Introduction to potential game

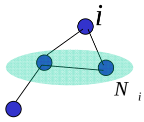
Goto Tatsuhiko
FL 09 - 14 -1
9th,Novenver,2009

outline

- Purpose of this paper
- Game theoretic approach
- Consensus Modeled as Potential Game
- Spatial adaptive play
- Restrictive Spatial Adaptive Play
- Simulation and application

Review of definition

$P = \{P_1, \cdots, P_n\}$ Group of players



Interacts with neighbors N_i

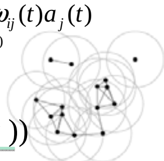
consensus

$a_1 = a_2 = \cdots = a_n$ a_i :State of player p_i

General consensus algorithm $a_i(t+1) = \sum_{p_j \in N_i(t)} \omega_{ij}(t) a_j(t)$

ω_{ij} :relative weight

Time varying directed graph $G(V, E(t))$



Purpose of this paper

Cooperative Control and Potential Games*
Jason R. Marden† Gürdal Arslan‡ Jeff S. Shamma§

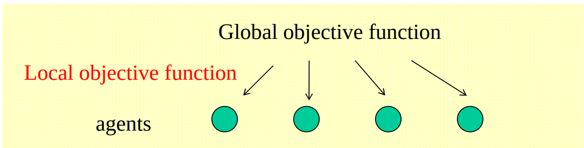
Little research of games for cooperative control

Establish a relationship
between cooperative control and game theory

1. Model the consensus problem as a potential game
2. Learning algorithm for potential game
3. New class of games "sometimes weakly acyclic game"

(ex)
Theoretical consensus accommodate obstructions

Game theoretic approach



A_i :Action set $A = \prod_{p_i \in p} A_i$:set of joint action

$a_{-i} = (a_1, a_2, \cdots, a_{i-1}, a_{i+1}, \cdots, a_n)$

joint action $a = (a_i, a_{-i})$

Local objective function $U_i : A \to R$

Control design
1. Designing the player objective function
2. Learning dynamics (repeated game)
   (ex)single stage memory dynamics

Potential game

Global planner $\phi : A \to R$
(potential function)

aligned

Player's objective function

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

Changing in the player's objective function
=
Changing in the potential function

Nash Equilibrium

For all players $p_i \in P$

$$U_i(a_i^*, a_{-i}^*) = \max_{a_i \in A_i} U_i(a_i, a_{-i}^*)$$

⇨ $a^* \in A$ is pure Nash equilibrium

In potential games

any action profile maximizing the potential function is a pure Nash equilibrium

→ Exists at least one such equilibrium

There may also exist suboptimal pure Nash equilibrium

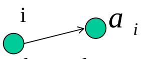
→ Don't maximizing the potential function

Consensus Modeled as Potential Game

1. Establish global objective function that captures the notion of consensus

$$\phi(a) = -\sum_{p_i \in P} \sum_{p_j \in N_i} \frac{\|a_i - a_j\|}{2}$$

⇨ $\phi(a) = 0 \Leftrightarrow a_1 = \cdots = a_n$



Player's action set a_i ⟶ Location that a player could select

Potential game's graph ⇨ Time –invariant and undirected

Consensus Modeled as Potential Game

2. Assign each player an objective function

first approach $U_i(a) = \phi(a)$ $\quad \phi(a) = -\sum_{p_i \in P} \sum_{p_j \in N_i} \frac{\|a_i - a_j\|}{2}$

Require to observe the decision of all players

⇨ May be infeasible

second approach

Captures the player's marginal contribution to the potential function

$$U_i(a) = -\sum_{p_j \in N_i} \|a_i - a_j\|$$

Wonderful life utility (WLU)

⇩

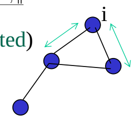
Only depend on the Neighbor's

claim

Claim 2.1

Player's objective function $U_i(a) = -\sum_{p_j \in N_i} \|a_i - a_j\|$ constitute a potential game with the potential function $\phi(a) = -\sum_{p_i \in P} \sum_{p_j \in N_i} \frac{\|a_i - a_j\|}{2}$

(Time-invariant interaction graph and undirected)



(Proof) $\phi(a) = -\sum_{p_i \in P} \sum_{p_j \in N_i} \frac{\|a_i - a_j\|}{2}$

undirected ⇨ $\phi(a) = -\sum_{p_j \in N_i} \|a_i - a_j\| - \sum_{p_j \neq p_i} \sum_{p_k \in N_j \setminus p_i} \frac{\|a_j - a_k\|}{2}$ ← not depend on i

action change $p_i : a_i^1 \rightarrow a_i^2$

$$U_i(a_i^2, a_{-i}) - U_i(a_i^1, a_{-i}) = \sum_{p_j \in N_i} -\|a_i^2 - a_j\| + \|a_i^1 - a_j\|$$

$$= \phi(a_i^2, a_{-i}) - \phi(a_i^1, a_{-i})$$

Time-invariant ⇨ $\sum_{p_j \neq p_i} \sum_{p_k \in N_j \setminus p_i} \frac{\|a_j - a_k\|}{2}$ is same

attention

Consensus point

⇨ Nash equilibrium of the game characterized by the

✗ Player's objective function $U_i(a) = -\sum_{p_j \in N_i} \|a_i - a_j\|$

⇨ $A^c \subset A^*$

set of Consensus point ↗ ↖ set of Nash equilibrium

⇩

Learning algorithm

Guarantees probabilistic convergence to a pure Nash equilibrium that maximize the potential function

Spatial adaptive play

Probability distribution

$p_i(t) \in \Delta(A_i)$ Set of probability distribution Over the set A_i

step1 Randomly choose one player P_i

(Another player do a same action)

step2 According to $p_i(t)$, player P_i randomly select an action .

$$p_i^{a_i}(t) = \frac{\exp\{\beta U_i(a_i, a_{-i}(t-1))\}}{\sum_{\bar{a}_i \in A_i} \exp\{\beta U_i(\bar{a}_i, a_{-i}(t-1))\}}$$

Spatial adaptive play

$$p_i^{a_i}(t) = \frac{\exp\{\ \beta U_i(a_i, a_{-i}(t-1))\}}{\sum_{\bar{a}_i \in A_i} \exp\{\ \beta U_i(\bar{a}_i, a_{-i}(t-1))\}}$$

$\beta = 0$ ➡ P_i select any action a_i with same probability

$\beta \to \infty$ ➡

Action a_i that satisfy

$$\{a_i \in A_i : U_i(a_i, a_{-i}(t-1)) = \max_{a_i \in A_i} U_i(a_i', a_{-i}(t-1))\}$$

has high probability

Spatial adaptive play

Repeated potential game with SAP

➡ The stationary distribution μ

$$\mu(a) = \frac{\exp\{\ \beta\phi(a)\}}{\sum_{\bar{a} \in A} \exp\{\ \beta\phi(\bar{a})\}} \quad , \mu \in \Delta(A)$$

(After large time t , $\mu(a)$ equals the probability that $a(t) = a$)

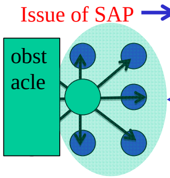
➡ $\beta \to \infty$ All the weight of μ is on the joint actions that maximize the potential function

If all players update their actions using SAP with large β , then the Players will reach a consensus with high probability

Restrictive Spatial Adaptive Play

Issue of SAP → Permit any action in its action set

→ From mobility limitation , it's impossible



← Restricted action set $R_i(a_i(t-1))$

Restrictive Spatial Adaptive Play

step1 Randomly choose one player P_i

(Another player do a same action)

step2 Player P_i selects **one trial action** $\hat{a}_i$ **from** $R_i(a_i(t-1))$. $z_i = \max_{a_i \in A_i} |R_i(a_i)|$

$$\Pr[\hat{a}_i = a_i] = 1/z_i \quad , \quad a_i \in R_i(a_i(t-1)) \setminus a_i(t-1)$$

$$\Pr[\hat{a}_i = a_i(t-1)] = 1 - (|R_i(a_i(t-1))| - 1)/z_i$$

Restrictive Spatial Adaptive Play

step3 Player P_i chooses its action at time t.

$$\Pr[a_i(t) = \hat{a}_i] = \frac{\exp\{\beta U_i(\hat{a}_i, a_{-i}(t-1))\}}{\exp\{\beta U_i(\hat{a}_i, a_{-i}(t-1))\} + \exp\{\beta U_i(a(t-1))\}}$$

→ Choose trial action

$$\Pr[a_i(t) = a_i(t-1)] = \frac{\exp\{\beta U_i(a(t-1))\}}{\exp\{\beta U_i(\hat{a}_i, a_{-i}(t-1))\} + \exp\{\beta U_i(a(t-1))\}}$$

→ Don't move

Theorem

Theorem

Consider a finite n-player potential game with potential function ϕ

If restricted action set satisfy

$$a_i^2 \in R_i(a_i^1) \leftrightarrow a_i^1 \in R_i(a_i^2) \quad \text{(Reversibility)}$$

and $a_i^k \in R_i(a_i^{k-1})$ (Feasibility)

then RSAP induce a Markov process over the state space A where the unique distribution $\mu \in \Delta(A)$ is

$$\mu(a) = \frac{\exp\{\ \beta\phi(a)\}}{\sum_{\bar{a} \in A} \exp\{\ \beta\phi(\bar{a})\}}$$

proof

(Proof) Show $\mu(a) P_{ab} = \mu(b) P_{ba}$

$$P_{ab} = \Pr[\ a(t) = b | a(t-1) = a\ ]$$

$$(a_{-i} = b_{-i}\ , a_i \neq b_i\ , b_i \in R_i(a_i))$$

Player P_i is chosen with probability $1/n$

Trial action $\hat{a}_i$ is chosen with probability $1/z_i$

$$\mu(a) P_{ab} =$$

$$\frac{\exp\{\beta\phi(a)\}}{\sum_{z \in A} \exp\{\beta\phi(z)\}} \frac{1}{n} \frac{1}{z_i} \frac{\exp\{\beta U_i(b)\}}{\exp\{\beta U_i(a)\} + \exp\{\beta U_i(b)\}}$$

$$\Pr[a_i(t) = \hat{a}_i] = \frac{\exp\{\beta U_i(\hat{a}_i, a_{-i}(t-1))\}}{\exp\{\beta U_i(\hat{a}_i, a_{-i}(t-1))\} + \exp\{\beta U_i(a(t-1))\}}$$

proof

$$\lambda = \frac{1}{\sum_{z\in A}\exp\{\beta\phi(z)\}}\frac{(1/n)(1/z_i)}{\exp\{\beta U_i(a)\}+\exp\{\beta U_i(b)\}}$$

$\rightarrow$ $\mu(a)P_{ab} = \lambda\exp(\beta\phi(a)+\beta U_i(b))$

Potential game $\rightarrow$ $U_i(b) - U_i(a) = \phi(b) - \phi(a)$

$$\mu(a)P_{ab} = \lambda\exp(\beta\phi(b)+\beta U_i(a))$$

$\Rightarrow$ $\mu(a)P_{ab} = \mu(b)P_{ba}$

Conclusion of RSAP

All players adhere to RSAP in consensus problem

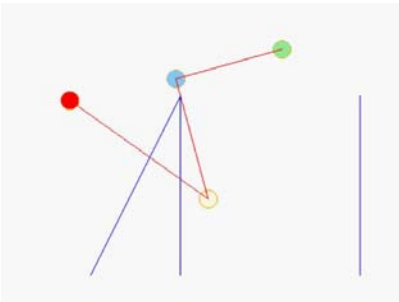
(interaction graph is Time-invariant and undirected)

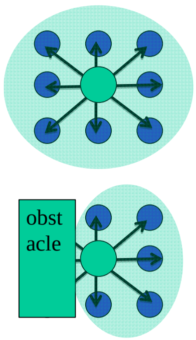
1.Player's collective behavior will maximize the potential Function with high probability.

2. interaction graph is connected and consensus is possible

$\Rightarrow$ player's actions constitute a consensus with high probability even in an environment filled with non-convex obstruction

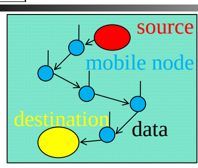
Consensus with obstruction





RSAP $\beta = t/70$ Restricted action set

Sensor Deployment Problem



object

Positioning the intermediate nodes to minimize the total transmission power

Power of transmitting information from P_i to P_j

$$e(a_i,a_j) = \alpha_1 + \alpha_2\|a_i - a_j\|^2$$

object $\Rightarrow$ minimize $\sum_{P_i\in P}\sum_{P_j\in N_i} e(a_i^*,a_j^*)$

equivalent

maximizing $\phi(a) = -\sum_{p_i\in p}\sum_{p_j\in N_i}\|a_i - a_j\|^2$

Equivalent consensus problem

Sensor Deployment Problem

objective function

$$U_i(a) = -2\sum_{p_j\in N_i}\|a_i - a_j\|^2$$

Do SAP or RSAP

$\Rightarrow$ The stationary distribution μ

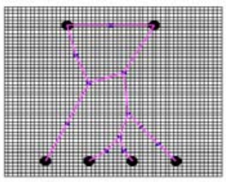
$$\mu(a) = \frac{\exp\{\beta\phi(a)\}}{\sum_{\bar{a}\in A}\exp\{\beta\phi(\bar{a})\}}, \quad \mu\in\Delta(A)$$

$\beta\rightarrow\infty$ stationary distribution is placed on action profiles which maximize the potential function

$\Rightarrow$ Action profiles represent minimum power allocation

Sensor Deployment Problem

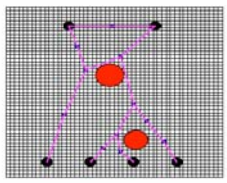
RSAP



$e(a_i,a_j) = \|a_i - a_j\|^2$

$\beta = 1 + t/300$

RSAP



There is no obstruct in the path $e(a_i,a_j) = \|a_i - a_j\|^2$

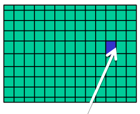
There is obstruct in the path $e(a_i,a_j) = 1.3\|a_i - a_j\|^2$

$\beta = 1 + t/300$

Interaction topology is fixed

Dynamic sensor coverage

Mission space S



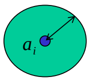
Sector s

object
To allocate sensors across a given space

Density function

$V(s) \geq 0 \quad , \sum_{s \in S} V(s) = 1$

Action set $A_i = S$



$\|s - a_i\| < r_i \Leftrightarrow p_i(s, a_i) > 0$

Probability of detecting an event in sector s

$$P(s,a) = 1 - \prod_{p_i \in p} [1 - p_i(s, a_i)]$$

$$\phi(a) = \sum_{s \in S} V(s) P(s, a)$$

WLU $\quad U_i(a) = \phi(a_i, a_{-i}) - \phi(a_i^0, a_{-i})$

Dynamic sensor coverage

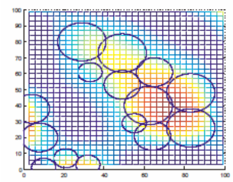
$$U_i(a) = \phi(a_i, a_{-i}) - \phi(a_i^0, a_{-i})$$

Sensor can evaluate utility function using only local information

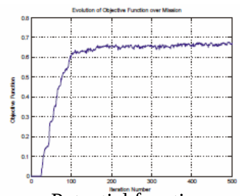
Simulation result

RSAP $\quad \beta = 0.6 \qquad$ 6: 6 radius sensor $\quad$ 6: 18 radius sensor

↳ range restricted action $\qquad$ 6: 12 radius sensor



Sensor configuration



Potential function

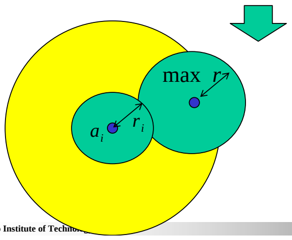
Dynamic sensor coverage

$$P(s,a) = 1 - \prod_{p_i \in p} [1 - p_i(s, a_i)]$$

$$\phi(a) = \sum_{s \in S} V(s) P(s, a)$$

$$U_i(a) = \phi(a_i, a_{-i}) - \phi(a_i^0, a_{-i})$$

Sensor can evaluate utility function using only local information



Communication range is needed at least $r_i + \max r$

to calculate utility function