[計測自動制御学会論文集
Vol.44, No.2, 191/198 (2008)]

知的車椅子ロボットにおける環境パラメータを介した協調行動の獲得

濱上知樹*・平田廣則**

Cooperative Behavior Acquisition by Using Local Environment Parameters for Intelligent Wheelchairs

Tomoki HAMAGAMI* and Hironori HIRATA**

This paper proposes a technique for realizing cooperation behavior emerging among intelligent wheel chairs (IWC). In order to acquire the behavior in various environments, a system with a broadcasting framework and a decision tree is implemented, which uses a frustration level indicating a dynamical internal state of an IWC . The system brings diversity among each IWC's, and resolves a deadlock of them. Moreover, in order to design an adaptive behavior, the decision tree is generated with the genetic programming. As results of experiment with simulation and real prototype IWC, we confirm the mutual concession behavior emerges in narrow aisle and cross road environment .

Key Words: intelligent wheelchair, cooperative behavior, stigmergy, genetic programming

1. はじめに

現在，日本における65歳以上の高齢者人口は2,560万人に達し，2050年の高齢化率は35.7%に達する見込みである[1]．超高齢社会を目前として，高齢者の自立生活を支援するさまざまな福祉機器開発が進められている．その中でも，電動車椅子は，高齢者の自立した生活，社会参加に欠かせない補助具である．しかし，普及に伴う事故の増加や公共の場での使用制限などさまざまな課題が残されている．今後，電動車椅子が広く社会で使われるようになるためには，より高い安全性と機能性を実現する知的な働きが必要である．以上を背景として，本研究では，利用者や環境中のさまざまな機器・ネットワークと連携しながら，知的な振る舞いにより快適で安全な空間を作り出す，ネットワークロボットとしての知的車椅子の研究開発を進めている[2)~4)]．

知的車椅子ロボットに求められる機能として，障害物回避・目的場所への移動等の自律性[5)]や，操作性向上のための高度なインタフェース機能[6)]に加え，単独のロボットでは対処できない複雑な問題を解決する協調性が重要である．このような協調性を実現した例として，館内での道案内機能をもったロボットを複数台用いて，大規模な人員誘導や整理を実現するシステムが提案されている[7)]．しかし，ロボットが動作する場所や，ロボットの数が時間とともに変動する環境においては，限られたセンサ情報とコミュニケーション手段を用いて有効な協調プランニングを確立することは容易ではない．

また，知的車椅子ロボットが動作する環境は，車道や工場のような整理された環境とは異なり，施設や屋内等，狭く多様な構造を有している．そのような環境において，事前にロボット間の明確な役割やルールを決めておくことは困難である．このような環境においては，動作する環境に応じて柔軟な振る舞いが可能な制御機構が望ましい．たとえば，中央集中制御機構を持たないマルチエージェント自律分散型制御が有望である[8),9)]．また，知的車椅子をはじめとする家庭用ロボットが混在する環境においては，少ない計算資源と簡単な情報処理だけで知的な行動を実現できる共通のアーキテクチャが必要になる[10),11)]．すなわち，近傍ロボットのIDを用いた直接的な交渉や，双方向の要求・指令などの意味情報を陽に含まない簡易なコミュニケーションから，協調・適応的動作が自律分散的に獲得されることが望ましい．

そこで本研究では，知的車椅子同士の協調行動を，一部の社会性昆虫等が有する環境情報を媒介とする簡便な非明示的コミュニケーションの手段を用いて創発させることを試みる．そして，制約のある環境観測能力と通信条件下においても，環境に応じた協調行動を自律的に獲得する手法を提案する．

以下，2章では，本研究で試作している知的車椅子ロボットの仕様について述べる．3章では，車椅子ロボット同士の協調行動が必要とされるタスクについて述べる．4章では，非明示的コミュニケーションを実現するパラメータとこれを用

* 横浜国立大学大学院工学研究院　横浜市保土ヶ谷区常盤台79–5
** 千葉大学大学院工学研究科　千葉市稲毛区弥生町1–33
* Graduate School of Engineering, Yokohama National University, 79–5 Tokiwadai, Hodogaya-ku, Yokohama
** Graduate School of Engineering, Chiba University, 1–33 Yayoi-cho, Inage-ku, Chiba
(Received June 12, 2007)
(Revised January 7, 2008)

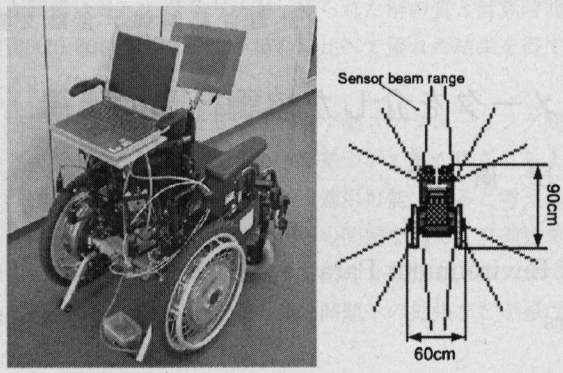


Fig. 1 The appearance of intelligent wheelchair robot, and the beam distance of range sensors on simulation screen

いた行動決定木の進化的獲得法を提案する．5章ではシミュレータと実機による実験と結果について述べる．6章では先行研究事例との関係について触れる．7章はむすびである．

2. 知的車椅子ロボット

本研究で開発を進めている知的車椅子ロボット（以下ロボット）の特徴は，自律・協調・協同行動を獲得する点にある[3),4)]．

ハードウェア構成

Fig. 1 に，知的車椅子ロボットの外観と，上からみたセンサビームの有効範囲を示す．本研究では，市販の電動車椅子をベースに，周囲に12個の測距センサ（PSDセンサ）を配し，周囲の障害物までの距離を測定する．車椅子の背面には制御用PCと，AD/DAコンバータ等のインターフェースボードが搭載されている．また，Bluetoothによる近傍のロボット同士のアドホックな無線ネットワークを構成することができる．

ソフトウェア構成

ロボットの振る舞いは，各ロボットに搭載された計算機で動作するエージェントにより制御されている．各エージェントは，前述のセンサからロボット周辺の局所的な情報を獲得し，その情報に基づき自律的な行動選択を行なう機構を有する．一方で，エージェントは以下のような制約条件のもとで適切な行動を獲得しなくてはならない[12)]．

- エージェントは動的に変化する環境情報を観測するセンサを持たない．よって，自分と他の車椅子の位置関係を正確に知ることはできない．
- 測距センサは，周囲の障害物との距離を離散的に観測するが，障害物の種類を特定することはできない．すなわち，壁やロッカー等の固定された障害物と，他の車椅子の区別はない．
- 車椅子は，無線ネットワークを介して近傍の車椅子同士で簡単なメッセージの送受信を行なう．しかし，その通信範囲，通信容量には制限がある．

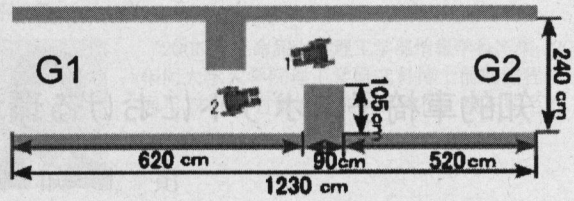


Fig. 2 An example of environment with narrow gap causing difficulties for robots going through each other

3. 協調行動が必要なタスク

実環境においては，知的車椅子ロボット同士が動的に協調し合わなくてはならない場面が多数存在する．たとえば，**Fig. 2** に示すような障害物のある通路における進路の譲り合いを考える．

Fig. 2 に示す2台のロボット (1,2) を制御する各エージェントは，それぞれの目標方向G1,G2にむけて，各ステップごとに状態を観測し行動を決定する．互いのロボットの位置関係が把握できていない場合，正面にいるロボットを障害物とみなしてしまう．互いに衝突を避け合うことで，双方とも目的地にたどり着くことができない．本論文では，このようにそれぞれのロボットが所望の行動をとろうとした際に，お互いの存在によりどちらもそれを妨げられてしまい，状態が変化しなくなることをデッドロックと呼ぶ．

従来，デッドロックを回避するためには，環境の正確な認識と，高度なプランニングや協調学習が必要であった[13)]．しかし，限られたセンサしか持たないロボットの場合，完全な環境の情報を得ることは難しい．また，たとえ正確な情報が得られる場合でも，考慮すべき状態数が多くなると，プランニングに時間がかかったり学習すべき状態数が多くなり収束が困難になるなどの問題が生じる．

一方，マルチエージェントにおける競合状況を，逐次的な報酬に基づくオンラインの強化学習で解決をはかる方法も提案されている[14)]．しかし，車椅子が使用される実環境を考えた場合，環境からの報酬を逐次的に得ることは現実的ではない．また，各エージェントが独立した学習を行ない，互いに役割分担をすることで組織的な協調行動を実現することもできるが，その場合，集団の規模やエージェントの種類に依存した結果となり，任意の集団における協調行動が保証されないなどの問題がある．

4. フラストレーションレベルを用いた協調行動の獲得

4.1 局所的な環境を利用した協調行動の創発

一方，単純なセンサしか持たない一部の社会性昆虫等では，局所的な環境を媒介にした動的な情報の生成・相互作用を利用して，さまざまな状況に対応可能な協調行動を生み出している．このような環境を媒介とする局所的相互作用・コミュニケーションのしくみはstigmergyと呼ばれ，群知能の創発メカニズムとして研究が進められている[15),16)]．知的車椅子ロ

ボットのような人工システムにおいても，人工的なstigmergyを発生させ，これを利用する行動獲得機構を実現できれば，限られたセンサと単純な行動決定機構しか持たないロボット同士からでも，高度な協調行動の創発が期待できる．

4.2 フラストレーションレベルの導入

本研究では，そのような局所的な環境情報として，各ロボットが有する動的なパラメータを定義し，これを互いにブロードキャストしあうことで局所的な情報空間を集団内で共有する方法を提案する．

本論文では，動的なパラメータとして，各ロボットのステップごとの行動の成功・失敗により変化するスカラー値を用いる．具体的には，エージェントが選択した行動通りにロボットが振る舞えた場合は減衰し，失敗した場合に増加する内部状態値を用いる．以下，この内部状態値をフラストレーションレベル (Frustration Level: FL) と呼ぶ．(1) 式に以下で用いるFLを示す．

$$\phi^{(i)}(t+1)= \begin{cases} \text{if the action is success} \\ \phi^{(i)}(t)+\alpha\left(\phi^{(i)}(t)-\phi^{(i)}(t-1)\right)-(1-\alpha)u_i \\ \text{else} \\ \phi^{(i)}(t)+\alpha\left(\phi^{(i)}(t)-\phi^{(i)}(t-1)\right)+(1-\alpha)u_i \end{cases} \quad (1)$$

$\phi^{(i)}(t)$ は $i \in I$ 番目のロボットのエージェント A_i が，時刻 t で持つFLである．I はアドホックネットワークにより通信可能な近傍エージェントの数を表わす．u_i は A_i のFL増減幅であり，後述の範囲内でエージェントごとにランダムな値が設定される．u_i が異なっていることで，$\phi^{(i)}(t)$ の増加・減衰のパターンに変化が生じる．すなわち，同じ回数だけ行動の成功・失敗があったとしても，ロボットごとに異なるFLの変化が生じることで，ロボット集団中にFLの相違が発生する．後述するように，集団内でFLの相違が発生することにより異なる行動が発生し，デッドロック解消のきっかけをつくる．

A_i は近傍の他 $A_j(j \in I, j \neq i)$ 同士との間でアドホックネットワークを動的に構築し，$\phi^{(i)}(t)$ をブロードキャストする．同時に，$\phi^{(j)}(t)(j \in I, j \neq i)$ を受信する．

4.3 FLに基づく行動獲得

エージェントはFLを用いた以下の手順でロボットを操作する．**Fig. 3** に手順を示す．

(1) A_i はAction selector(AS)の計算に基づき行動を決定する．ここでASとは，センサからの情報をもとに静的な環境において合理的な行動をとるようにあらかじめ学習・設計されたサブシステムである．具体的には，文献2),12)で提案した手法に基づき，他の移動ロボットが存在しない環境において，目的地移動までの行動プランニングをシミュレータと実機上のティーチングとQ-learningを併用した学習により，あらかじめ構築しておく．ASからの出力される行動とは，左右の車輪の回転速度である．

(2) 行動を受け取ったBehavior controller(BC)は，

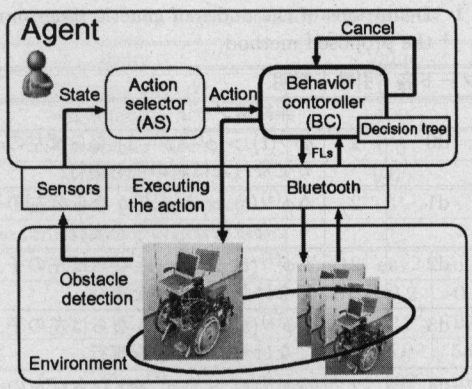


Fig. 3 The process of action execution of robot with an action selector and behavior controller

$\phi^{(i)}(t)$ を計算する．また，近傍のエージェントから $\phi^{(j)}(t)$ を受信し，$\phi_\Sigma^{(i)}(t)$，$\phi_\mu^{(i)}(t)$，およびこれらの一次差分を求める．

$$\phi_\Sigma^{(i)}(t) = \sum_{j \neq i}^{I} \phi^{(j)}(t) \quad (2)$$

$$\phi_\mu^{(i)}(t) = \phi_\Sigma^{(i)}(t)/(I-1) \quad (3)$$

$$\Delta\phi^{(i)}(t) = \phi^{(i)}(t) - \phi^{(i)}(t-1) \quad (4)$$

$$\Delta\phi_\Sigma^{(i)}(t) = \phi_\Sigma^{(i)}(t) - \phi_\Sigma^{(i)}(t-1) \quad (5)$$

$$\Delta\phi_\mu^{(i)}(t) = \phi_\mu^{(i)}(t) - \phi_\mu^{(i)}(t-1) \quad (6)$$

(3) BCは，(1)式で示した自身の状態であるFLと，(1)〜(6)式で示した環境パラメータの値をもとに，ASが選択した行動（両輪の回転速度）をそのまま実行するか，停止するか，反対の行動をとるかを決定木を用いて選択する．反対の行動が選択されると，ロボットはこれまで進んできた経路から外れて，後退するような動作となる．このように，本手法においては，近傍ロボット同士の位置情報を使わずに行動を決定している点が特徴である．

一方，さまざまな環境において，協調行動を誘発する適切な行動決定をする決定木を人手で設定することは容易ではない．そこで，決定木の構造を最適化するために，メタヒューリスティスクの一手法である遺伝的プログラミング (GP) [17] を用いる．

4.4 進化的プログラミングによる行動の獲得

以下に，GPを用いたBCにおける決定木獲得のアルゴリズムを示す．また，**Fig. 6** に示すように，ここで進化をする決定木はすべてのエージェントが共通して用いる．

(1) エージェント集団の生成

シミュレータ上に協調行動を創発させるロボットを I 台生成する．一定の距離以内にいるロボット同士は互いに通信することができる．各エージェントは，一定の範囲内で u_i をランダムに決定する．

(2) 初期遺伝子集団の生成

Table 1 に示すノードをランダムに選んだ初期遺伝子を

Table 1 Definitions of the nodes of genetic programming in the proposed method

ID	ノード名	引数	説明
			非終端ノード
1	d0	2	$\phi^{(i)}(t) > \phi^{(i)}(t-1)$ ならば左の子，そうでなければ右の子を実行.
2	d1	2	$\Delta\phi^{(i)}(t) > \Delta\phi_{\Sigma}^{(i)}(t)$ ならば左の子，そうでなければ右の子を実行.
3	d2	2	$\Delta\phi^{(i)}(t) > \phi_{\mu}^{(i)}(t)$ ならば左の子，そうでなければ右の子を実行.
4	d3	2	$\Delta\phi^{(i)}(t) > \phi_{\Sigma}^{(i)}(t)$ ならば左の子，そうでなければ右の子を実行.
5	d4	2	$\Delta\phi^{(i)}(t) > \Delta\phi_{\mu}^{(i)}(t)$ ならば左の子，そうでなければ右の子を実行.
6	p1	2	$\phi^{(i)}(t) > \Delta\phi_{\Sigma}^{(i)}(t)$ ならば左の子，そうでなければ右の子を実行.
7	p2	2	$\phi^{(i)}(t) > \phi_{\mu}^{(i)}(t)$ ならば左の子，そうでなければ右の子を実行.
8	p3	2	$\phi^{(i)}(t) > \phi_{\Sigma}^{(i)}(t)$ ならば左の子，そうでなければ右の子を実行.
9	p4	2	$\phi^{(i)}(t) > \Delta\phi_{\mu}^{(i)}(t)$ ならば左の子，そうでなければ右の子を実行.
			終端ノード
10	F	0	AS の指令に従う
11	B	0	AS の指令と反対の動作をする
12	S	0	なにもしない
13	R	0	F,B,S を等確率で発生させる

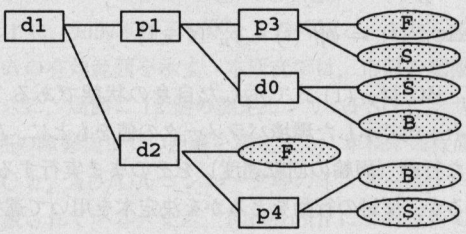


Fig. 4 An example of gene representing by S equations, and a decision tree as the phenotype of the gene

N 個体生成する．このとき，木の深さは一定にそろえられる．このようにしてつくられた個体群を，個体集団と呼ぶ．深さ 3 の初期遺伝子の例と，対応する決定木を **Fig. 4** に示す．

（3） 適応度の計算

個体集団から 1 つの個体を取り出し，その表現形である行動決定木を作成する．そして，すべてのエージェントがこの 1 つの決定木を用いてタスクを実行する．タスク終了後，後述の適応度を評価する．この操作を，すべての個体について行なう．

（4） 遺伝的操作

各個体の適応度を用いて，集団内からルーレット選択により 2 個体を選択し，交叉率と突然変異率に基づいた交叉，突然変異を行なう．交叉・突然変異の方法を **Fig. 5** に示す．交叉はランダムに選ばれた非終端ノードを根とする部分木を入れ替える操作である．突然変異は，任意のノードを 1 つ選び，非終端ノード，終端ノードごとにランダムなノードに置き換える操作である．また，交叉により深くなりすぎた決定木を除去するために，一定の深さ以上の構造になった遺伝子に対しては，強制的に終端ノードと非終端ノードを入れ替える処理を行なう．

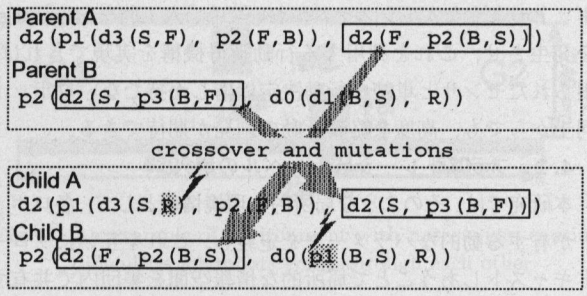


Fig. 5 Crossover and mutation operations

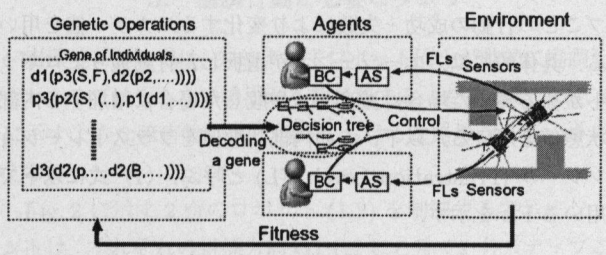


Fig. 6 The details of the genetic algorithm for acquiring cooperative behavior

（5） 世代交代

世代を 1 つ進め，(3) に戻る．規定世代数の後，エリート個体を最終的に得られる行動決定木とする．

4.4.1 適応度

ロボット集団のタスクは，それぞれの初期座標 $(x_S^{(i)}, y_S^{(i)})$ から行動を開始 $(t=0)$ した後，いずれかのロボットが (1) 壁または互いに衝突する，(2) 既定の上限ステップ数を超える，(3) 目的座標 $(x_G^{(i)}, y_G^{(i)})$ に到達する，のいずれかの場合に終了する．タスクの終了時刻を T と表わす．以下のパラメータを用いて適応度を計算する．

（1） 全個体平均 FL: $\phi_\mu(T)$

時刻 T での平均 FL.

（2） タスク達成度：$D(T)$

タスクが終了した時点の各ロボットの座標 $(x_T^{(i)}, y_T^{(i)})$ が，どれだけそれぞれの目標座標に近づいていたかを表わす報酬パラメータ．すべてのロボットが目標座標に到達していれば最大値 1 をとり，目標座標から離れるほど小さな値をとる．

$$D(T) = \frac{1}{I}\sum_{i=1}^{I}\Big(|x_G^{(i)} - x_S^{(i)}| - |x_G^{(i)} - x_T^{(i)}| + |y_G^{(i)} - y_S^{(i)}| - |y_G^{(i)} - y_T^{(i)}|\Big) \Big/ \Big(|x_G^{(i)} - x_S^{(i)}| + |y_G^{(i)} - y_S^{(i)}|\Big) \quad (7)$$

（3） タスク達成時間：$R(T)$:

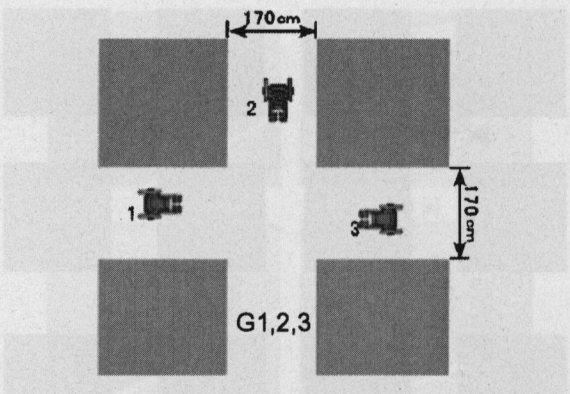


Fig. 7　A mutual concessions task by three robots in crossroad environment (env-2)

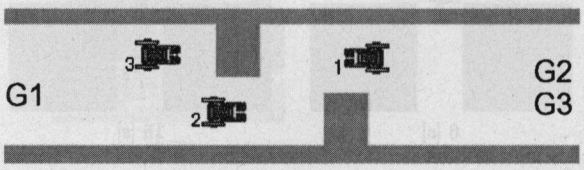


Fig. 8　A mutual concessions task by three robots in aisle with narrow gap (env-3)

既定上限ステップ数（$T_{\max}$）以内に目標座標に到達した場合に与えられる報酬パラメータ．上限ステップで終了した場合に 0，早く到達するほど 1 に近づく．

$$R(T) = \frac{T_{\max} - T}{T_{\max}} \tag{8}$$

（4）　衝突ペナルティ：P:
衝突によりタスクが終了した場合に与えられるペナルティ．衝突の場合に-1，衝突がない場合は 0 とする．
以上の評価値を以下のようにまとめて，個体の適応度とする．

$$F = W_\phi \phi_\mu(T) + W_{\rm d} D(T) + W_{\rm r} R(T) + P \tag{9}$$

ここで，W_ϕ, $W_{\rm d}$, $W_{\rm r}$ は，各評価値に対してそれぞれの評価値を均等に適応度に反映させるためのスケール調整を目的とした重みである．また，$\phi_\mu(T)$ の重みである $W_{\phi'}$ は，集団中で AS からの所望の動作が実行できていないロボットが多数存在する場合のペナルティ項となるため，負の値をとっている．

5.　協調行動の獲得

5.1　実験環境

本手法の有効性を確認するために，シミュレータによる実験を行った．用いた環境は以下の 3 つである．

環境 1　障害物のある通路において 2 台のロボットが同時に進入する場合．(Fig. 2)

環境 2　十字路において，3 台のロボットが同時進入する場合．(**Fig. 7**)

環境 3　障害物のある通路において，3 台のロボットが同時に進入する場合（**Fig. 8**）

Table 2　The parameters for experiments

Item	Value		
Renewal rate α	0.3		
Individuals N	100		
Crossover rate	0.8		
Mutation rate	0.05		
The depth of initial tree	4		
The limitation of depth of tree	7		
Weight parameters	env1	env2	env3
W_ϕ	-0.067	-0.2	-0.2
$W_{\rm d}$	19	24	27
$W_{\rm r}$	5.0	5.0	5.0
$T_{\max}$	100	100	150

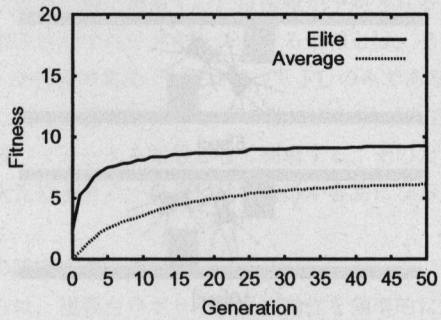


Fig. 9　Changes of fitness of elite and average against generations in env-1 with two robots

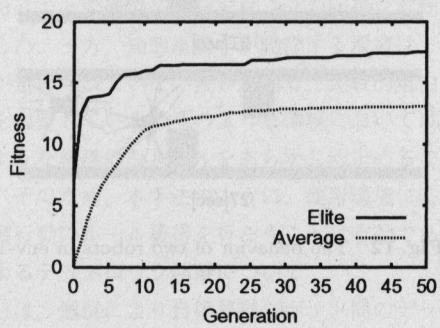


Fig. 10　Changes of fitness of elite and average against generations in env-2 with three robots

それぞれの手法におけるパラメータを **Table 2** に示す．また，本実験では，u_i の値を，$0 < u_i < 1.0$ の範囲で，0.05 刻みの離散値をランダムに選択して設定をしている．

5.2　結果，考察

各実験環境における個体集団の適応度の変化を **Fig. 9**, **Fig. 10**, **Fig. 11** に示す．各図は横軸が世代であり，縦軸は適応度となっている．図中の Elite は，集団内における各世代ごとの最良個体について，15 試行分の適応度の平均値をとった結果である．また，Average は，集団内におけの各世代ごとの全個体の平均適応度について，さらに 15 試行分の平均値をとった結果である．

各環境において世代を経るほど，適応度が上昇し，適切な行動が得られていることがわかる．環境 1 より環境 3 の行動獲得に時間がかかるのは，同じ探索空間でもより深い木を形成する必要があったためである．

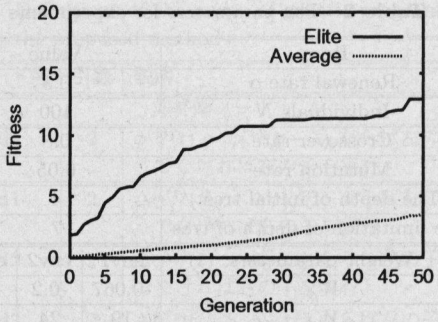


Fig. 11 Changes of fitness of elite and average against generations in env-3 with three robots

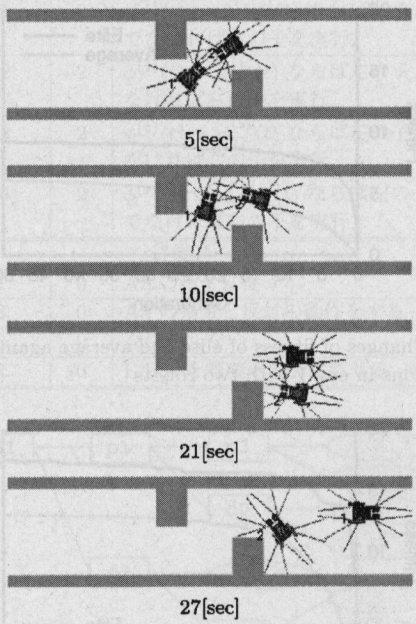


Fig. 12 The behavior of two robots in env-1

シミュレータにおける各協調行動の獲得例を **Fig. 12**, **Fig. 13**, **Fig. 14** に示す．環境 1 では，障害物のある付近でロボット同士が向き合った後，FL の変化に伴い一方のロボットが後退する行動を起こす．その間にもう一方のロボットが狭い箇所を通過し，進路を譲ったロボットが後から通路への進入を図っている．このように，単純な FL のやり取りから，協調的な動作が発現していることがわかる．15 回の試行中，GP がタスク達成可能なエリート個体を獲得する確率は 73%(11/15) であった．

環境 2 では，3 台のロボットが十字路中央で同時に下の通路に進入しようとするためいったん交差点の中央で 3 台とも停止する．しかし，FL の増加に伴い不均衡が生じ，FL が最も早く上昇したロボットを他の 2 台が優先させる振る舞いをとる．その後，1 台ずつ通路に進入する．どのロボットが優先的に通過するかは，タスク開始時にランダムに決定される u_i に依存して異なった結果となる．15 回の試行中，GP がタスク達成可能なエリート個体を獲得する確率は 80%(12/15) である．

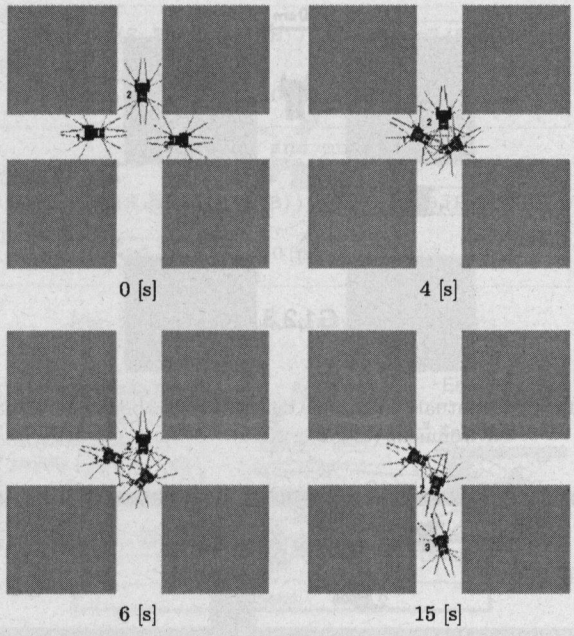


Fig. 13 The behavior of three robots in env-2

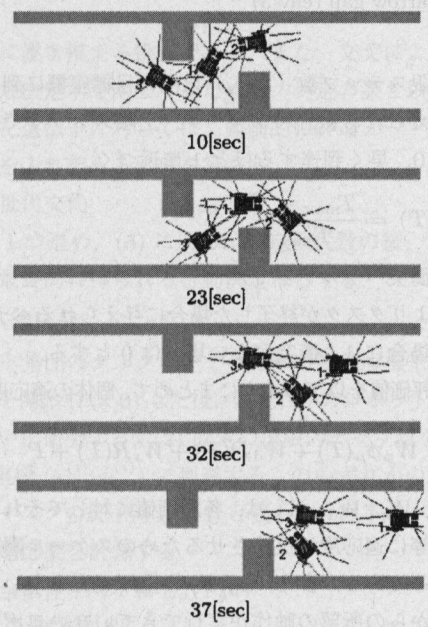


Fig. 14 The behavior of three robots in env-3

環境 3 では，2 台に挟まれる中央のロボット 1 に複雑な振る舞いが要求される．3 台中で唯一進行方向の異なるロボット 2 が，一度後退することで，先にロボット 1,2 を優先的に通過させ，ロボット 2 が最後に通過する．15 回の試行中，GP がタスク達成可能なエリート個体を獲得する確率は 80% (12/15) であった．

さらに，環境 3 で得られたタスク達成可能なエリート個体を用いて，環境 1,2 を走行させた結果，100% の確率 (15/15) でタスクを達成することが確認できた．また，それぞれの環境で 5 台，6 台の例を試行した結果，それぞれ 73%(11/15), 60%(9/15) という成功率が得られた．これより，複雑な環境

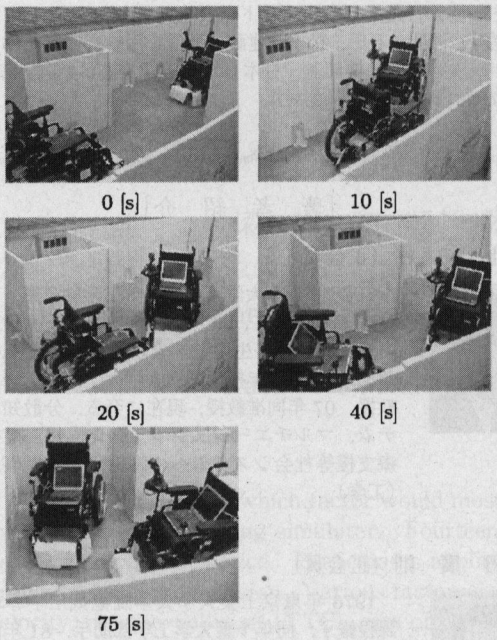


Fig. 15 The behavior of real robots in env-1

にて協調行動を達成可能な決定木は，一定の規模のタスクに対して有効な行動選択機構になるうることが確認できた．

以上の結果中，u_i が互いに近かったり，極端に異なる値が設定されていた場合に，タスクに失敗したり，協調行動に無駄な動きが生じる場合があった．一方，u_i が分散されている場合は，確実に協調行動の獲得ができることが確認できた．

一方，4.2節で述べたように，u_i は，デッドロック状態においてロボットごとのFLの変化パターンを生じさせ，異なる行動の発生を促し，このことがデッドロック回避・解消に役立っている．これより，u_i を決定する際には，近傍ロボット同士で一定のネゴシエーションをはかることが有効だと考えられる．

実機における協調行動の獲得例を **Fig. 15** に示す．実環境は，Fig. 2の障害物のある狭い通路と同様の環境である．行動決定木はシミュレータ実験により求めたエリート遺伝子木を用いる．実機においては，測距センサの特性によるセンサ値のばらつきや，ロボットの左右のモータの誤差などがあるにもかかわらず，シミュレータと同様の協調行動が得られた．

通過にかかる時間をシミュレータ実験と比較すると，実機が通路を通過するためには，シミュレータ実験の3倍弱の時間が必要であった．これは，通路の両端が壁により区切られていたため，前後に別の障害物があるような振る舞いになったこと，シミュレータ上では問題にならなかったロボット間のセンサ干渉の影響を避けるために，測距を複数回繰り返す操作が必要だったためである．

6. 先行研究事例との関係

本研究では，局所的な環境情報を介したコミュニケーションによる衝突回避等の協調行動を獲得しているが，先行研究においても類似の手法がされている[18)~21)]．これらの先行研究事例との相違点について考察を加える．

ロボットの状態をブロードキャストする手法について

油田らは，複数ロボットによる空間資源共有のモデルとして，ロボット間の一対一通信，一対多通信を組み合わせた自律ロボットネットワークを提案した[18),19)]．各ロボットは「自身の状態」をブロードキャストし，集団内で情報を共有する．この点において，本手法は共通の枠組みを有している．一方，先行研究においては「自身の状態」の明確な定義はなされておらず，複数ロボットが存在する場合の集団の状態空間は大きな次元を有することが予想される．そのため，状態空間から適切な協調行動を獲得するには困難が予想される．これに対し，本手法がブロードキャストする情報とは，それぞれのロボットの行動の成功・失敗に基づくFLのみである．FLを受信した後は，近傍のロボットの数によらず(1)-(6)式で示した環境パラメータを状態として利用する．そのため，扱う情報の次元は小さく，協調行動の獲得が容易になる点が有利である．

ルールの設定方法について

加藤らは，複数台ロボットの協調動作を効率的に実現するために，交通ルールを設計・構築することを提案した[20)]．交差点において，あらかじめ移動方向や位置による優先順位を決めておくことで衝突・デッドロックの回避に有効であることを示した．一方，知的車椅子が動作する環境は，明確な移動規則が定められていない狭い屋内で，比較的自由に移動する用途を想定している．そのような環境においては，道路交通流のような組織的なルールをあらかじめ定めることは困難である．そのため，本手法のように，使用環境に応じた適応的な協調行動のルール獲得を行なうことは有効である．

交渉によるデッドロック回避について

尾崎らは，通信により自律移動ロボット間のデッドロックを解消する交渉プロセスを提案した[21)]．衝突対象となるロボットを特定し，そのIDと位置をもとに交渉を行なうことでデッドロックを解消できる．しかし，扱われている衝突回避は1対1の関係のみであり，3台以上の構成になった場合は交渉が複雑になることが予想される．これに対し本手法は，近傍の知的車椅子の存在は，ブロードキャストされたFLから導き出した環境パラメータを介して間接的に知りうるが，そのIDや位置関係の情報を必要としていない．このように，本手法は，台数が増えても処理の複雑さが変化しない点が特長である．

7. むすび

知的車椅子の協調行動創発のために，フラストレーションレベルと行動決定木を用いた行動決定機構を提案した．本手法の特徴は，局所的な情報と単純なアルゴリズムから，いくつかの単純な環境に対して，有効な協調行動を自律的に獲得できる点である．シミュレータ上に作成された障害物のある狭い通路と十字路の通路を用いた計算機実験の結果，GPの

進化を経て効果的な行動決定木を得ることができた．更に，シミュレータで得られた知的エージェントを用いて，実機においても協調行動が得られることを確認した．

今後は，集団中における u_i の組み合わせに依存しない改良を検討を進めるとともに，本手法を基礎としたより複雑な協調行動の獲得を検討する．

参 考 文 献

1) 平成 18 年版高齢社会白書，内閣府 (2006)
2) 阿部，濱上，小圷，平田：ティーチングと Q 学習を併用した知的車椅子の自律走行，電気学会論文誌 D, **123**-11, 1384/1385 (2003)
3) 濱上，平田：知的車椅子における環境適応のための状態空間の構成法，電気学会論文誌 D, **123**-10, 1118/1124 (2003)
4) T. Hamagami and H. Hirata: Development of Intelligent Wheelchair Acquiring Autonomous, Cooperative, and Collaborative Behavior, Proc of IEEE International Conference on Systems, Man and Cybernetics, 3235/3530 (2004)
5) 吉岡，登尾：障害物の幾何学的性質にもとづくセンサベーストデッドロックフリーナビゲーションアルゴリズム，計測自動制御学会論文集，**33**-4, 280/288 (1997)
6) 足立，中西，久野，島田，白井：顔の視覚情報処理を用いた知的車椅子，日本ロボット学会誌，**117**-3, 423/431 (1999)
7) 小林，河野，木戸出：館内誘導システムのための複数台ロボット協調機構の提案，第 2 回計測自動制御学会システムインテグレーション部門学術講演会，**1**, P2_32 (2001)
8) M. Quin, L. Smith, G. Mayley and P. Husbands: Evolving Teamwork and Role-Allocation with Real Robots, Proceedings of the 8th International Conference on Artificial Life, 302/311 (2002)
9) 小林，細江：マルチエージェントシステムの自律分散—群移動ロボットによる協調捕獲行動，計測と制御，**46**-3, 178/184 (2007)
10) R.A. Brooks: Intelligence Without Representation, Artificial Intelligence Journal, **47**, 139/159 (1991)
11) R.A. Brooks: A Robust Layered Control System for a Mobile Robot, IEEE J. Robotics and Automation, **2**-1, 14/23 (1986)
12) 今井，川戸，濱上，平田：知的車椅子のための進化型シミュレータと行動統括機構の開発，電気学会論文誌 D，**125**-1, 67/75 (2005)
13) 大沢英一：マルチエージェント環境における交渉のモデル，人工知能学会誌，**10**-5, 690/696 (1995)
14) 森山，沼尾：環境状況に応じて自己の報酬を操作する学習エージェントの構築，**17**-6, 676/683 (2002)
15) O. Holland and C. Melhuish: Stigmergy, Self-Organization, and Sorting in Collective Robotics, Artificial Life **5**-2, 173/202 (1999)
16) V. Trianni, T.H. Labella and M. Dorigo: Evolution of direct communication for a swarm bot performing hole avoidance, In Ant Colony Optimization and Swarm Intelligence, Springer (2004)
17) 伊庭斉志：遺伝的プログラミング，東京電気大学出版局，東京 (1996)
18) プレムウッティ，油田：複数台ロボットによる空間資源共用のための自律ロボット間通信ネットワーク，第 9 回日本ロボット学会学術講演会講演論文集，595/598 (1991)
19) S. Yuta, S. Premvuti and S. Yuta: Coordinating Autonomous and Centralized Decision Making to Achieve Cooperative Behaviors Between Multiple Mobile Robots IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) '92, 1566/1574 (1992)
20) 加藤，西山，武野：交通ルールの適用による複数の移動ロボットの協調行動，日本ロボット学会誌，**12**-2, 785/790 (1994)
21) 尾崎，淺間，石田，松本，遠藤：通信を用いた複数自律移動ロボットの相互衝突回避，日本ロボット学会誌，**14**-7, 961/967 (1996)

[著 者 紹 介]

濱 上 知 樹（正会員）



1988 年千葉大学工学部電気工学科卒業．セコム (株) 勤務を経て，99 年在職のまま千葉大学大学院自然科学研究科生産科学専攻後期課程修了．2001 年同助手．04 年横浜国立大学大学院工学研究院助教授．07 年同准教授，現在に至る．分散知能システム，マルチエージェント，強化学習，福祉・医療支援等社会システムへの応用研究に従事．博士（工学）．

平 田 廣 則（正会員）



1976 年東京工業大学大学院電気工学専攻博士課程修了．同年千葉大学工学部助手．81 年同助教授．94 年同教授．97 年同大学院自然科学研究科教授．2007 年同大学院工学研究科教授，現在に至る．大規模システムのモデリング，解析と設計，特に生態システム，VLSI レイアウト，並びに分散型コンピュータの基礎的理論の研究に従事．工学博士．