Stochastic Dynamic Teams and Games with Asymmetric Information

TAMER BAŞAR
ECE, CAS, CSL and ITI, UIUC
basar1@illinois.edu

November 11, 2014

ITN SADCO
Rome, Italy
November 10-12, 2014

OUTLINE

- **Introduction to decision making in stochastic environments**
- **Dynamic teams, games, role of information, asymmetry**
- **Some caveats and counter-examples on existence and characterization of sols**
- **Existence and characterization of team-optimal policies under asymmetry**
- **Existence and characterization of Nash equilibrium policies under asymmetry**
- **Conclusions**

General Framework

INGREDIENTS

- *Uncertainty* (uncertain environment)
- *Decision makers* (DMs) (players, agents)
- *Perceptions* of DMs on uncertainty
- *Objective(s)*
- Description of interactions (underlying *network*)
- Common information / Private information

General Framework

- *Uncertainty* (uncertain environment); *decision makers* (DMs) (players, agents); *perceptions* of DMs on uncertainty; *objective(s)*; description of interactions (underlying *network*); common information
- DMs pick *policies* (decision laws, strategies) leading to *actions* that evolve over time

General Framework

- *Uncertainty* (uncertain environment); *decision makers* (DMs) (players, agents); *perceptions* of DMs on uncertainty; *objective(s)*; description of interactions (underlying *network*); common information
- DMs pick *policies* (decision laws, strategies) leading to *actions* that evolve over time
- Policies are constructed based on information received (active as well as passive) and guided by individual utility or cost functions over the DM horizon

General Framework

- *Uncertainty* (uncertain environment); *decision makers* (DMs) (players, agents); *perceptions* of DMs on uncertainty; *objective(s)*; description of interactions (underlying *network*); common information
- DMs pick *policies* (decision laws, strategies) leading to *actions* that evolve over time
- Policies are constructed based on information received (active as well as passive) and guided by individual utility or cost functions over the DM horizon

  - Single DM => stochastic control
  - Single objective => stochastic teams
  - Otherwise ZS or NZS games, with NE

Coupling of Information and Actions

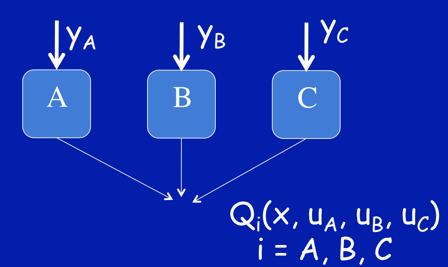
- Is the quality of active and relevant information received by a player affected by actions of other players ?
  - If no => underlying game (team) is generally "simple"
  - If yes => it is generally "difficult"

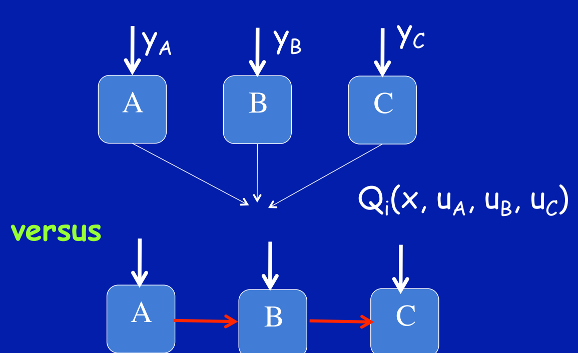
Asymmetric Information

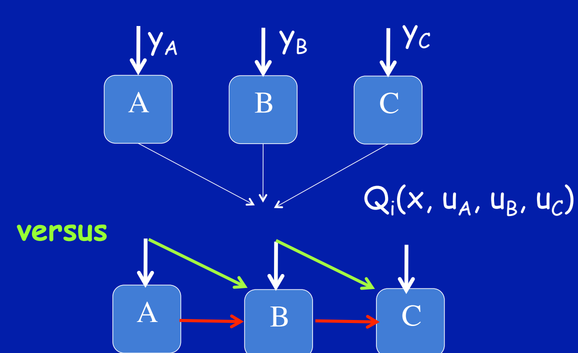
A stochastic decision problem is one with *asymmetric information*, if different decision units (players, agents, decision makers) acting at the same time instant do not have access to the same information (of relevance).

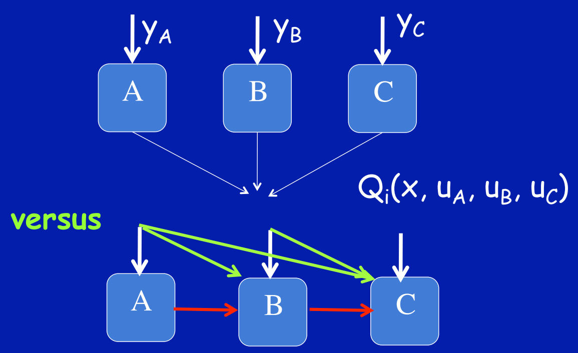
Non-classical Information

A stochastic decision problem is one with *non-classical information*, if a decision unit, **B**, that "*follows*" another one, **A**, does not have all the information acquired and used by **A**.









Questions / *Challenges*

- Existence of team-optimal (T-O) solutions
- Existence of Nash equilibrium (NE) policies (in dynamic stochastic games)
- Characterization of T-O solutions
- Characterization of NE policies

Questions / *Challenges*

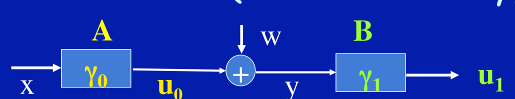
- Existence of team-optimal (T-O) solutions
- Existence of Nash equilibrium (NE) policies (in dynamic stochastic games)
- Characterization of T-O solutions
- Characterization of NE policies
- *Challenges*
  - Possibility of non-existence with asymmetry
  - Triple role (caution, probing, signaling)
  - Signaling (deception/threat) through action
  - Tension between signaling and optimization

Questions / *Challenges*

NEXT: An example exhibiting the subtleties, such as the significance of failure of continuity of information content of channels in the limit of a sequence of discretized problems

- *Challenges*
  - Possibility of non-existence with asymmetry
  - Triple role (caution, probing, signaling)
  - Signaling (deception/threat) through action
  - Tension between signaling and optimization

SCP with non-classical information (limited memory)



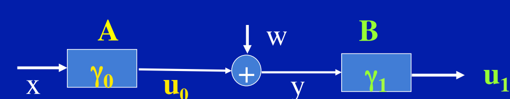
x, w ~ independent random variables

u_0, u_1 are real valued actions

$J(\gamma_0, \gamma_1) = E[\, Q(x, u_0, u_1) \mid \gamma_0, \gamma_1 \,]$

$J^* = \min \min J(\gamma_0, \gamma_1)$

Non-classical SCP (2-pt pmf's)

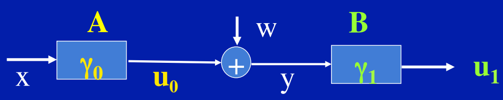


$x \sim \pm 1$ with equal prob; likewise w

$J(\gamma_0, \gamma_1) = E[k(u_0 - x)^2 + (u_0 - u_1)^2 \mid \gamma_0, \gamma_1],\ k > 0$

$J^* = \inf \inf J(\gamma_0, \gamma_1)$

Infimizing policies



$J^* = \inf \inf J(\gamma_0, \gamma_1) = 0$ – lowest possible

Consider the policies:

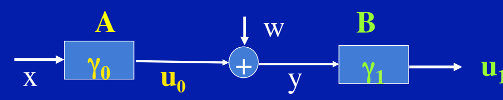
$\gamma_0(x) = x + \varepsilon\, \mathrm{sgn}(x)$

$\gamma_1(y) = 1+\varepsilon$, if $y=2+\varepsilon$ or ε

$= -1-\varepsilon$, if $y=-2-\varepsilon$ or $-\varepsilon$

➔ $J = k\, \varepsilon^2$ → 0 as $\varepsilon \to 0$

Infimizing policies



But, as $\varepsilon \to 0$, limiting policies do not lead to J = 0 !

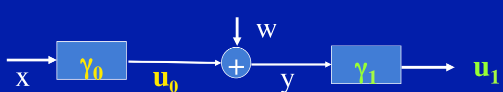
Loss of crucial information in the limit !

$\gamma_1(y) = 1+\varepsilon$, if $y=2+\varepsilon$ or ε

$= -1-\varepsilon$, if $y=-2-\varepsilon$ or $-\varepsilon$

➔ $J = k\, \varepsilon^2$ → 0 as $\varepsilon \to 0$

Same SCP but with Gaussian RVs
(Witsenhausen, 1968)

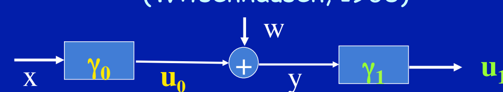


$x \sim N(0, \sigma_x^2)$ $w \sim N(0, \sigma_w^2)$

$Q_W(x, u_0, u_1) = k\,(u_0 - x)^2 + (u_0 - u_1)^2$

➡ optimal team solution exists, but its structure is not known

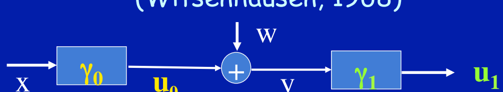
Same SCP but with Gaussian RVs
(Witsenhausen, 1968)



$Q_W(x, u_0, u_1) = k\,(u_0 - x)^2 + (u_0 - u_1)^2$

➡ optimal team solution exists, but its structure is not known
affine policies are not optimal

Same SCP but with Gaussian RVs
(Witsenhausen, 1968)

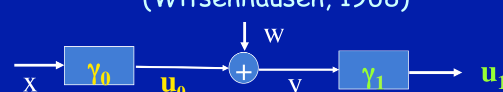


$Q_W(x, u_0, u_1) = k\,(u_0 - x)^2 + (u_0 - u_1)^2$

Note that if $u_0 = \gamma_0(x)$, best γ_1 is
$u_1 = \gamma_1(y) = E[\gamma_0(x) \mid y]$

Hence optimization is with respect to the probability distribution of u_0

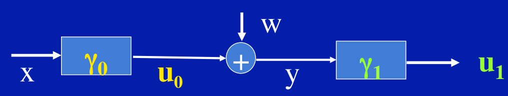
Same SCP but with Gaussian RVs
(Witsenhausen, 1968)



$Q_W(x, u_0, u_1) = k\,(u_0 - x)^2 + (u_0 - u_1)^2$

Original existence proof given by Witsenhausen involved calculus of variations type arguments, and does not extend to more general scenarios.

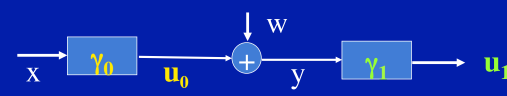
A class of non-linear policies



$Q_W(x, u_0, u_1) = k\,(u_0 - x)^2 + (u_0 - u_1)^2$

A policy pair that beats the best linear one:
$u_0 = \gamma_0(x) = \varepsilon \text{ sgn } (x) + \lambda\, x$
$u_1 = \gamma_1(y) = E[\varepsilon \text{ sgn } (x) + \lambda\, x \mid y]$
optimize wrt ε and λ

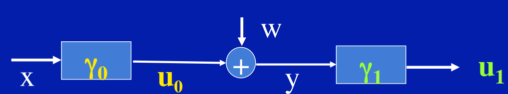
A class of non-linear policies



$Q_W(x, u_0, u_1) = k\,(u_0 - x)^2 + (u_0 - u_1)^2$

A more general one with improved perf:
$u_0 = \gamma_0(x) = \varepsilon \text{ Quant } (x) + \lambda\, x$
$u_1 = \gamma_1(y) = E[\varepsilon \text{ Quant } (x) + \lambda\, x \mid y]$
optimize wrt ε and λ for different Quant(ization) schemes (Bansal, TB '87)

A class of non-linear policies



$Q_W(x, u_0, u_1) = k\,(u_0 - x)^2 + (u_0 - u_1)^2$

$u_0 = \gamma_0(x) = \varepsilon \text{ Quant } (x) + \lambda\, x$
$u_1 = \gamma_1(y) = E[\varepsilon \text{ Quant } (x) + \lambda\, x \mid y]$
Best perf under linear could be arbitrarily *bad* against best performance under Quant
=> $\sup_{k,\sigma} [Q_W^{\text{linear*}} / Q_W^{\text{Quant*}}] = \infty$ (Mitter, Sahai '99)

BUT
Is the solution to Wits (68) piecewise affine ?

NO

It is strictly increasing with a real analytic left inverse

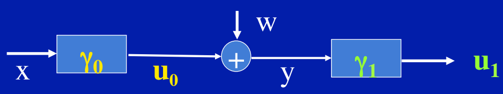
(Wu, Verdú – CDC'11)

NO
It is strictly increasing with a real analytic left inverse

(Wu, Verdú CDC'11)

Uses optimal transport theory:
Given prob measures P and Q, and cost function c: $R^2 \rightarrow R$, find inf over all joint distributions of X and Y, with marginals P and Q, of
$E[c(X,Y)]$
(Monge-Kantorovich)

Alternative existence proof

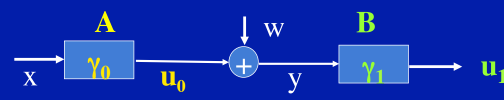


$$Q_W(x, u_0, u_1) = k\,(u_0 - x)^2 + (u_0 - u_1)^2$$

$E[c(x,u_0)] = E[Q_W] = k\,W_2(P,Q)^2 + \text{mmse}(Q)$ where P and Q are PDF's of x and u_0, and W_2 is the quadratic Wasserstein distance between P and Q.

For any P, ∃ minimizing Q, with "smooth" γ_0 .

Earlier general set-up with a different Q



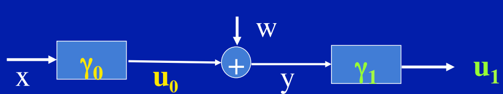
$x \sim N(0, \sigma_x^2)$ $\quad$ $w \sim N(0, \sigma_w^2)$

u_0, u_1 are real valued actions

$J(\gamma_0\,, \gamma_1) = E\,[\,Q(x, u_0, u_1) \mid \gamma_0\,, \gamma_1\,]$

$$Q(x, u_0, u_1) = k\,(u_0)^2 + (u_1 - x)^2$$

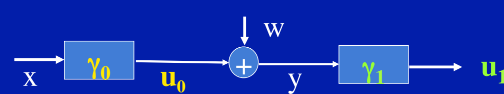
Gaussian Test Channel



$$Q(x, u_0, u_1) = k\,(u_0)^2 + (u_1 - x)^2$$

➡ optimal control law (encoder/decoder) exists, and is linear

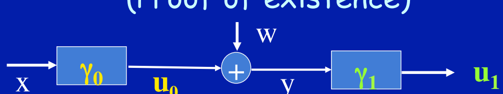
Generalized Gaussian Test Channel



$$Q(x, u_0, u_1) = k\,(u_0)^2 + (u_1 - x)^2 + b_0\, u_0\, x$$

➡ optimal control law (encoder/decoder) exists, and is linear

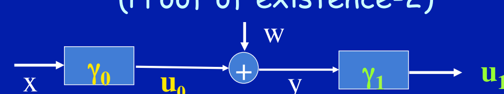
Generalized Gaussian Test Channel (Proof of existence)



$$Q(x, u_0, u_1) = k\,(u_0)^2 + (u_1 - x)^2 + b_0\, u_0\, x$$

$$E[Q] = F(\gamma_0, \gamma_1) \ge k\,\alpha + \beta + \inf_\gamma b_0\, E[\,\gamma_0(x)x] \ge k\,\alpha + \beta - \text{sgn}(b_0)\,\sigma_x \sqrt{\alpha}$$

Generalized Gaussian Test Channel (Proof of existence-2)

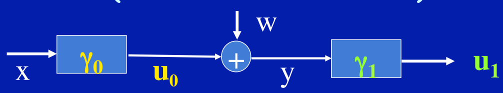


$$Q(x, u_0, u_1) = k\,(u_0)^2 + (u_1 - x)^2 + b_0\, u_0\, x$$

$$E[Q] = F(\gamma_0, \gamma_1) \ge k\,\alpha + \beta + \inf_\gamma b_0\, E[\,\gamma_0(x)x] \ge k\,\alpha + \beta - \text{sgn}(b_0)\,\sigma_x \sqrt{\alpha}$$

DPT: $I(X;Y) \ge I(X;U_1)$

Generalized Gaussian Test Channel (Proof of existence-3)



$Q(x, u_0, u_1) = k (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$

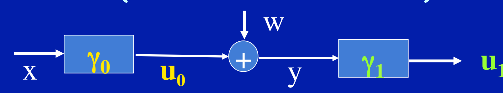
$E[Q]=F(\gamma_0, \gamma_1) \geq k\,\alpha + \beta + \inf_\gamma b_0 E[\gamma_0(x)x]$

$\geq k\,\alpha + \beta - \mathrm{sgn}(b_0)\, \sigma_x \sqrt{\alpha}$

$(1/2)\log(1+(\alpha/\sigma_w^2)) \geq I(X;Y) \geq I(X;U_1) \geq (1/2)\log(\sigma_x^2/\beta)$

$C(\alpha)$ $R(\beta)$

Generalized Gaussian Test Channel (Proof of existence-4)



$Q(x, u_0, u_1) = k (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$

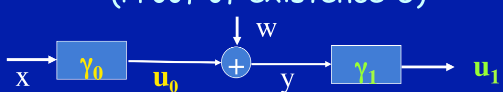
$E[Q]=F(\gamma_0, \gamma_1) \geq k\,\alpha + \beta + \inf_\gamma b_0 E[\gamma_0(x)x]$

$\geq k\,\alpha + \beta - \mathrm{sgn}(b_0)\, \sigma_x \sqrt{\alpha}$

$(1/2)\log(1+(\alpha/\sigma_w^2)) \geq I(X;Y) \geq I(X;U_1) \geq (1/2)\log(\sigma_x^2/\beta)$

==> $\beta \geq \sigma_w^2 \sigma_x^2 / (\sigma_w^2 + \alpha)$

Generalized Gaussian Test Channel (Proof of existence-5)



$Q(x, u_0, u_1) = k (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$

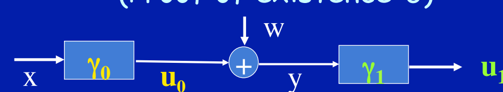
$E[Q]=F(\gamma_0, \gamma_1) \geq k\,\alpha + \beta + \inf_\gamma b_0 E[\gamma_0(x)x]$

$\geq k\,\alpha + \beta - \mathrm{sgn}(b_0)\, \sigma_x \sqrt{\alpha}$

==> $\beta \geq \sigma_w^2 \sigma_x^2 / (\sigma_w^2 + \alpha)$

Inequality is tight with $\gamma_0(x) = -\mathrm{sgn}(b_0)(\sqrt{\alpha} / \sigma_x)\, x$

Generalized Gaussian Test Channel (Proof of existence-6)



$Q(x, u_0, u_1) = k (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$

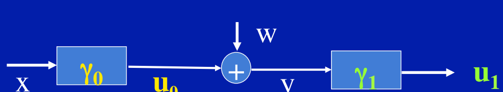
$E[Q]=F(\gamma_0, \gamma_1) \geq k\,\alpha + \beta - |b_0|\, \sigma_x \sqrt{\alpha}$

$\geq k\,\alpha + \sigma_w^2 \sigma_x^2 / (\sigma_w^2 + \alpha) - |b_0|\, \sigma_x \sqrt{\alpha}$

Obtain the α that minimizes the bound --> α^*

Then, $\gamma_0^*(x) = -\mathrm{sgn}(b_0)(\sqrt{\alpha^*} / \sigma_x)\, x$, $\gamma_1^*(y) = E[x|y]$

Generalized Gaussian Test Channel



$Q(x, u_0, u_1) = k (u_0)^2 + (u_1 - x)^2 + b_0 u_0 x$

One of the *few* instances when static/causal coding (and linear in this case) leads to attainment of equality in $C(\alpha) \geq R(\beta)$

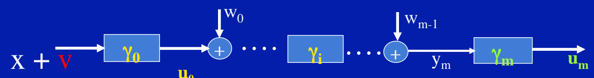
Relay Channels Multiple Serial Decision Units

$x \sim N(0, \sigma_x^2)$, $w_i \sim N(0, \sigma_w^2)$, $v \sim N(0, \sigma_v^2)$

$J(\gamma_0, \gamma_{[1,m]}) = E[Q(x, u_0, u_{[1,m]}) \mid \gamma_0, \gamma_{[1,m]}]$

$Q(x, u_0, u_{[1,m]}) = (u_m - x)^2 + \Sigma_{i=1}^m (u_{i-1})^2$

Multiple Serial Decision Units



$x \sim N(0, \sigma_x^2)$, $w_i \sim N(0, \sigma_w^2)$, $v \sim N(0, \sigma_v^2)$

$J(\gamma_0 , \gamma_{[1,m]}) = E [Q(x, u_0, u_{[1,m]}) \mid \gamma_0 , \gamma_{[1,m]}]$

➔For m > 1, affine laws no longer optimal for this extended GTC model

Does a general unifying existence proof exist?

One that would apply to

- Witsenhausen counter-example
- Gaussian test channel
- their multi-dimensional versions
- relay channels
- stochastic control problems with no memory or limited memory
- LQG teams with asymmetric information
- etc

Nov 11, 2014 ITN-SADCO

Answer is YES
(Gupta, Yüksel, TB -- CDC'14)

Assumptions on the dynamic M-agent T stage stochastic team problem:

- Sequential decision problem: Agent i observes and recalls at time t, $y_{(i,t)}$, and constructs $u_{(i,t)} = \gamma_{(i,t)}(y_{(i,t)})$
- Inf $J(\gamma)$ is finite (e.g. use zero control)
- Witsenhausen's static reduction holds

Nov 11, 2014 ITN-SADCO

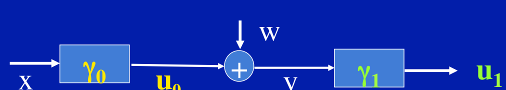
Answer is YES
(Gupta, Yüksel, TB -- CDC'14)

Static reduction:

- Convert M-agent team to MT-agent one where each agent acts only once
- Measure and cost transformation that turns the dynamic problem into static one, with independent measurements (measurements now enter into cost)

Nov 11, 2014 ITN-SADCO

Static Reduction for the C-example



$x \sim N(0, \sigma_x^2)$ $w \sim N(0, \sigma_w^2)$

$Q_W(x, u_0, u_1) = k (u_0 - x)^2 + (u_0 - u_1)^2$

➔ $\int P(dx)P(dy)\, Q_W(x, u_0, u_1)$

$\bullet \exp (u_0 (2y - u_0)/2)$

$x \sim N(0, \sigma_x^2)$ $y \sim N(0, 1)$

independent

Challenges in the Proof

- Extracting a convergent subsequence of the infimizing sequence
- Showing lower semi-continuity of J
- Making sure that informational constraints are preserved in the limit

Nov 11, 2014 ITN-SADCO

Approach to Overcome Challenges in the Proof

- With static reduction, redefine optimization over randomized strategies
- Because of independence of measurements information constraints are preserved
- Identify a compact set that contains the optimal solution (Markov's inequality)
- Use Blackwell's *irrelevant information theorem* to go to deterministic policies
- Invoke equivalence of reduced static and dynamic teams

Assumptions & Approach Apply to Various Settings

- Multi-dimensional Witsenhausen C-E
- Multi-dimensional Gaussian Test Channel
- Relay channel and multi-dimensional version
- LQG with static output feedback (solution is generally nonlinear)
- Countable/quantized observation spaces in static teams with observation sharing IS (uses a *lifting* technique)

Details in: A. Gupta, S. Yüksel & TB, "Existence of optimal strategies in a class of dynamic stochastic teams," CDC 2014

Also: Gupta, Yüksel, Langbort & TB (ACC'14)

And: Gupta, Yüksel, TB & Langbort, On the existence of optimal policies in a class of sequential dynamic teams," 2014 (arxiv)

Comprehensive coverage of analysis, optimization and performance of stochastic networked systems in

Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints -- Yüksel & TB

Birkhäuser, 2013

Stochastic Dynamic Games with Asymmetric Information

"Lifting" to a symmetric game utilizing the common information of the players

Formulation

- **Game G1**: n player game with state & measurement for Pi:

  $x^i_{t+1} = f^i_t(x^i_t, x_t, \underline{u}^i_t, w^i_t)$

  $y^i_t = h^i_t(x^i_t, \underline{x}_t, v^i_t), \quad i = 1, \ldots, n$
- Some sharing of past measurements and control actions by the players → $\mathcal{I}_{(i,t)}$
- Player i's policy at time t: $g_{(i,t)}: \mathcal{I}_{(i,t)} \to U_{(i,t)}$
- $J^i(g^1, \ldots, g^n) = E[\sum_{[1,T]} c^i(x^i_t, x_t, \underline{u}_t) | \underline{g}]$
- Interested in Nash equilibrium $\underline{g}^*$

Transformation to a game with symmetric information (say, n=2)

- Decompose into common and private information: $C_t = \mathcal{I}_{(1,t)} \wedge \mathcal{I}_{(2,t)}$, $P_{(i,t)} = \mathcal{I}_{(i,t)} \setminus C_t$
- Replace Pi with FPi who has access to C_t, and selects $\Gamma^i_t: P_{(i,t)} \to U^i$ according to φ^i_t (his strategy): $\Gamma^i_t = \varphi^i_t(C_t)$
- Then, control action is: $u^i_t = \Gamma^i_t(P_{(i,t)})$
- Cost: $L^i(\varphi^1, \ldots, \varphi^n) = E[\sum_{[1,T]} c^i(x^i_t, x_t, \underline{u}_t) | \underline{\varphi}]$

➔ ***Game G2***

Game G2: Replace Pi with FPi who has access to C_t, and selects $\Gamma^i_t: P_{(i,t)} \to U^i$ according to φ^i_t (his strategy): $\Gamma^i_t = \varphi^i_t(C_t)$
Then, control action is: $u^i_t = \Gamma^i_t(P_{(i,t)})$
Cost: $L^i(\varphi^1, \ldots, \varphi^n) = E[\sum_{[1,T]} c^i(x^i_t, x_t, \underline{u}_t) | \underline{\varphi}]$

Theorem: Let $\underline{g}$ be a NE for ***G1***. Define $\varphi^i_t(C_t) = g^i_t(\bullet, C_t)$
Then, $\underline{\varphi}$ is a NE for ***G2***. Conversely, if φ is a NE for ***G2***, $\underline{g}$ as above is a NE for ***G1***.

What does this lead to (conceptually)

- Since ***G2*** is a symmetric full information game, its NE can be obtained by backward induction (strongly-time consistent, sub-game perfect, etc), albeit with optimization on a function space $\{\Gamma^i_t\}$

➔ Markov perfect equilibrium (MPE) – generally one of many other NE which however are not STC

What does this lead to (conceptually)

- Since ***G2*** is a symmetric full information game, its NE can be obtained by backward induction (strongly-time consistent, sub-game perfect, etc), albeit with optimization on a function space $\{\Gamma^i_t\}$

➔ Markov perfect equilibrium (MPE) – generally one of many other NE which are not STC

- For ZS version, non-uniqueness is not an issue

What does this lead to (conceptually)

- Since ***G2*** is a symmetric full information game, its NE can be obtained by backward induction (strongly-time consistent, sub-game perfect, etc), albeit with optimization on a function space $\{\Gamma^i_t\}$

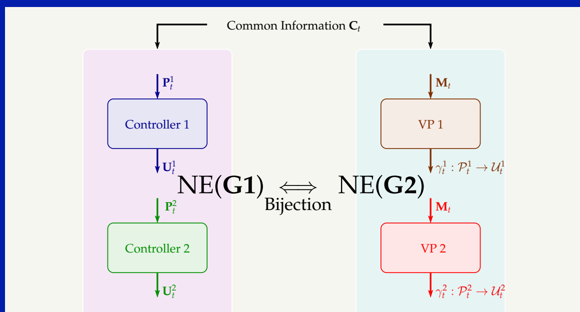
➔ Markov perfect equilibrium (MPE) – generally one of many other NE which are not STC

- For ZS version, non-uniqueness is not an issue
- NEED some conditions for this to go through

Assumptions / Approach

- Common information C_t does not contract
- Common information based common belief is independent of the policies of the players
- --
- Introduced common information based game (G2) on lifted strategy spaces
- MPE of G2 is CIMPE for G1

Relationship between G1 & G2



Assumptions / Approach / Result

- Common information C_t does not contract
- Common information based common belief is independent of the policies of the players
- ---
- Introduced common information based game (G2) on lifted strategy spaces as before
- MPE of G2 is CIMPE for G1

➔ For LQG, CIMPE is unique & CIMPE strategies are affine; computation involves solving linear eqs

Recent papers:

Nayyar & TB (CDC, 2012)
"Dynamic Stochastic Games with Asymmetric Information"

Nayyar, Gupta, Langbort, TB (TAC, 59(3), 2014)
"Common information based Markov perfect equilibria for stochastic games with asymmetric information: Finite games"

Gupta, Nayyar, Langbort, TB (SICON, 52(5), 2014)
"Common information based Markov perfect equilibria for Linear-Gaussian games with asymmetric information"

Recap

- We now have a general theory of existence of team-optimal solutions in stochastic dynamic teams with asymmetric and non-classical IS
- Caveats in computations based on discretization/quantization
- Lifting of SGs with asymmetric information to ones with symmetric information at the expense of increase in complexity
- Common information based MPE offers computational advantages

THANKS !