Introduction to State-based Potential Game

Yasuaki WASA
FL11-18-02
8, November, 2011

Outline

- Background
- Potential Game
  - Strategic Game and Features
  - Limitations
- State-based Potential Game
  - Framework
  - Utility Design
  - Learning Design
- Summary

Background

Multi-agent Systems

A system composed of multiple interacting autonomous agents

Applications

Sensor Network, Robotic Network
Power Network

Wireless Selection (access point/channel)


Power Control

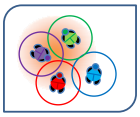
Cooperative Control

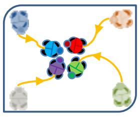
Each agent has decision-making components with limited local information

⇒ All agents seek to collectively accomplish a global objective

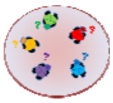
Distributed Approach [2]

- Game-Theoretic Control [1]


Sensor Coverage Problem


Consensus Problem


Centralized Approach


Distributed Approach

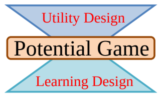
Background: Game Theoretic Approach

Game Theoretic Approach

Interaction between agents →	Agents are self-interested ⇒ Non-Cooperative Game
The solution of the coop. control problem ⇄	The convergence of the equilibrium of the game

Advantages

- Robustness to dynamic uncertainties
- Scalability and real-time adaptation
- Reduction of communication requirements



Design

- Utility Design (game design for the optimization problem)
- Learning Design (local decision-making rules)

Hierarchical decomposition between these designs

Potential Game

- the existence of Nash equilibrium invariably
- Local maxima of (global) objective function are Nash equilibria

Outline

- Background
- Potential Game
  - Strategic Game and Features
  - Limitations
- State-based Potential Game
  - Framework
  - Utility Design
  - Learning Design
- Summary

Strategic Game [1]

Players (agents) $i \in \mathcal{N}$ $\mathcal{N} = \{1, \cdots, n\}$

1. Action (sets) $a_i \in \mathcal{A}_i$ $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$

   Agent i's action a_i

   Other actions $a_{-i} = (a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_n) \in \mathcal{A}_{-i} = \prod_{j \neq i} \mathcal{A}_j$

   All actions $a = (a_i, a_{-i}) \in \mathcal{A}$

2. Utility function $U_i : \mathcal{A} \to \mathbb{R}$ $U = \{U_i\}_{i \in \mathcal{N}}$

   (same as local objective function)

   Each agent basically chooses the action a_i to maximize the own function U_i

[option] Restricted Action Set $\mathcal{R}_i : \mathcal{A}_i \to 2^{\mathcal{A}_i}$ $\mathcal{R} = \{\mathcal{R}_i(a_i)\}_{i \in \mathcal{N}}$

Agent i's restricted action set $\mathcal{R}_i(a_i)$

Assumptions

(Reversibility) $\forall i \in \mathcal{N}\ \forall a_i^1, a_i^2 \in \mathcal{A}_i$ s.t. $a_i^2 \in \mathcal{R}_i(a_i^1) \Leftrightarrow a_i^1 \in \mathcal{R}_i(a_i^2)$

(Feasibility) $\forall i \in \mathcal{N}\ \forall a_i^1, a_i^m \in \mathcal{A}_i$

$\exists\ a_i^1 \to a_i^2 \to \cdots \to a_i^m$ s.t. $a_i^l \in \mathcal{R}_i(a_i^{l-1}), \forall l \in \{2, \cdots, m\}$

Strategic Game $\Gamma = < \mathcal{N}, \mathcal{A}, U >$

Restricted Strategic Game $\Gamma_{res} = < \mathcal{N}, \mathcal{A}, U, \mathcal{R} >$

Potential Game (setup)

Game Strategic Game (Non-cooperative Game)

Objective Functions Global: Potential Function $\phi : \mathcal{A} \to \mathbb{R}$

Local: Utility Function $U_i : \mathcal{A} \to \mathbb{R}$

Equilibrium pure Nash Equilibrium

an action profile $a^* \in \mathcal{A}$ s.t. $\forall i \in \mathcal{N}$ $U_i(a_i^*, a_{-i}^*) = \max_{a_i \in \mathcal{A}_i} U_i(a_i, a_{-i}^*)$

If each agent changes an own action only, U_i cannot be more increased

Goal Both functions are maximized

In general, each agent just seeks to maximize U_i, while might even decrease ϕ

Constraint Condition

$\exists \phi$ s.t. $\forall i \in \mathcal{N}, \forall a_i, a_i' \in \mathcal{A}_i, \forall a_{-i} \in \mathcal{A}_{-i}$

$$U_i(a_i', a_{-i}) - U_i(a_i, a_{-i}) = \phi(a_i', a_{-i}) - \phi(a_i, a_{-i})$$

An increment of utility function An increment of potential function

Boarder definition: Generalized ordinal potential game (OPG)

$$U_i(a_i', a_{-i}) - U_i(a_i, a_{-i}) > 0 \quad \Rightarrow \quad \phi(a_i', a_{-i}) - \phi(a_i, a_{-i}) > 0$$

Any OPG has the finite improvement property

Potential Game (features)

Potential Game

Goal Both functions U_i, ϕ are maximized

Features

- the existence of Nash equilibrium invariably

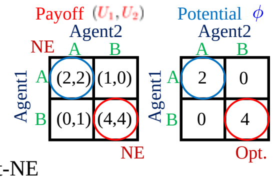
ϕ is maximum = Optimal NE

(global objective)

↑ ?

Sub-optimal NE

(local maxima)



➡ Learning Algorithm to lead to Opt-NE

Efficiency of games The social welfare function $W(a)$

Price of Stability (PoS) $PoS(\mathcal{G}) = \inf_{\Gamma \in \mathcal{G}} \left(\max_{a^* \in \mathcal{E}(\Gamma)} \frac{W(a^*)}{W(a^{opt})} \right) \leq 1$

(the worst-case efficiency of the best equilibrium across all games)

Price of Anarchy (PoA) $PoA(\mathcal{G}) = \inf_{\Gamma \in \mathcal{G}} \left(\min_{a^* \in \mathcal{E}(\Gamma)} \frac{W(a^*)}{W(a^{opt})} \right) \leq PoS(\mathcal{G})$

(the worst-case efficiency of any equilibrium across all games)

Distributed Welfare Game (DWG) [3]

A finite set of resources $\mathcal{R}$ ➡ An action set $\mathcal{A}_i \subseteq 2^{\mathcal{R}}$

Global welfare function $W(a) = \sum_{r \in \mathcal{R}} W^r(\{a\}_r)$ $W : \mathcal{A} \to \mathbb{R}$

$W^r : 2^{\mathcal{N}} \to \mathbb{R}_+$ The welfare function for resource $r \in \mathcal{R}$

$\{a\}_r := \{i \in \mathcal{N} : r \in a_i\}$ The subset of agents that are allocated to $r \in \mathcal{R}$

Utility function $U_i(a) = \sum_{r \in a_i} f^r(i, \{a\}_r)$ $U_i : \mathcal{A} \to \mathbb{R}$

$f := \{f^r(1, \{a\}_r), \cdots, f^r(n, \{a\}_r))\}_{r \in \mathcal{R}, \{a\}_r \subseteq \mathcal{N}}$ The distribution rule

How the welfare garnered from resource r is distributed across the players

Properties: $\forall i \in \mathcal{N}, \forall r \in \mathcal{R}, \forall \{a\}_r \subseteq \mathcal{N}$

(i) $f^r(i, \{a\}_r) \geq 0$ (ii) $i \notin \{a\}_r \Rightarrow f^r(i, \{a\}_r) = 0$ (iii) $\sum_i f^r(i, \{a\}_r) \leq W^r(\{a\}_r)$

To satisfy (iii) with equality ➡ Budget balanced distribution rule

Submodular [5]

A set valued function $W : 2^{\mathcal{A}} \to \mathbb{R}$ is submodular if

$$W(X) + W(Y) \geq W(X \cap Y) + W(X \cup Y), \ \forall X, Y \subseteq 2^{\mathcal{A}}$$

Same as, $W^r(X) + W^r(Y) \geq W^r(X \cap Y) + W^r(X \cup Y), \ \forall X, Y \subseteq \mathcal{N}$

Specifications of Utility Design

Utility Design

Desirability
- Existence and efficiency of NE
- Budget balance
- Computational Complexity
- Locality of information

Rule	PG	Equilibrium existence	Budget balanced	Tractable	Informational Requirement	PoS(G)	PoA(G)
ESU	○	○	○	○	Low	1/2	1/2
WLU	○	○	×	○	Middle	1	1/2
SVU	○	○	○	×	High	1/2	1/2

Limitation Theorem[3], [4]

Consider the set of distributed welfare games with submodular welfare functions and a budget balanced distribution rule that guarantees the existence of an equilibrium in all games. The price of stability is $\leq 1/2$

Furthermore, if the protocol is scalable and guarantees the existence of an equilibrium across all games then the price of stability is equal to $= 1/2$

Specifications of Learning Design

Learning Design

Desirability
- Asymptotic global behavior
- Equilibrium selection
- Informational dependencies
- Convergence rates

System
- Deterministic system
- Stochastic system

ex. PIPIP

$a(k-1), a(k-2)$

$U_i(k-1), U_i(k-2)$

$\varepsilon(k) = k^{(-1/N(D+1))}$

$\lim_{k \to \infty} P(z(k) \in \text{diag}(\zeta(\Gamma))) = 1$

(Limitation) [7]

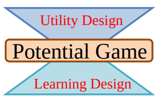
The result in [6] demonstrates that such levels of heterogeneity will not impact the asymptotic behavior of such learning algorithms.

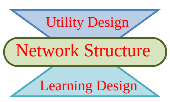
Network [9,10]

Behavior analysis depends on $G = (\mathcal{N}, \mathcal{E}(k))$

Environment

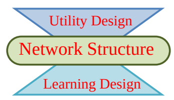
Design Structure [2] ➡ New Design Structure





Outline

- Background
- Potential Game
  - Strategic Game and Features
  - Limitations [3]
- State-based Potential Game
  - Framework [7]
  - Utility Design [8]
  - Learning Design [7]
- Summary



Extension to State-based Potential Game

Motivation

Potential Game (PG) [1]

To build upon existing game-theoretic results to better accommodate a broader class of cooperative control problems ➡ Maximize the functions

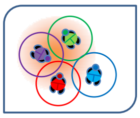
State-based Potential Game (SPG)

ex. Sensor Coverage, Resource Allocation

➡ Minimize their energy for actions

To use Cost function in application scenarios

➡ Minimize the functions



SPG Setup [8]

Optimization Problem $\min_{a_i} \phi(a_1, a_2, \cdots, a_n)$ s.t. $a_i \in \mathcal{A}_i,\ \forall i \in \mathcal{N}$ (1)

Action Renewal Law $a_i(k) = \mathcal{F}_i\left(\{a_j(k-1)\}_{j \in \mathcal{N}_i(k-1)}, x(k); U_i\right)$

(local control law: virtual payoff-based)

State Transition Rule $x(k+1) = F(a(k), x(k))$ (deterministic)

or $x(k+1) \sim P(a(k), x(k))$ (stochastic)

General Form [7] $\mathcal{T} = \{0, \cdots, k-1\}$

$a_i(k) = \mathcal{F}_i(\{a(\tau)\}_{\tau \in \mathcal{T}}, \{x(\tau)\}_{\tau \in \mathcal{T}}, x(k); U_i)$

Game Components

Strategic Game: PG	State-based Strategic Game
$\Gamma = < \mathcal{N}, \mathcal{A}, \{U_i\} >$	$\Gamma = < \mathcal{N}, \mathcal{A}, \mathcal{X}, \{U_i\}, * >$
$\mathcal{N}$ agents	$\mathcal{N}$ agents
$\mathcal{A}$ action sets	$\mathcal{A}$ action sets
	$\mathcal{X}$ state sets
	$* = \{F, P\}$ state transition rule; F (deterministic); P (stochastic)
U_i local func. (maximization)	U_i local func. (maximization)
$\phi(a)$ potential func. (welfare func.)	$\Phi(a, x)$ potential func.; $\phi(a)$ (welfare func.)

Learning Design

Utility Design

Note[7]:
The state is introduced as a coordinating entity used to improve system level behavior and can take on a variety of interpretations ranging from dynamics for equilibrium selection to the addition of dummy players in a strategic form game that are preprogrammed to behave in a set fashion.

Utility Design (II)

Approach

1. $U_i(a, x)$

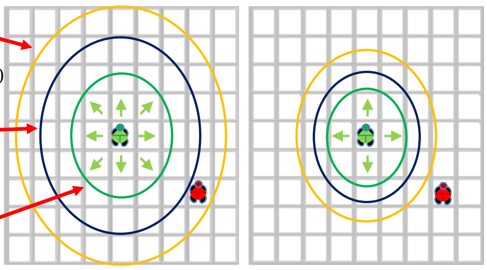
ex. x_i Residual Energy of agent i (buttery)

$U_i(a, x) = u_i(a, x) - f_i(x_i)$ (energy consumption)

Neighbor Area $\mathcal{N}_i(a, x)$ (agents' information with communication)

Sensing Area $\mathcal{D}_i(a_i, x_i)$ (environmental information) $W(q)$

Action Area $\mathcal{R}_i(a_i, x_i)$



$x_i \simeq E_i^{\max}$ $x_i \simeq 0$

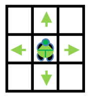
Utility Design (II)

Approach

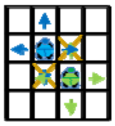
2. $\mathcal{R}_i(a_i, x)$

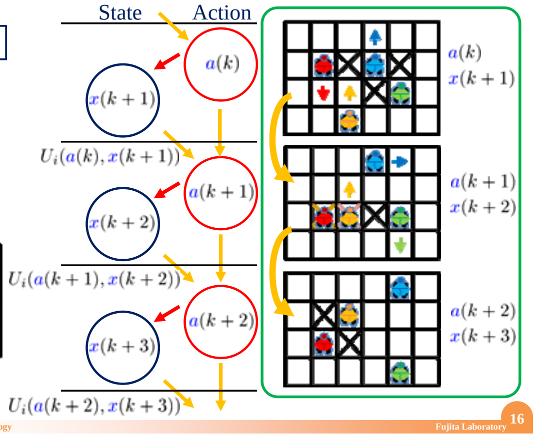
ex.

action set



constraints





Game Components

State-based Strategic Game	SPG
$\Gamma = < \mathcal{N}, \mathcal{A}, \mathcal{X}, \{U_i\}, * >$	$\Gamma = < \mathcal{N}, \mathcal{X}, \mathcal{A}, \{J_i\}, * >$
$\mathcal{N}$ agents	$\mathcal{N}$ agents
$\mathcal{A}$ action sets	$\mathcal{X} = \mathcal{A} \times \mathcal{X}$ state sets
$\mathcal{X}$ state sets	$\mathcal{A} = \Pi \times \mathcal{X}$ action sets
	Π control strategy set
$* = \{F, P\}$ state transition rule; F (deterministic); P (stochastic)	$* = \{F, P\}$ state transition rule; F (deterministic); P (stochastic)
U_i local func. (maximization)	J_i local func. (minimization)
$\Phi(a, x)$ potential func.; $\phi(a)$ (welfare func.)	$\Phi(x, a)$ potential func.; $\phi(a)$ objective func.
Learning Design; Utility Design	Utility Design; Learning Design

Notations for Utility Design (III) [8]

$\mathcal{N}_i(k) \subset \mathcal{N}$ Neighbor sets of agent i $\mathcal{N}_i(k) := \{j \in \mathcal{N} : (i, j) \in \mathcal{E}(k)\}$

$x = (v, e) \in \mathcal{X}$ (finite) State (space)

$v = (v_1, \cdots, v_n) \in \mathbb{R}^n$ The profile of values $v_i \in \mathcal{V}_i$

$e = (e_1, \cdots, e_n)$ The profile of estimation terms

$e_i = (e_i^1, \cdots, e_i^n) \in \mathbb{R}^n$ Agent i 's estimation for the joint action profile

$a = (\hat{v}, \hat{e}) \in \mathcal{A}$ Action (change terms)

$\hat{v} = (\hat{v}_1, \cdots, \hat{v}_n) \in \mathbb{R}^n$ The change in the profile of values

$\hat{e} = (\hat{e}_1, \cdots, \hat{e}_n)$ The change in the profile of estimation terms

$\hat{e}_i = (\hat{e}_i^1, \cdots, \hat{e}_i^n) \in \mathbb{R}^n$ The change in the agent i's estimation

$\hat{e}_{i \to j}^k \in \mathbb{R}$ The estimation value that i passes to j regarding to the value of k

$\hat{e}_{i \to j}^k = 0\ \forall j \notin \mathcal{N}_i, k \in \mathcal{N}$

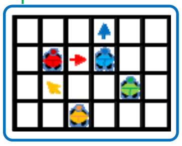
$J_i(x, a)$ Utility Function $\Phi(x, a)$ Potential Function $\phi(v)$ Objective Function

- Value v_i is defined with scalar variable not vector variable. $v_i \in \mathbb{R}^1$
  If it is a vector variable, do transition rules hold?
- How is the change of values and estimation terms calculated actually?
- Is the objective function $\phi(v)$ equal to the welfare function?

Utility Design (III)

ex. Based on [8]

Output



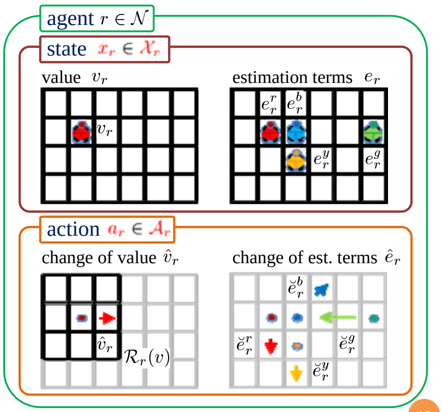
State $x = (v, e)$

$a \to v$

$x \to e$

Action $a = (\hat{v}, \hat{e})$

It is possible that the estimation terms e_i don't exist over $\mathcal{V}$



Functions for Utility Design (III) [8]

Estimation profile

Initialization $e_i^k(0) = \begin{cases} n \cdot v_i(0) & k = i \\ 0 & \text{otherwise} \end{cases}$

constraint $\sum_{i \in \mathcal{N}} e_i^k(t) = n \cdot v_k(t) \quad \forall t \geq 1, \forall k \in \mathcal{N}$

State transition function

Action value $f_i^v(x, a) = v_i + \hat{v}_i$

Estimation value $f_{i,k}^e(x, a) = e_i^k + n\delta_i^k \hat{v}_i + \check{e}_i^k$

Estimation error $\check{e}_i^k := \sum_{j \in \mathcal{N}_i} \hat{e}_{j \to i}^k - \sum_{j \in \mathcal{N}_i} \hat{e}_{i \to j}^k$

Cost function $J_i(x, a) = J_i^\phi(x, a) + \alpha \cdot J_i^e(x, a)$

$J_i^\phi(x, a) = \sum_{j \in \mathcal{N}_i} \phi(\tilde{e}_j^1, \cdots, \tilde{e}_j^n)$

$J_i^e(x, a) = \sum_{j \in \mathcal{N}_i} \sum_{k \in \mathcal{N}} (\tilde{e}_i^k - \tilde{e}_j^k)^2$

where $\tilde{x} = (\tilde{v}, \tilde{e}) = F(x, a)$

Potential function $\Phi(x, a) = \Phi^\phi(x, a) + \alpha \cdot \Phi^e(x, a)$

$\Phi^\phi(x, a) = \sum_{i \in \mathcal{N}} \phi(\tilde{e}_i^1, \cdots, \tilde{e}_i^n)$

$\Phi^e(x, a) = \frac{1}{2} \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_i} \sum_{k \in \mathcal{N}} (\tilde{e}_i^k - \tilde{e}_j^k)^2$

Note: the other expression about $J_i^e(\cdot, \cdot)$ exists[9]. They are two different things.

Theorem about Utility Design (III) [8]

Theorem[8]

Suppose the objective $\phi(\cdot)$ *and the designed communication graph* $G = (\mathcal{N}, \mathcal{E})$ *satisfies* ***at least one*** *of the following conditions*

(i) $\phi(\cdot)$ is convex over $\mathcal{V} \subset \mathbb{R}^n$ and G is non-bipartite

(ii) $\phi(\cdot)$ is convex over $\mathcal{V} \subset \mathbb{R}^n$ and $n = |\mathcal{N}|$ is odd

(iii) $\phi(\cdot)$ is convex over $\mathbb{R}^n$ and $\exists i, j \in \mathcal{N}, \ |\mathcal{N}_i| \neq |\mathcal{N}_j|$

Then the state action pair $[x, a]$ *is a recurrent state equilibrium in* G *if and only if the following conditions are satisfied:* $x(k+1) = F(x(k), a(k))$

(a) Value profile: v is an optimal solution for the problem(1)

(b) Estimation profile: $e_i^k = v_k, \ \forall i, k \in \mathcal{N}$

(c) Change in Value profile: $\hat{v}_i = 0, \ \forall i \in \mathcal{N}$

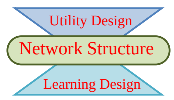
(b) Change in Estimation profile: $\check{e}_i^k = 0, \ \forall i, k \in \mathcal{N}$ (balanced)

Note:

1. A non-bipartite graph is a graph that contains an odd-length cycles
2. Each $\mathcal{V}_i$ is convex set
3. In case of the time-variant connected communication graph $G(t) = (\mathcal{N}, \mathcal{E}(t))$, if the condition (i) is satisfied, the theorem holds. $x(k+1) \sim P(x(k), a(k))$

Outline

- Background
- Potential Game
  - Strategic Game and Features
  - Limitations [3]
- State-based Potential Game
  - Framework [7]
  - Utility Design [8]
  - Learning Design [7]
- Summary



Game Components

State-based Strategic Game

$\Gamma = <\mathcal{N}, \mathcal{A}, \mathcal{X}, \{U_i\}, * >$

$\mathcal{N}$ agents

$\mathcal{A}$ action sets

$\mathcal{X}$ state sets

$* = \{F, P\}$

state transition rule

F (deterministic)

P (stochastic)

U_i local func. (maximization)

$\Phi(a, x)$ potential func.

$\phi(a)$ (welfare func.)

Learning Design

Utility Design

Learning Design: Binary Log-Linear Learning [7]

PG

1) One player $i \in \mathcal{N}$ is randomly (uniformly) chosen. $a_{-i}(t) = a_{-i}(t-1)$
2) Select one trial action $\hat{a}_i$ uniformly from $\mathcal{A}_i \backslash \{a_i(t-1)\}$
3) Select action from the following actions with probability $\beta > 0$

$$a_i^{tmp} \leftarrow \begin{cases} a_i(t-1), & \text{w.p.} \ \dfrac{e^{\beta U_i(a(t-1))}}{e^{\beta U_i(a(t-1))} + e^{\beta U_i(a_i, a_{-i}(t-1))}} \\ \hat{a}_i, & \text{w.p.} \ \dfrac{e^{\beta U_i(a_i, a_{-i}(t-1))}}{e^{\beta U_i(a(t-1))} + e^{\beta U_i(a_i, a_{-i}(t-1))}} \end{cases}$$

SPG

1) Choose randomly a player $i \in \mathcal{N}$ with probability q

$q_i > 0, \ \sum_{i \in \mathcal{N}} q_i < 1, \ q_0 = 1 - \sum_{i \in \mathcal{N}} q_i > 0$

Note: There are the cases where no player is selected

2) 3) If a player i is selected, select an action in the same way as PG

$U_i(a(t-1)) \to U_i(a(t-1), x(t))$

4) The ensuing state $x(t+1)$ is chosen randomly according the transition probability $P(a(t), x(t))$

Note: In terms of the process of the proof, it is hard to use the deterministic transition function $F(a(t), x(t))$

Note[7]: Reversibility is not satisfied in our setup

Learning Design and Constraints [7]

Tokyo Institute of Technology

Theorem[7]

Let $\Gamma = \{\mathcal{N}, \mathcal{A}, \mathcal{X}, \{U_i\}, P\}$ *be an ordinal state based potential game with a state invariant potential function* $\phi : \mathcal{A} \to \mathbb{R}$ *that satisfies the following three conditions* $\forall a \in \mathcal{A}, \forall x \in \mathcal{X}$

(i) The action invariant state transition process $P(a, \cdot)$ is aperiodic and irreducible over $\mathcal{X}(a)$

(ii) $\forall i \in \mathcal{N}, \forall a_i' \in \mathcal{A}_i \;\; U_i(a_i', a_{-i}, x) - U_i(a, x) \leq \phi(a_i', a_{-i}) - \phi(a)$

(iii) $\forall i \in \mathcal{N}, \forall a_i' \in \mathcal{A}_i, \exists x' \in \mathcal{X}(a)$ s.t.
$$U_i(a_i', a_{-i}, x') - U_i(a, x') = \phi(a_i', a_{-i}) - \phi(a)$$

For such state-based potential games, the process log-linear learning guarantees that an action state pair $[a^*, x^*]$ *is stochastically stable if and only if the action profile* $a^* = \arg\max_a \phi(a)$ *and state* $x^* = \mathcal{X}(a^*)$

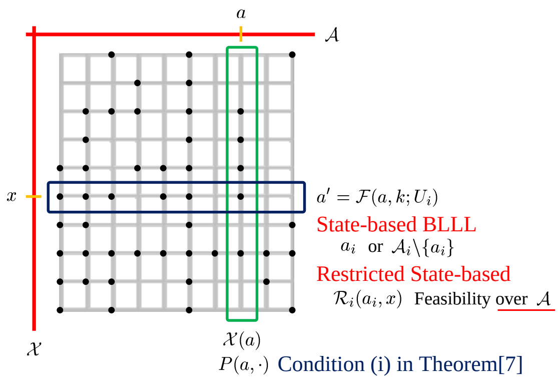
(ii) and (iii) provide a relaxation to the PG structure by relaxing the equality constraint

conditions (i) Regular Perturbation $\quad q_0$, Binary: Resistance of transition
(ii) Resistance and $\phi(= W)$ $\quad$ (iii) Resistance of feasible transition path

When each agent selects an action which depends on the current state, it is logically impossible that he select (go back) the previous action which depends on the previous state from example of Utility Design (II).

[Appendix] Aperiodic and Irreducible

Tokyo Institute of Technology



$P(a, \cdot)$ Condition (i) in Theorem[7]

Outline

Tokyo Institute of Technology

- Background
- Potential Game
  - Strategic Game and Features
  - Limitations [3]
- State-based Potential Game
  - Framework [7]
  - Utility Design [8]
  - Learning Design [7]
- **Summary**

Summary

Tokyo Institute of Technology

Conclusion

- Limitation of Potential Game
- Framework of State-based Potential Game

Future Work

- State-based Learning Design
  - extension to PIPIP (meaningless?)
  - application to payoff-based algorithm (lead to Optimal Equilibrium)

$$a_i(k) = \mathcal{F}_i\left(\{a_i(k-1), U_i(a(k-1), x(k))\}, x(k)\right)$$

- State-based Potential Game
  - strictly framework (practical usability)
- Application: State-based Utility Design
  - [Target] Robotic Network, Power Network, (Camera Network?)
  - [Scenario] coverage with collision avoidance,
    resource allocation with environmental change,
    wind farm optimal control [11]

Reference

Tokyo Institute of Technology

Potential Game (PG)

[1] J. R. Marden, G. Arslan and J. S. Shamma, "Cooperative Control and Potential Games," *IEEE Transacions on Systems, Man and Cybernetics*, Vol. 39, No. 6, pp. 1393-1407, 2009.

Framework of PG

[2] R. Gopalakrishnan, J. R. Marden and A. Wierman, "An Architectural View of Game Theoretic Control," *Proc. of ACM Hotmetrics: Third Workshop on Hot Topics in Measurement and Modeling of Computer Systems*, 2010.

Limitations of PG

[3] J. R. Marden and A. Wierman, "Overcoming Limitations of Game-theoretic Distributed Control," *Proc. of 48th IEEE Conference on Decision and Control*, pp. 6466-6471, 2009.

[4] J. R. Marden and A. Wierman, "Limitations of Utility Design for Multiagent Systems," *under submission*, 2011.

[5] J. R. Marden and A. Wierman, "Distributed Welfare Games," *Operations Research special issue on Computational Economics*, submitted, 2008.

[6] H. P. Young, *Strategic Learning and Its Limits*, Oxford University Press, 2004.

Reference

Tokyo Institute of Technology

State-based Potential Game (SPG)

[7] J. R. Marden, "State Based Potential Games," *IEEE Transactions on Automatic Control*, under review, 2011.

Utility Design of SPG (SPG)

[8] N. Li and J. R. Marden, "Designing Games for Distributed Optimization," *Proc. of 50th IEEE Conference on Decision and Control, submitted*, 2011.

[9] N. Li and J. R. Marden, "Designing Games for Distributed Optimization with a Time-Varying Communication Graph," *International Conference on Network Games, Control and Optimization*, submitted, 2011.

Network Formation of SPG (SPG)

[10] G. C. Chasparis and J. S. Shamma, "Network Formation: Neighborhood Structures, Establishment Costs, and Distributed Learning," *IEEE Transactions on Automatic Control, submitted*, 2011.

Wind Farm

[11] J. R. Marden, S. D. Ruben and L. Y. Pao, "Surveying Game Theoretic Approaches for Wind Farm Optimization," *30th ASME Wind Symposium, submitted* , 2011.

Appendix

Welfare Function and Potential Function

Utility Function $U_i(a) = \sum_{r \in a_i} f^r(i, \{a\}_r)$

Easiest Distributed Rule $U_i(a) = \frac{1}{n} W(a)$

Equally Shared Utility (ESU) $I\{\cdot\}$ The usual indicator function

$$f^r(i, \{a\}_r) := \frac{1}{\sum_{j \in \mathcal{N}} I\{r \in a_j\}} W^r(a)$$

In general, such a design cannot guarantee the existence of an equilibrium

Wonderful Life Utility (WLU)

$$f^r(i, \{a\}_r) := W^r(\{a\}_r) - W^r(\{a\}_r \backslash \{i\})$$

$$U_i(a) = W(a) - W(\emptyset, a_{-i})$$

$\Rightarrow$ $\phi(a) \equiv W(a)$ (not budget-balanced) $\sum_{i \in \mathcal{N}} f^r(i, \{a\}_r) \leq W^r(\{a\}_r)$

Shapley Value Utility (SVU)

$$f^r(i, \{a\}_r) := \sum_{S \subseteq \{a\}_r : i \in S} \omega_s (W^r(S) - W^r(S \backslash \{i\}))$$

$$\omega_s := \frac{(|\{a\}_r| - 2)!(|S| - 1)!}{|\{a\}_r|!}$$

$\Rightarrow$ $\phi(a) \not\equiv W(a)$ (budget-balanced) $\sum_{i \in \mathcal{N}} f^r(i, \{a\}_r) = W^r(\{a\}_r)$

Utility Design (I) Priority-based distribution Rule [3]

Equilibrium State-based Nash Equilibrium

$[a^*, x^*]$ s.t. $\forall i \in \mathcal{N}\ \forall x' \sim P(a^*, x^*)\ U_i(a_i^*, a_{-i}^*, x') = \max_{a_i \in \mathcal{A}_i} U_i(a_i, a_{-i}^*, x')$

Constraint Condition of SPG

$\exists \phi(a)$ s.t. $\forall [a, x] \in \mathcal{A} \times \mathcal{X}, \forall i \in \mathcal{N}, \forall a_i' \in \mathcal{A}_i$

$U_i(a_i', a_{-i}, x) - U_i(a_i, a_{-i}, x) > 0 \Rightarrow \phi(a_i', a_{-i}) - \phi(a_i, a_{-i}) > 0$

Concept WLU+SVU (+state-based)

Utility function $U_i(a', x) = \boldsymbol{E}_{P(a', x)} V_i(a', x')$

$V_i(a, x) = \sum_{r \in a_i} (W^r(\{\bar{x}_i\}_r) - W^r(\{\bar{x}_i\}_r \backslash \{i\}))$

$\{\bar{x}_i\}_r := \{j \in \mathcal{N} : x_j^r \leq x_i^r\}$

First in first out (FIFO) $a(t) = a(t-1) \Rightarrow x(t+1) = x(t)$

Theorem[3]

Consider any distributed welfare game with submodular welfare functions, priority-based utility functions, and FIFO state dynamics. The resulting game is a state-based potential game with potential function $\phi = W$ and a price of stability is 1. (Note: this design also satisfies budget balanced)

Definition of State-based Equilibrium [7]

State Invariant Equilibrium

An action profile a^* is a state invariant equilibrium if $\forall i \in \mathcal{N}, \forall x \in \mathcal{X}$

$$U_i(a_i^*, a_{-i}^*, x) = \max_{a_i \in \mathcal{A}_i} U_i(a_i, a_{-i}^*, x)$$

Recurrent State Equilibrium

The action state pair $[a^*, x^*]$ is a recurrent state equilibrium if the following two conditions are satisfied:

(i) $\forall i \in \mathcal{N}, \forall x \in \bar{X}(a^*, x^*)$ $U_i(a^*, x) = \max_{a_i \in \mathcal{A}_i} U_i(a_i, a_{-i}^*, x)$

(ii) $\forall x \in \bar{X}(a^*, x^*)$ $x^* \in \bar{X}(a^*, x)$

Notation

$\bar{X}(a^0, x^0) := \{x^0, x^1, x^2, \cdots\} \subseteq \mathcal{X}$ where $x^{k+1} \sim P(a^0, x^k)\ \forall k \in \{0, 1, \cdots\}$

or $x^{k+1} = F(a^0, x^k)\ \forall k \in \{0, 1, \cdots\}$

the set of reachable states by an action invariant state trajectory for $[a^0, x^0]$

$\sim$ The state is selected randomly according to the probability distribution $P(\cdot, \cdot)$

$\bar{P}(a, x) := \{x' \in \mathcal{X} : x' \sim P(a, x)\}$ The set of states in the support of $P(a, x)$

When we express an equilibrium in the SPG, using "Nash" is strange.

Definition of State-based Potential Game

[7] *A state based game Γ is an (exact) state based potential game if there exists a potential function $\Phi(\cdot, \cdot)$ that satisfies the following two properties:* $\forall a \in \mathcal{A}, \forall x \in \mathcal{X}(a)$

(i) $\forall i \in \mathcal{N}, \forall a_i' \in \mathcal{A}_i$ $U_i(a_i', a_{-i}, x) - U_i(a, x) = \Phi(a_i', a_{-i}, x) - \Phi(a, x)$

(ii) $\forall x \in \bar{P}(a, x)$ $\Phi(a, x') \geq \Phi(a, x)$

Ordinal SPG (i) ☞ $U_i(a_i', a_{-i}, x) - U_i(a, x) > 0 \Rightarrow \Phi(a_i', a_{-i}, x) - \Phi(a, x) > 0$

Complete SPG $\forall a \in \mathcal{A}, \forall x \in \mathcal{X}(a)$ ☞ $\forall [a, x] \in \mathcal{A} \times \mathcal{X}$

Lemma *A recurrent state equilibrium exists in any Ordinal SPG*

[8] *A (deterministic) state based game Γ is a (deterministic) state based potential game if there exists a potential function $\Phi(\cdot, \cdot)$ that satisfies the following two properties:* $\forall a \in \mathcal{A}, \forall x \in \mathcal{X}(a)$

(i) $\forall i \in \mathcal{N}, \forall a_i' \in \mathcal{A}_i$ $U_i(a_i', a_{-i}, x) - U_i(a, x) = \Phi(a_i', a_{-i}, x) - \Phi(a, x)$

(ii) $\forall [a, x] \in \mathcal{A} \times \mathcal{X}$ $\Phi(a, x) = \Phi(\tilde{x}, \mathbf{0})$ where $\tilde{x} = F(a, x)$

Proposition *Given a (deterministic) state based game Γ, if a state action pair $[x^*, a^*]$ satisfies for $a^* = \arg\max_a \Phi(x^*, a)$ and $x^* = F(x^*, a^*)$ then its pair a recurrent state equilibrium.*

[Appendix] Proof Method of Stochastic Stability

(Perturbed) Markov Process $\{P_t^0\}$ $(\{P_t^\varepsilon\})$ over a finite state space $\mathcal{X}$

Transition probability P_{xy}^ε from $x \in \mathcal{X}$ to $y \in \mathcal{X}$ along with $\{P_t^\varepsilon\}$

A family of stochastic processes $\{P_t^\varepsilon\}$ is a regular perturbation of $\{P_t^0\}$ if

(i) $\exists \varepsilon^* > 0, \forall \varepsilon \in (0, \varepsilon^*]$ s.t. $\{P_t^\varepsilon\}$ is aperiodic and irreducible

(ii) $\forall x, y \in \mathcal{X}$ s.t. $\lim_{\varepsilon \to 0} P_{xy}^\varepsilon = P_{xy}^0$

(iii) $\exists \varepsilon, P_{xy}^\varepsilon > 0 \Rightarrow \exists \chi(x \to y) \in \mathbb{R}_+$ s.t. $\lim_{\varepsilon \to 0} \frac{P_{xy}^\varepsilon}{\varepsilon^{\chi(x \to y)}} \in (0, \infty)$

Theorem H. P. Young

(assumption) $\{P_t^\varepsilon\}$ is a regular perturbation of $\{P_t^0\}$

(i) $\exists \lim_{\varepsilon \to 0+} \mu(\varepsilon)$ and the limiting distribution $\mu(0)$ is a stationary distribution of $\{P_t^0\}$

(ii) the stochastically stable states are contained in the recurrent communication classes with minimum stochastic potential

Theorem Freidlin and Wentzell

(assumption) P is an irreducible Markov process on $\mathcal{X}$

Its stationary distribution μ has $\mu(x) = \frac{\nu(x)}{\sum_{w \in \mathcal{X}} \nu(w)}$ where $\nu(x) = \sum_{T \in \mathcal{T}_x} P(T)$