related problem: MFE in average reward models

[Glynn et al. 2004]

(6) Interaction models

MFE is valid with *full temporal mixing*:
  Interact with a small number of agents each period, but resample i.i.d. every time period

MFE is valid with *full spatial mixing*:
  Interact with everyone at every time period

What about more complex interaction models (e.g., random graphs that evolve over time?

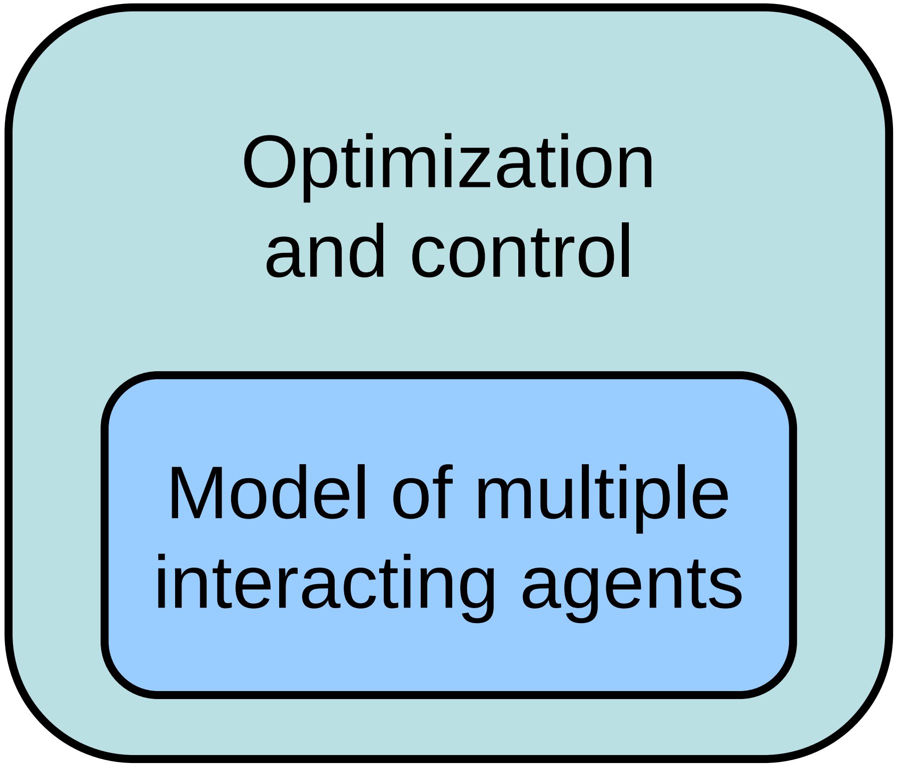
Conclusion

Big picture

Model of multiple interacting agents

Traditional game theory makes *the model* so complex…

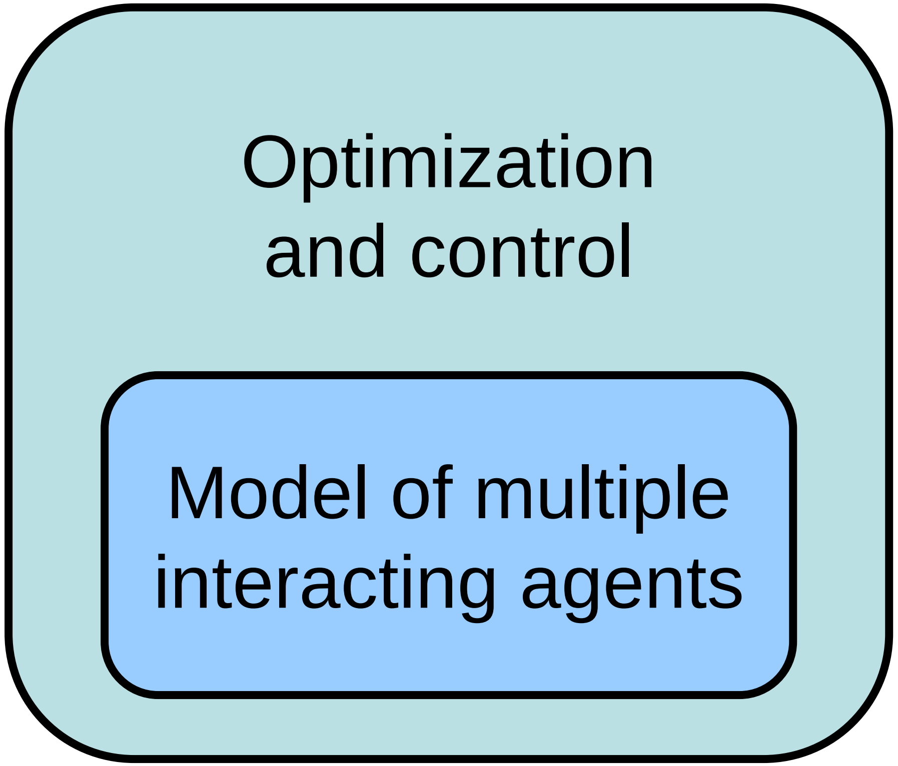
Big picture



Traditional game theory makes *the model* so complex…

…that *optimization and control* are intractable.

Big picture



MFE simplifies the model, so optimization and control become tractable.