ポテンシャルゲーム理論的姿勢協調 -同期・平衡の達成

木下大輔

富山県立大学　電子・情報工学科

May 21, 2021

はじめに

はじめに

安価で消費電力の低いセンサや無線通信機器などの要素技術の発展によってセンサネットワーク、ロボティックネットワーク、交通ネットワーク、電力ネットワークに代表されるネットワークのスマート化が望まれている。

目的

- ポテンシャルゲーム理論に基づいて、姿勢同期と姿勢平衡なる姿勢強調を考察する。
- 姿勢強調問題をゲームとして表現する。
- 適当な仮定の下でこれらの姿勢強調ゲームがポテンシャルゲームを構成することを証明する。
- 目的を達成する均衡解への到達手法年て Restrictive Special Adaptive Play(RSAP) を加速する学習アルゴリズムを提案する。

姿勢強調制御

協調制御

個々の構成要素が限られた情報交換のもとで協調的に動作し、所望の目的を達成することが求められる。このような目的を達成する制御は協調制御と呼ばれ様々な協調制御問題が提案されてきた。

姿勢協調問題

姿勢強調問題は 2 次元及び 3 次元空間上の剛体の姿勢を、分散制御則によってある秩序だった配置に収束させる問題である。

TOYAMA Prefectural University

問題設定 1

2 次元空間に存在する N 台のエージェント $\nu :=$ 1,..,N から構成されるロボティックネットワークを考える。本論を通して以下の仮定をおく。

仮定 1

拘束行動集合は、以下の条件を満たす

1. **可逆性** 任意の $i \in \nu$ 及び行動 $a_i^1, a_i^2 \in A_i$ に対して $a_i^2 \in R_i(a_i^1) \longleftrightarrow a_i^1 \in R_i(a_i^2)$ が成立する。
2. **可解性** 任意の $i \in \nu$、行動 $a_i^0, a_i^m \in A_i$ に対して,$\exists(a_i^0, .., a_i^m)$s.t
$a_i^l \in R_i(a_i^{l-1}), l \in 1, .., m$ が成立する。

エージェント $i \in \nu$ の姿勢 $\theta_i \in [0, 2\pi)$、エージェント i と j の相対姿勢 $\theta_{ij} = \theta_i - \theta_j$、各エージェントの姿勢の動き $\theta_i(k+1) = a_i(k), a_i(k) := \theta_i(k) + u_i(k)$、エージェント i に対する速度入力 $u_i(k)$、設計対象 $a_i(k) = \theta_i(k) + u_i(k) \in A \subset [0, 2\pi)$、時刻 k における行動 $a_i(k) \in A \subset [0, 2\pi)$, 行動集合 A_i, 全てのエージェントの行動集合を集めた $A := A_1 \times .. \times A_n$, 全てのエージェントの行動まとめ $a = (a_1, ..a_N) \in A$, 行動の偏差 $a_{ij} = a_i - aj$, エージェント i が行動 a_i を選択している際の実行可能な行動集合を拘束行動集合 $R_i(a_i) \subset A_i$

制御アルゴリズム

今回の目的は、以下の二つの姿勢配置を達成する制御アルゴリズムを提案することである。

$$\theta_i = \theta_j, \forall i, j \in \nu \tag{1}$$

$$\sum_{i=1}^{N} \cos\theta_i = 0, \sum_{i=1}^{N} \sin\theta_i = 0 \tag{2}$$

エージェント群 ν が (1) を満たすとき、姿勢同期を達成するという、また (2) を満たすとき姿勢平衡を達成するという。すなわち姿勢同期とは全てのエージェントの姿勢が一致することを指し、姿勢平衡とはエージェントの姿勢が釣り合うことを意味する。

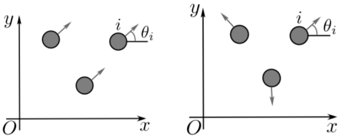

図 1: 姿勢同期と姿勢平衡

姿勢強調ゲーム

姿勢強調ゲーム

姿勢強調問題をゲームとして捉える。戦略型ゲームはエージェント ν、行動集合群 A 及び利得関数 $\{U_i\}_{i\in\nu}, U_i : A \to R$ の組として定義される。個々のエージェントは自身の利得関数を U_i を最大化するために行動 $a_i(k)$ を選択する。よって行動制約のもとで大域的に姿勢協調を達成する利得関数と行動の決定アルゴリズムを設計する。

ポテンシャルゲーム

定義

エージェント集合 ν、行動集合群 A および利得関数 $\{U_i\}_{i\in\nu}, U_i : A \to R$ からなるゲーム Γ を考える。このとき、ある関数 $\o$ が存在していて任意の $i \in \nu$,$a_i', a_i'' \in A_i$ に対して

$$U_i(a_i'', a_{-i}) - U_i(a_i', a_{-i}) = \o(a_i'', a_{-i}) - \o(a_i', a_{-i}) \tag{3}$$

を満たすとき Γ はポテンシャルゲームを構成する。

ポテンシャル関数

ラプラシアン姿勢ポテンシャルを導入する。

定義

$$W(\theta) = \frac{1}{2N}\sum_{i=1}^{N}(d_i - \sum_{j\in}\cos\theta_{ij})$$

d_i は近傍 N_i の要素数である。以降は簡単化のため $D := \sum_{j\in}\cos\theta_{ij}$ とする。ポテンシャル $W(\theta)$ は以下の性質を満足する。

- $W(\theta)$ が最小の時エージェント群は同期する
- G が環状グラフ $(N_i = \{N, 2\}, N_N = \{1, N-1\}, N_i = \{i-1, i+1\}, i \neq \{1, N\})$ であれば $W(\theta)$ が最大の時、エージェント群は姿勢平衡を達成する。

同期、平衡のポテンシャル関数を以下のように定義する

$$\phi_s(a) = -W(a) \tag{4}$$

$$\phi_b(a) = W(a) \tag{5}$$

利得関数

姿勢強調問題に対する利得関数を設計し、結果として与えられるゲームがポテンシャルゲームを構成することを証明する。

利得関数

エージェント群 ν、行動集合 $\{A_i\}_{i=1}^{N}$ に加えて利得関数を

$$U_i(a) = \frac{1}{N}\sum_{j\in N}\cos a_{ij} \tag{6}$$

とする姿勢同期ゲーム Γ_s を考える。このとき仮定 2 のもとで、ゲーム Γ_s はポテンシャル関数 (4) に対するポテンシャルゲームを構成する。また利得関数を

$$U_i(a) = -\frac{1}{N}\sum_{j\in N}\cos a_{ij} \tag{7}$$

とする姿勢平衡ゲーム Γ_b を考える。このとき仮定 2 のもとで、ゲーム Γ_b はポテンシャル関数 (5) に対するポテンシャルゲームを構成する

証明

証明

仮定 2 のもとでポテンシャル関数は

$$\phi(\theta) = -\frac{1}{2N}\left(D - \sum_{j\in N_i} 2\cos a_{ij} - \sum_{j\neq i}\sum_{l\in\frac{N_i}{\{i\}}} \cos a_{jl}\right.$$

と書き換えることが出来る。行動を a_i' から a_i'' に変化させたときのポテンシャル関数の変化は、

$$\begin{aligned}
&\phi(a_i'', a_{-i}) - \phi(a_i', a_{-i}) \\
=\ &\frac{1}{2N}\sum_{j\in N_i} 2\cos(a_i'' - a_j) - \frac{1}{2N}\sum_{j\in N_i} 2\cos(a_i' - a_j) \\
=\ &U_i(a_i'', a_{-i}) - U_i(a_i', a_{-i})
\end{aligned}$$

群ポテンシャルゲーム

群ポテンシャルゲームでは、エージェント集合 ν、行動集合群 A に加えて ν の部分集合族 $\nu_i, i=1,..,m$ の存在を仮定し、各集合 ν_i に利得関数を対応させる。すなわち利得関数を $U_{\nu_i}: A \to R$ とする。

定義

エージェント群 ν, 行動集合 A_i, 集合族 $\nu_i, i=1,..,m$ および $\{U_{\nu_i}\}_{i=1}^m$ から構成されるゲーム Γ を考える。任意の $\nu_i, a'_{\nu_i}, a''_{\nu_i} \in \prod_{i\in\nu_i} A_i, a-\nu_i = (a_j)_{j\notin\nu_i} \in \prod_{i\notin\nu_i} A_i$ に対して

$$U_{\nu_i}(a''_{\nu_i}, a_{-\nu_i}) - U_i(a'_{\nu_i}, a_{-\nu_i}) = \phi(a''_{\nu_i}, a_{-\nu_i}) - \phi(a'_{\nu_i}, a_{-\nu_i}) \tag{8}$$

を満たすときを満たす関数 $\phi: A \to \mathbb{R}$ が存在するときゲーム Γ は群ポテンシャルゲームを構成するという。

郡利得関数

集合 ν_i を N_i とする。このとき次の定理が成り立つ。

群利得関数

エージェント群 ν、行動集合 $\{A_i\}_{i=1}^{N}$、集合族 $\{N_i\}_{i-1}^{N}$ に加えて群利得関数を

$$U_{N_i}(a) = \frac{1}{N}\sum_{j \in N_i} \cos a_{ij} + \frac{1}{N}\sum_{j \in N_i}\sum_{r \in \frac{N_i}{N}} \cos a_{jr} + \frac{1}{2N}\sum_{j \in N_i}\sum_{l \in N_i \cap N} \cos a_{jr}$$

とする姿勢同期ゲーム $\Gamma_g s$ を考える。このとき仮定 2 のもとで、ゲーム $\Gamma_g s$ はポテンシャル関数 (4) に対する群ポテンシャルゲームを構成する。

郡利得関数

群利得関数

$$U_{N_i}(a) = -\frac{1}{N}\sum_{j\in N_i}\cos a_{ij} - \frac{1}{N}\sum_{j\in N_i}\sum_{r\in\frac{N_i}{N}}\cos a_{jr}$$
$$\frac{1}{2N}\sum_{j\in N_i}\sum_{l\in N_i\cap N}\cos a_{jr}$$

とする姿勢平衡ゲーム $\Gamma_g b$ を考える。このとき仮定２のもとで、ゲーム $\Gamma_g b$ はポテンシャル関数 (5) に対する群ポテンシャルゲームを構成する。
以上より姿勢強調問題は群ポテンシャルゲームを構成する。すなわち群 N_i が U_i を増加する行動を取れば、ポテンシャル関数も増加する。

提案手法の概要

Algorithm 1 :Restrictive Spacial Adaptive Play with Neighbor-based Decisions

初期状態: $k := 0$

1: ランダムに一つのエージェント $i \in \mathcal{V}$ を選択する．エージェント i は，選ばれた事実を近傍のエージェント $\mathcal{N}_i$ に伝える．

2: 近傍のエージェント $j \in \mathcal{N}_i$ は，その近傍 $\mathcal{N}_j$ の情報 $\{a_l(k-1)\}_{l \in \tilde{\mathcal{N}}_j}$ を集め，エージェント i に伝える．

3: エージェント i は，一つの試行行動集合 $\hat{a}_{\tilde{\mathcal{N}}_i}$ を確率 (9)，(10) に従って選択する．

4: エージェント i は，近傍からの情報をもとに群目的関数を計算する．

5: エージェント i は，確率 (11)，(12) に従って行動群 $a_{\tilde{\mathcal{N}}_i}(k)$ を選択する．

6: エージェント i は近傍に $a_{\tilde{\mathcal{N}}_i}(k)$ を伝える．

7: $j \in \tilde{\mathcal{N}}_i$ は行動 $a_{\tilde{\mathcal{N}}_i}(k)$ を実行する．

8: $k \leftarrow k+1$ として Step 1 へ．

図 2: 概要

TOYAMA Prefectural University

学習アルゴリズム

$$\Pr[\hat{a}_{\tilde{\mathcal{N}}_i} = a_{\tilde{\mathcal{N}}_i}] = 1/\tilde{z}_i \quad \text{if } a_{\tilde{\mathcal{N}}_i} \in \tilde{\mathcal{R}}_i(a_{\tilde{\mathcal{N}}_i}(k-1)) \tag{9}$$

$$\Pr[\hat{a}_{\tilde{\mathcal{N}}_i} = a_i(k-1)] = 1 - |\tilde{\mathcal{R}}_i(a_{\tilde{\mathcal{N}}_i}(k-1))|/\tilde{z}_i \tag{10}$$

自然数 z_i' は障害物がない場合の選択可能な行動の数 z_i を用いて $z_i' = \prod_{j \in N_i} z_i$ と定義される。すなわちハードウェア制約に起因する制限がなく、自由に姿勢を変更できる。状況下では $z_i = |A_i|$ である。最後に、試作行動 $\hat{a}_{N_i}$ を実際に実行するかを以下の確率に従い決定する。

$$\Pr[a_{\tilde{\mathcal{N}}_i}(k) = \hat{a}_{\tilde{\mathcal{N}}_i}] = \frac{\exp\{\beta U_i(\hat{a}_{\tilde{\mathcal{N}}_i}, a_{-\tilde{\mathcal{N}}_i}(k-1))\}}{\exp\{\beta U_i(\hat{a}_{\tilde{\mathcal{N}}_i}, a_{-\tilde{\mathcal{N}}_i}(k-1))\} + \exp\{\beta U_i(a(k-1))\}} \tag{11}$$

$$\Pr[a_{\tilde{\mathcal{N}}_i}(k) = a_{\tilde{\mathcal{N}}_i}(k-1)] = \frac{\exp\{\beta U_i(a(k-1))\}}{\exp\{\beta U_i(\hat{a}_{\tilde{\mathcal{N}}_i}, a_{-\tilde{\mathcal{N}}_i}(k-1))\} + \exp\{\beta U_i(a(k-1))\}} \tag{12}$$

収束性解析

集合 A 内の行動群に順番をつけて $a^{(1),..,a^{(|A|)}}$ と表現する。つぎに、時刻 k において行動群 $a^{(i)}$ を実行する確率を μ_i^k と表現し、それらの行ベクトルを $\mu(k)$ と表記する。この遷移はマルコフ連鎖として表現される。

$$\mu(k+1) = \mu(k)P \tag{13}$$

ここで行列 P は遷移行列を表し、その（i,j）成分 P_{ij} は $Pr[a(k+1) = a^{(j)}|a(k) = a^{(j)}]$

仮定 1,2 を満足する姿勢強調群ポテンシャルゲームを考える。この時 Algorithm1 を用いるならば $\mu(k)$ の定常分布 μ は次式で与えられる。

$$\mu_i = \frac{\exp\{\beta\phi(a^{(i)})\}}{\sum_{\bar{a}\in A}\exp\{\beta\phi(\bar{a})\}} \tag{14}$$

RSAP のシミレーション

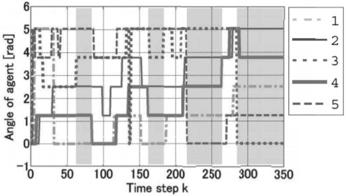

図 3: 各エージェントの姿勢

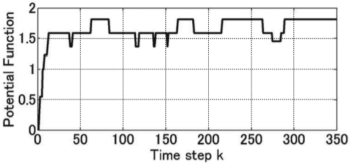

図 4: ポテンシャル関数の応答

Algorithm1 のシミレーション

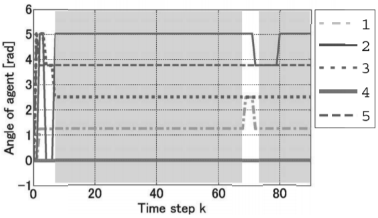

図 5: 各エージェントの姿勢

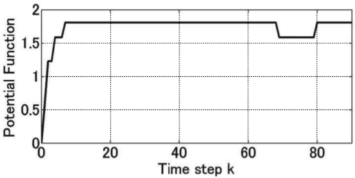

図 6: ポテンシャル関数の応答

姿勢同期

姿勢同期

ビジュアルセンサネットワークにおける姿勢制御を想定して、障害物が存在する状況における姿勢同期の達成を検証する。４台のカメラが存在する増強を踏まえて、行動集合を $A_i = \{l\pi/4\}_{l=0}^{7}, i_1, 2, 3, 4$ と設定する

図 7: カメラ環境

比較

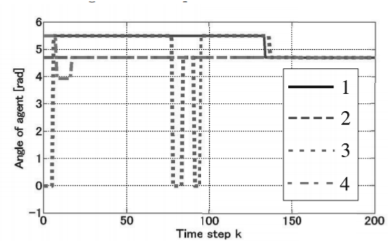

図 8: RSAP

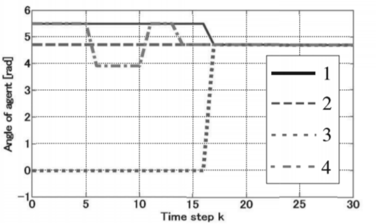

図 9: Algorithm 1

まとめ

まとめ

- 適切なポテンシャル関数と目的関数を用いることで、姿勢強調問題をポテンシャルゲームとして表現できることを示した。
- 姿勢強調問題が群ポテンシャルゲームを構成することを示し、複数のエージェントを同時に動かすことで、定常分布に収束する速さを加速するアルゴリズムを提案した。
- 有効性および妥当性をシミュレーションにより検証した。