

動作合成の反復によるロボット協調行動の自律的実現

五十嵐 祐樹^{*1} 根谷 清順^{*1} 鈴木 昌和^{*2}

Autonomous Realization of Cooperative Robot Actions by Iterative Motion Composition

by

Yuki IGARASHI^{*1}, Seijun NEYA^{*1} and Masakazu SUZUKI^{*2}

(Received on Mar. 31, 2009 and accepted on May 25, 2009)

Abstract

This article presents a method for autonomous realization of cooperative robot actions by iterative motion composition. Using the Intelligent Composite Motion Control (ICMC), complex actions are gradually realized from fundamental motions by repeating optimal motion compositions. Although it is not easy to obtain the action intelligence, which gives adequate control from a given situation, of a complex cooperation at once, the intelligence is constructed in a multi-stage manner using the Multi-stage Genetic Algorithm (MGA). MGA solves the cooperative action realization problem as iterative combinatorial optimization. The empirical knowledge obtained is effectively utilized to solve similar problems efficiently and moreover for more complex actions. By applying the method to complex cooperative robot soccer actions, it has been shown that the action intelligence with large applicability can be efficiently obtained by the learning for only typical situations. And the applicability can be adaptively enlarged with additional learning according to need. In addition, in order to make the action planning more versatile, the Variable-chromosome-length Genetic Algorithm (VGA) is introduced and combined with MGA. It was successfully demonstrated that a wide variety of optimal cooperative robot soccer actions where the number of element motions is not fixed in advance can be autonomously planned according to the situation.

Keywords: Multiple robots, Cooperative action, Autonomous planning, Multi-stage optimization, Genetic algorithm

1. はじめに

知能ロボットへの要求は高まるばかりである。さまざまな複雑タスクに対して、人のように自律的・適応的に行動を計画し実行する高度に知的なロボットの実現が求められている。ロボットシステムに実用的なタスクを実行させることは非常に多くの決定変数をもつ大規模最適化問題を解くことに相当し、設計者があらゆる複雑行動に対して大規模問題を解いて行動の細部まで計画し、制御を求めておくことは現実的ではない。複数ロボットによる協調行動では問題は一層大規模複雑となるが、複雑な行動も比較的単純な動作の組合せが基本になっていることが多く、したがってそれらの単純動作の実現をサブタ

クとして最適実現して組み合わせる方法が有効な手法となる。^{1)~3)}我々はそのような方法論の1つとして、基本的な動作群の学習から始め、それらの合成と継続的な学習により漸次的に複雑な行動を実現する進化的学習法である知的合成動作制御法 (ICMC: Intelligent Composite Motion Control)^{4)~6)}を提唱している。ICMCは、獲得した行動を同じ状況で再現するだけでなく、類似状況でも準最適に実行でき、また、より複雑な新しいタスクを既獲得のサブタスクを活用して効率的に実現できる。しかし、大規模タスクを実現する行動知能を状況変化に対応できる形で最適構成することは容易ではない。しかも実際の計算時間での効率的意思決定や、タスク実行中の状況変化への柔軟な適応が必要となる。そこでこのような行動知能構成に遺伝的アルゴリズム (GA) の拡張で上記のような特性を備えた大規模最適化問題の効率的近似解法

^{*1} 工学研究科機械工学専攻修士課程

^{*2} 工学部航空宇宙学科教授

である多段階遺伝的アルゴリズム (Multi-stage Genetic Algorithm, MGA)⁷⁾を用いた複数ロボットによる協調行動実現法の検討を進めている。これまでに MGA を合成型大規模ロボット行動最適化手法として用いることを提案し⁸⁾, GA を 3 段階に用いた 3 台の移動ロボットによる協調サッカー行動を実現した⁹⁾。

本論文では、行動知能を動作合成の反復により漸次的に構成する本手法が、複雑協調行動を効果的に最適計画でき、かつ、継続的な学習により適用可能範囲を必要に応じて徐々に拡張して複雑行動の行動知能を効率的に構成できることを複数ロボットによる協調サッカー問題に適用した例を通して示す。更に、染色体長の固定された GA では協調行動として最適合成する要素動作数が限定されるという問題点の解決法として、染色体長を可変とした GA を用いる協調行動計画法を提案し、ロボットが自律的に状況に応じて多様な協調行動を計画できることを示す。

以下では、まず知的合成動作制御法に基づいて複雑行動が動作合成の反復によりどのように実現されるかを述べ、行動知能の効率的構成法として多段階遺伝的アルゴリズムを紹介する。次に車輪移動ロボットの協調ブロックセンタリングシュートの実現過程を例に、行動知能の獲得・蓄積・活用法を述べる。そして、知能適用可能範囲の漸次的拡張、可変染色体長 GA による多様な協調の自律的計画などを提案手法の有効性を示す。

2. 動作合成の反復と行動知能の多段階最適化

2.1 ロボット行動知能とその多段階構成

知的合成動作制御法 (ICMC)^{4)~6)}は、基本的な動作の制御さえ与えれば自律的適応的に複雑で器用な行動を実現できるロボットを目指した知能化法で、未経験状況への適応や、経験知識を活用したより複雑な行動の効率的実現に特徴がある。ある複雑な行動は可調整制御パラメータをもつ複数のより単純な要素動作の合成として定式化され、望ましい行動はパラメータ最適化により実現される。この最適動作合成により獲得された経験知識は行動知能として保存・活用される。

ICMC における行動知能とは、状況を規定する環境情報等が与えられると、その状況で最適な行動を実現する制御情報を出力する状況から制御への写像である。与えられた状況から、その状況で望ましい行動を選択して行動内容詳細を指定する行動選択知能と、それにより指示された行動内容を与えると、その行動を最適に実現する制御パラメータを出力する行動制御知能とに分けて考えられるが、両者の明確な分離は簡単ではない。本稿では、これらを単に行動知能と呼ぶ。

行動知能の多段階構成 幅広い状況で複雑な行動を最適に実行できるような行動知能の構成は容易でなく、多段階構成が有効である。ここで言う行動知能構成の多段

階性の意味は、大まかに以下の 3 つに分けられる。

(1) 行動生成の多段階性: 最初から複雑タスクの行動知能を構成するのではなく、まず実現の容易な単純動作の行動知能を構成し、これらの合成を繰り返して段階的に複雑タスクの行動知能を構成する。

(2) 適用範囲拡張の多段階性: あらゆる状況でタスクを実行できる行動知能を一度に構成するのではなく、まず典型的な (あるいは遭遇した) 状況に対する行動知能を構成し、これを段階的に幅広い状況に対する行動知能へ拡張する。

(3) 知能精緻化の多段階性: 目的タスクを最適実現する行動知能を一気に構成するのではなく、まずは計算時間などの制約のもとで準最適行動のための行動知能を構成し、これを段階的に精緻化してゆく。

このような多段階構成が可能となるのは、最適動作合成の反復により複雑行動を漸次実現する ICMC の採用が本質的であるが、次項の多段階遺伝的アルゴリズムがその効率的構成に大きく貢献する。

多点探索手法による効率的知能構成 行動知能の構成には遺伝的アルゴリズム (GA) など求解過程で容易に準最適解集合が得られる多点探索手法を用いると効率的である。行動内容を規定する動作パラメータ \mathbf{b} を入力すると、それを最適実現する制御パラメータ \mathbf{a} を出力する行動知能を考えよう。最適実現問題 \mathbf{P} の準最適解集合 $\mathcal{X}_{\mathbf{P}} = \{\mathbf{a}^i | i = 1, 2, \dots, m\}$ は動作 \mathbf{b} を最適実現する制御 \mathbf{a}^* およびその近傍点 \mathbf{a}^i からなる集合である。“類似した制御により達成される動作結果は類似し、逆も成り立つ” ような状況を考えると、各 \mathbf{a}^i は \mathbf{b} に類似したある動作 \mathbf{b}^i を最適実現する制御となっている (Fig.1 左)。したがって $\mathcal{X}_{\mathbf{P}}$ に適当な処理を加えることで、これを $\mathbf{b} \in \mathcal{D}_{\mathbf{b}}$ から $\mathbf{a}^*(\mathbf{b}) \in \mathcal{R}_{\mathbf{a}}$ を近似的に生成する写像として活用できる (Fig.1 右)。このように特定の \mathbf{b} に対する最適化問題を多点探索手法により解くことで、各々の \mathbf{b}^i に対する最適化問題を解くことなく効率的に行動知能を構成できる。

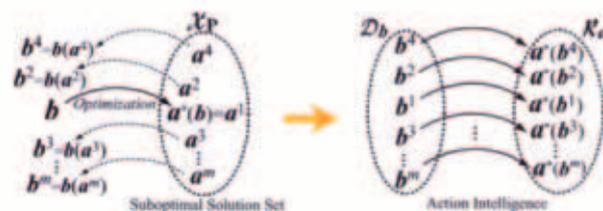


Fig. 1 Suboptimal solution set and its transformation to the action intelligence in the form of an approximate mapping

単純な動作の行動知能同士を最適合成することで、複雑な動作を実現でき、これにより得られる準最適解集合も行動知能として蓄積してより複雑な動作に用いる。こ

のように行動知能の合成を多段階に繰り返すことにより複雑な協調行動を系統的に実現できる。類似状況に遭遇した場合には実現問題を解き直さずとも、すでに得られている行動知能を活用して効率的に複雑行動を実現できる。様々な状況に対する行動知能を蓄積し続けることにより、複雑協調行動を効率的に獲得できるようになる。

2.2 多段階遺伝的アルゴリズム - MGA -

多段階遺伝的アルゴリズム (MGA)⁷⁾は、GA を多段階に用いて合成型最適化問題を近似的ではあるが効率的に解く方法であり、複雑な制約条件を課されても効率的に許容解を生成できる、類似問題を効率的に解ける等の特徴をもつ。

MGA では、まず最適化問題 $P_k, k = 1, 2, \dots, N$

$$P_k : \min_{\mathbf{x}_k \in X_k} f_k(\mathbf{x}_k) \\ \text{subj. to } \mathbf{x}_k \in G_k$$

が適当なパラメータ δ_k, m_k を用いて GA により解かれ、決定変数 $\mathbf{x}_k \in R^{m_k}$ の準最適解集合

$$X_k^* = \{\mathbf{x}_k^{i_k}, i_k = 1, 2, \dots, m_k\} \quad (1)$$

が求められているものとする。ここで δ_k は X_k^* の要素に多様性をもたせるための個体間最小距離 (一般には擬距離) であり、問題に応じて必要な多様性を確保できるよう定義される。各 X_k^* は複雑な問題を解く際の近似解の基底を与えることになる。解くべき主問題は次の最適化問題 P である。

$$P : \min_{\mathbf{x} \in X} f(\mathbf{x}) \\ \text{subj. to } \mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \in X \\ \mathbf{x}_k \in R^{m_k}, k = 1, 2, \dots, N \\ \mathbf{x} \in G_0 \\ \mathbf{x}_k \in G_k, k = 1, 2, \dots, N$$

決定変数 (全体変数) \mathbf{x} は各部分システムの変数 (部分変数) \mathbf{x}_k からなり、同様に制約条件は全体制約 $\mathbf{x} \in G_0$ と部分制約 $\mathbf{x}_k \in G_k$ からなる。

MGA では P を直接解くのではなく、 P_k で得られている X_k^* の要素を最適に合成する次の問題 P^Σ を解くことにより P を近似的に解く。

$$P^\Sigma : \min_{\mathbf{x} \in X^\Sigma} f(\mathbf{x}) \\ \text{subj. to } X^\Sigma = \{\mathbf{x} \mid \mathbf{x}_k \in X_k^*, k = 1, 2, \dots, N\} \\ \mathbf{x} \in G_0$$

P あるいは P^Σ の決定変数の次元は、各 P_k の次元の和 $\dim \mathbf{x} = \sum_{k=1}^N \dim \mathbf{x}_k$ であり大次元だが、 \mathbf{x} は表現型であり、その遺伝子型 \mathbf{g} を

$$\mathbf{g} = [i_1 \ i_2 \ \dots \ i_N] \in R^N, i_k \in \{1, 2, \dots, m_k\} \quad (2)$$

として、表現型との対応関係を

$$\mathbf{x}(\mathbf{g}) = \{\mathbf{x}_1^{i_1} \mathbf{x}_2^{i_2} \dots \mathbf{x}_N^{i_N}\} \quad (3)$$

とすれば、 P^Σ の解は \mathbf{g} で規定される。すなわち P^Σ は、実際には次のようなわずかな N 次元の整数値組合せ最適化問題として容易に解くことができる。

$$P^\Gamma : \min_{\mathbf{g} \in \Gamma} f(\mathbf{x}) \\ \text{subj. to } \mathbf{x}(\mathbf{g}) = (\mathbf{x}_1^{i_1}, \mathbf{x}_2^{i_2}, \dots, \mathbf{x}_N^{i_N}) \in X^\Sigma \\ \Gamma = \{\mathbf{g} = [i_1 \ i_2 \ \dots \ i_N] \mid i_k \in \{1, 2, \dots, m_k\}; \\ k = 1, 2, \dots, N\} \\ \mathbf{x} \in G_0$$

P を解く過程で得られる準最適解集合 X^* も保存し、より複雑な問題の近似解の基底に用いる。このプロセスを多段階繰り返すことにより徐々に複雑な合成型大規模問題を解くことができる (Fig.2)。MGA の求解効率を

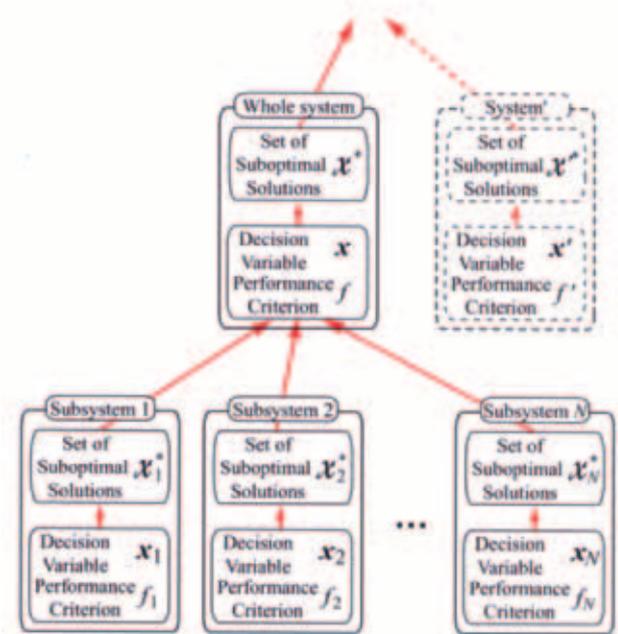


Fig.2 Synthetic solution in MGA for compositional large scale optimization problems

通常の単段階 GA と比較すると、MGA では決定変数が多いほど、また制約条件が複雑なほど問題を効率的に解ける^{8),9)}。また、適切な仮定をおけば、単段階 GA と比較して精度が低下することもない⁷⁾。

3. 協調ロボットサッカー行動の実現

3.1 問題設定と状況の表現

ここでは、Fig.3 に示すように、2 台の敵ロボット (E_1, E_2) が存在し、ロボット 1 (R_1) が E_1 をブロックし、ロボット 2 (R_2) がロボット 3 (R_3) にパスし、 R_3

$a_P \in G_P$ は旋回運動後に直進運動でキック可能であることを要求する。 P_{Pass} を GA で解くことで $\mathcal{A}_P^* = \{a_P^i\}$ を得る。ただし個体間最小距離 δ としてボール到達位置間距離を設定し、準最適解集合要素数 $m = 30$ とした (以下の問題でも同様)。

Fig.6 の左図は \mathcal{A}_P^* の各要素 a_P^i を用いた Pass による所定速度でのボール到達位置 $p_{B_i}^f$ の分布の例を示している。右図は左図の点線内を拡大したもので、 $J_P(a_P)$ のうち $\|p_B^f - p_B^d\|$ の評価部分の等高線を加えてある。また、数字 i は a_P^i を用いた時のボール到達位置であり、例えば 6 は $p_{B_6}^f(a_P^6)$ の位置を示す。

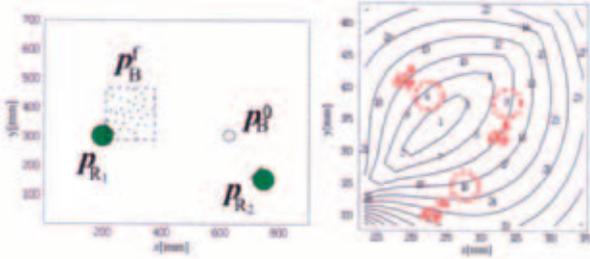


Fig. 6 Attained ball positions by Suboptimal Pass actions

シュート動作 動作最適化問題 P_{Shoot} を

$$P_{\text{Shoot}} : \min_{a_S \in \mathcal{A}_0} J_S(a_S) \\ \text{subj. to } a_S \in G_S$$

のように定式化する。 J_S は動作の迅速さと所定ボール速度でのゴール成否を評価している。 G_S は P_{Pass} の G_P と同じとする。 P_{Shoot} を GA で解くことにより $\mathcal{A}_S^* = \{a_S^i\}$ を得る。 Fig.7 は Fig.6 と同様、左図は \mathcal{A}_S^* の各要素を用いた Shoot によるボール到達位置の分布例、右図は左図点線内の拡大である。数字 i は $p_{B_i}^f(a_S^i)$ の位置を示す。

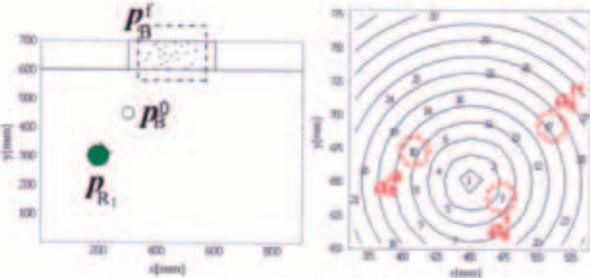


Fig. 7 Attained ball positions by Suboptimal Shoot actions

ブロック動作 動作最適化問題 P_{Block} を

$$P_{\text{Block}} : \min_{a_B \in \mathcal{A}_0} J_B(a_B) \\ \text{subj. to } a_B \in G_B$$

のように定式化する。 J_B はブロックの最終位置 p_R^f がボールに近く効果的なブロック位置 (l_{EB} 上) に素早く移動するほど高評価になる。 $a_B \in G_B$ はブロックと敵の衝突を防ぎ、かつ、ブロックが有効ブロック領域内に達することを要求する。 P_{Block} を GA で解くことで $\mathcal{A}_B^* = \{a_B^i\}$ を得る。

3.3 協調行動の合成的実現

センタリングシュート行動 \mathcal{A}_P^* と \mathcal{A}_S^* の要素を最適合成する問題 P_{CShoot}^Σ を次のように定式化する。

$$P_{\text{CShoot}}^\Sigma : \min_{a_C \in \mathcal{A}_C} J_C(a_C) \\ \text{subj. to } a_C = (a_P, t_S, a_S) \in G_C \\ \mathcal{A}_C = \{a_C | a_P \in \mathcal{A}_P^*, a_S \in \mathcal{A}_S^*\}$$

J_C は行動の迅速さとゴールの成否を評価し、 t_S はパスされたボールとのタイミングを合わせる Shoot 開始時刻である。 $a_C \in G_C$ は敵位置 p_E から一定距離離れた安全なボール軌跡を要求する。 Fig.6, Fig.7 に示したように、 \mathcal{A}_P^* , \mathcal{A}_S^* は多様な動作結果を与える制御パラメータを含むので、これらの要素を組み合わせることで Pass, Shoot を学習した状況を組み合わせた単一状況に対してだけでなく広い状況に対して適切な CShoot を実現できる。 P_{CShoot}^Σ を GA で解くことにより $\mathcal{A}_C^* = \{a_C^i\}$ を得る。

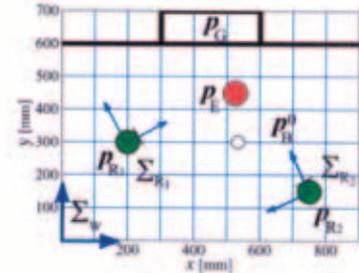


Fig. 8 Situation setting for Centering-Shoot action

ブロックセンタリングシュート行動 \mathcal{A}_B^* と \mathcal{A}_C^* の要素を最適合成する問題 $P_{\text{BCShoot}}^\Sigma$ を次のように定式化する。

$$P_{\text{BCShoot}}^\Sigma : \min_{a_{BC} \in \mathcal{A}_{BC}} J_{BC}(a_{BC}) \\ \text{subj. to } a_{BC} = (a_B, a_C) \in G_{BC} \\ \mathcal{A}_{BC} = \{a_{BC} | a_B \in \mathcal{A}_B^*, a_C \in \mathcal{A}_C^*\}$$

J_{BC} はゴールの成否を評価している。 $a_{BC} \in G_{BC}$ は R_1 と R_2, R_3 との衝突を防ぐ制約である。 Fig.3 の状況に対して実現されたブロックセンタリングシュート行動をその動作合成の過程と共に Fig.9 に示す。 図中の矢印はその行動がどの要素動作の合成であることを示している。

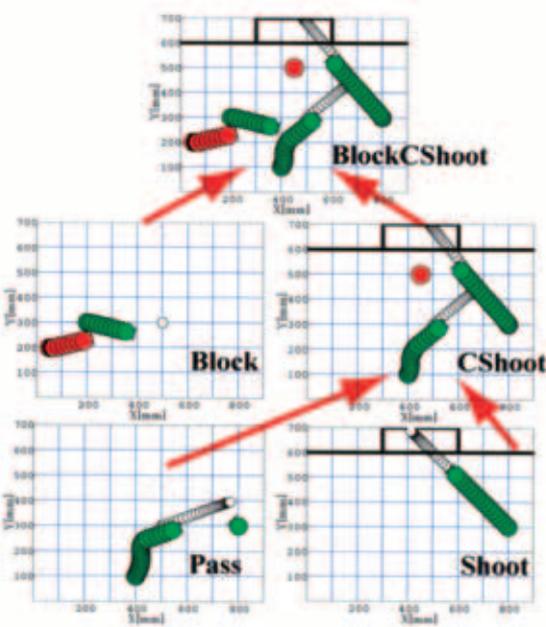


Fig. 9 Motion compositions and attained Block-Centering-Shoot action

4. 行動知能適用可能性とその適応的拡張

Fig.9のように協調行動が実現できたのは動作合成に必要な要素動作の行動知能が構成済みであったからであり、あらゆる状況に対してこれを実現するためにはあらゆる状況に対して全ての要素動作が学習済みでなければならない。しかし、幅広い状況に対して行動知能を構成するのは容易ではないから、まずは典型的な状況に対するタスクを実現可能としておき、必要に応じて適応的に適用範囲の拡張と知能の精緻化を進める。急ぐときは最適性は下げても速やかに意思決定しなければならない。場合によっては制約条件を緩和し、許容領域を拡げてでも解を求めなければならない。本節ではPassとShootで共通の運動となっているキック動作(Kick)を例に行動知能の適用可能性とその適応的拡張について述べる。

4.1 行動知能の適用可能性

ここでは、ある p_B^0 に対して $\|p_B^f - p_B^t\| < \epsilon_K$ ($\epsilon_K > 0$) となる制御パラメータ a_K が得られたとき、状況 $s_K = (p_B^0, p_B^t)$ におけるKickは実現されたと表現する。更に行動知能 $\mathcal{A}_K^* = \{a_K^i\}$ の要素のいずれかが s_K におけるKickを実現するとき \mathcal{A}_K^* はこの状況に適用可能であると言う。

状況 s_K における動作最適化問題 P_{Kick} を次のように定式化する。

$$P_{\text{Kick}} : \min_{a_K \in \mathcal{A}_0} J_K(a_K) \\ \text{subj. to } a_K = [\omega_1^T t_1 \ \omega_2^T t_2 \ \omega_3 \ t_3] \in G_K \cap G_t$$

J_K は素早い動作で、上手くボールを目標位置 p_B^t に蹴ると高評価となる。 G_K は P_{Pass} における G_P と同じとする。また p_B^0, p_B^t は極座標表示する。 G_t は現実的な動作時間で動作実現するための制約で

$$G_t = \{a_K \mid 0 \leq t_i \leq t^+, i = 1, 2, 3\} \quad (5)$$

とし、以下では $t^+ = 3.0$ とする。

Fig.10に示すような p_B^0, p_B^t について各12通りの組み合わせ、すなわち $12 \times 12 = 144$ 通りの典型的な状況の集合 \mathcal{S}_K^{144} を考え、 \mathcal{S}_K^{144} の全状況に対して P_{Kick} を

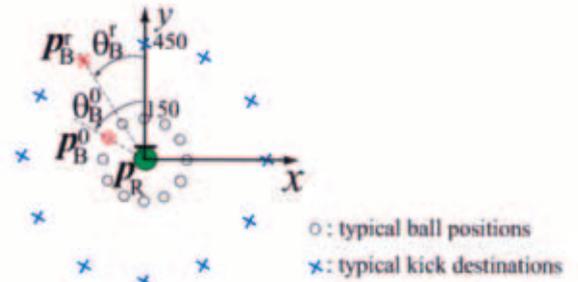


Fig. 10 Typical situations for Kick action

解いて得た \mathcal{A}_K^* によるボール到達位置の分布を調べた結果、144のうち51の状況でKickを実現できなかった。Fig.11に、典型的なボール方向 $\theta_B^0 = 0, \frac{\pi}{3}, \frac{2}{3}\pi, \pi$ の状況に対するボール到達位置分布を示す。動作を実現できなかった状況の p_B^t や分布範囲を太線で表した。頻繁に遭遇する基本的な状況に対しては十分に適用可能であるが、後方のボールを前方へ蹴るような難しい動作が実現できていない。

4.2 適用可能性の適応的拡張

最大動作時間と適用可能性 多くの状況でKickが実現できないのは、最大動作時間 t^+ の設定が原因と考えられたため、 $t^+ = 3, 7, 11$ [sec] とし、Fig.11の $p_B^0 = (150, \pi)$, $p_B^t = (450, \frac{\pi}{3})$ の状況でKickを最適化した(Fig.12)。 t^+ が大きくなるに従って、 p_B^t が p_B^0 に近づくことが分かる。そこで $t^+ = 7$ [sec] とし、 \mathcal{S}_K^{144} に対して動作最適化を行い、そのうち Fig.11と同じ状況に対するものを Fig.13に示す。144のうち13の状況で動作を実現できなかったが、最大動作時間を大きくすることで動作実現できない状況が大幅に減少し、キック動作知能の適用可能性が拡張された。

学習済状況数と適用可能性 ここでは基底動作を学習する状況数を変化させた場合の合成動作の実現可能性の変化を調べた。典型的なCShootの状況を10通り設定し、これを基にロボットやボールの位置・角度に摂動を加え、100通りのCShootの状況集合 \mathcal{S}_{CS}^{100} を設定し、典型的状況に対するキック動作知能のみで \mathcal{S}_{CS}^{100} の状況の

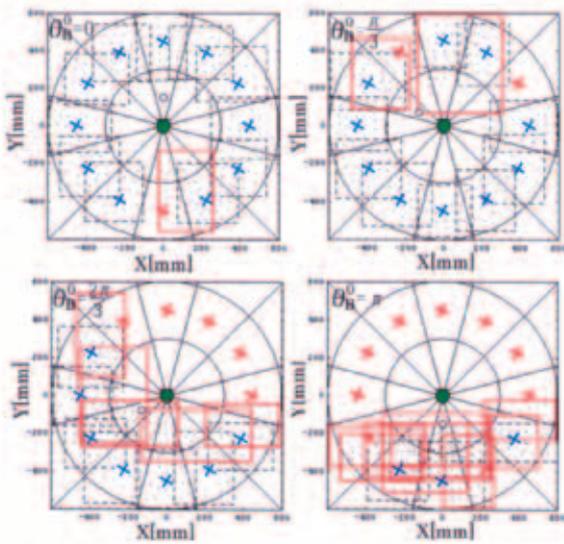


Fig. 11 Distribution of attained ball positions by Kick action ($t^* = 3[\text{sec}]$)

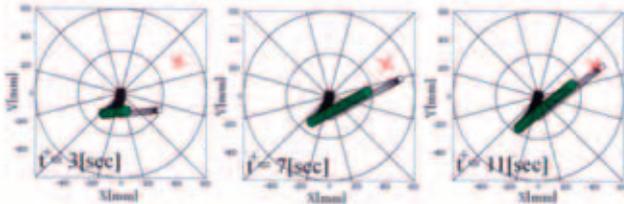


Fig. 12 Maximum action time and optimum kick action

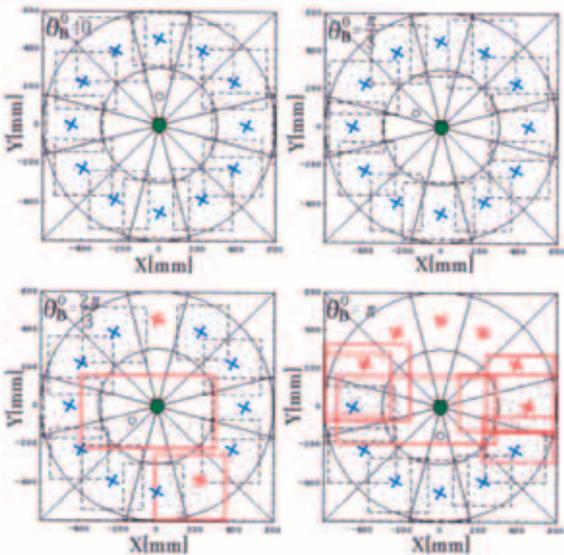


Fig. 13 Distribution of attained ball positions by Kick action ($t^* = 7[\text{sec}]$)

どれ程に対して CShoot が成功するかを検討した。 S_{CS}^{100} に対して CShoot が成功する状況の割合を適用可能性と定義し、学習した状況数との関係を示すと Table 1 のよ

うな結果となった。学習した状況数が多い程、CShoot の実現可能性が高いことが分かる。すなわち、必要に応じて基底動作の追加学習を継続することで行動知能の合成動作への適用可能性を高められる。

Table 1 Number of learned situations for the base action vs the applicability to composite Centering-shoot action

p_B^0	p_B^r	No. of learned situations	Applicability
4	4	16	29 %
6	6	36	51 %
12	12	144	59 %

5. 可変染色体長 GA による多様な行動の計画

5.1 可変染色体長遺伝的アルゴリズム -VGA-

3 節では Fig.3 の状況でブロックセンタリングシュートが最も望ましい行動と判断されたものとしている。選択された行動が与えられた状況で最も有望なものとして行動内容を最適化するには、染色体長を固定して GA を適用すればよい。しかし、状況によって 1 台のロボットによる単独行動が望ましいことも複数ロボットでの協調行動が望ましい場合もあり、異なる数の要素動作を最適に組み合わせる必要がある。そこで要素動作数を限定することなく多様な行動を計画できるように可変染色体長遺伝的アルゴリズム (VGA: Variable-chromosome-length Genetic Algorithm) ¹⁰⁾ を用いる。

VGA における染色体 (決定変数) x は遺伝子座数 N と各遺伝子情報 $x_k \in R^{n_k}$ を組み合わせて

$$x = (N, x_1, x_2, \dots, x_N) \quad (6)$$

と表す。このとき各個体の染色体長 l_c は $l_c = 1 + \sum_{k=1}^N n_k$ となる。VGA のアルゴリズムは基本的には通常の GA と同じであるが、交叉は染色体長を変化させる切断・接合を含んだより一般的な操作となる。

5.2 VGA による協調サッカー行動の自律計画

以下では VGA を 2 対 1 協調サッカー行動問題に適用する。ただし実行可能な要素動作は Pass, Block, Shoot の 3 種とし、これらの行動知能 (準最適解集合) は予め MGA で求められ保存されているものとする。

協調行動の染色体表現 要素動作種を一般に A と表してロボット R_n による動作 A を $R_n A$ と表現し、動作主ロボットと要素動作種を組み合わせた 6 つの対立遺伝子を考える (Table 2)。現実的な協調サッカーの実現のため $R_n B$ と $R_{\bar{n}} S$ など同時に行える動作は同時実行を許容する ($R_{\bar{n}}$ は R_n 以外のロボットを表す)。染色体 a は、協調行動を構成する要素動作数 N と要素動作を特定する遺伝子情報を要素動作の実行順に並べたものとし、一般に次の形をもつ。

$$\mathbf{a} = (N, R_{n_1} A_1, R_{n_2} A_2, \dots, R_{n_N} A_N) \quad (7)$$

Table 2 Element actions in 2 vs 1 cooperative soccer action optimization by VGA

Action	Gene
Pass by Robot 1/Robot 2	1P/2P
Block by Robot 1/Robot 2	1B/2B
Shoot by Robot 1/Robot 2	1S/2S

動作実行順序の制約とその取り扱い 協調行動において R_n によるある要素動作の後に実行可能な要素動作には制限がある。ボールを持たねば **Pass** や **Shoot** を実行できない。また、ここではゴールを最終目的としているので **Shoot** が各個体の最終遺伝子座になければならない。VGA においてはこのような先行動作-後続動作の順序関係や最終遺伝子座は制約条件として取り扱う。

5.3 MGA を用いた個体評価法

(7) は協調行動においてどのロボットによるどの動作がどのような順序で実行されるかを規定するのみであり、各要素動作の具体的な内容は未定で、協調行動の定量的評価もできない。そこで既に MGA で求められている要素動作の行動知能を活用して、計画した協調行動の評価値を定量的に推定する。

状況 \mathbf{s} を行動開始時の Σ_W での各種位置情報により $\mathbf{s} = (\mathbf{q}_{R_1}, \mathbf{q}_{R_2}, \mathbf{p}_B, \mathbf{p}_G)$ と定義し、ロボット R_i による要素動作 A を考える時にはこれを各 Σ_{R_i} での位置情報に変換し \mathbf{s}_{A, R_i} を求める。Fig.8 の状況を \mathbf{s}^1 とし、 \mathbf{s}^1 においてロボット 2 (R_2) がパスし、ロボット 1 (R_1) がシュートすることを規定する個体 $\mathbf{a}_1 = (2, 2P, 1S)$ を例に行動評価の流れを述べる。

まず要素動作 2P の具体化と評価のために \mathbf{s}^1 から \mathbf{s}_{P, R_2}^1 を求め、これに適用可能な **Pass** の行動知能 $\mathcal{A}_P^* = \{\mathbf{a}_P^i\}$ を決定し、その要素より適切な数、例えば 3 つの準最適パスをランダムに選択する (Fig.6 における $\mathbf{a}_P^6, \mathbf{a}_P^8, \mathbf{a}_P^{16}$)。同様に、要素動作 1S の具体化・評価のため **Pass** 実行後の R_1 にとっての状況 \mathbf{s}_{S, R_1} を計算し、これに適用可能な $\mathcal{A}_S^* = \{\mathbf{a}_S^j\}$ の要素より準最適シュートを選択する (Fig.7)。選択されたそれぞれ 3 つの準最適 **Pass** ($\mathbf{a}_P^i, i = 1, 2, 3$) と準最適 **Shoot** ($\mathbf{a}_S^j, j = 1, 2, 3$) を組み合わせた 9 通りの協調 **CShoot** を

$$J'_C(\mathbf{a}_P, \mathbf{a}_S) = c_1 J_{A_N} + c_2 \sum_{k=1}^N t_{M_k} + \frac{c_3}{N} \sum_{k=1}^N J_{A_k} + c_4 N \quad (8)$$

により評価する。各項はそれぞれゴールの是非、行動時間、各要素動作の平均評価値、要素動作数を評価している。 $c_1 \sim c_4$ は適当な重み係数である。

9 通りの評価 $J'_C(\mathbf{a}_P^i, \mathbf{a}_S^j), i = 1, 2, 3, j = 1, 2, 3$ を行

い、評価の最も良い組み合わせ ($\mathbf{a}_P^{i^*}, \mathbf{a}_S^{j^*}$) を \mathbf{a}_1 の具体的行動内容とし、また $\mathbf{a}_P^{i^*}$ を用いた R_2 による **Pass** と $\mathbf{a}_S^{j^*}$ を用いた R_1 による **Shoot** の結果 $J'_C(\mathbf{a}_P^{i^*}, \mathbf{a}_S^{j^*})$ を \mathbf{a}_1 の評価値とする。すなわち

$$J(\mathbf{a}_1) = J'_C(\mathbf{a}_P^{i^*}, \mathbf{a}_S^{j^*}), (i^*, j^*) = \arg \min_{(i, j)} J'_C(\mathbf{a}_P^i, \mathbf{a}_S^j)$$

5.4 協調サッカー行動自律計画アルゴリズム

アルゴリズムをまとめると以下のようになる。

- (1) 状況設定: 状況 $\mathbf{s} = (\{\mathbf{q}_{R_n}\}, \{\mathbf{q}_{E_n}\}, \mathbf{p}_B^0, \mathbf{p}_G)$ を与える。
- (2) 初期個体群生成: 乱数により各個体の要素動作数 N 、動作主ロボット R_k と実行する動作種 $A_k, k = 1, 2, \dots, N$ を決定し、個体群 $\{\mathbf{a}_i\}$ を生成する。
- (3) 評価: \mathbf{s} から各要素動作の状況 \mathbf{s}_{A_k, R_k} を計算し、これに適用可能な動作 A_k の行動知能 $\mathcal{A}_{A_k}^* = \{\mathbf{a}_{A_k}^i\}$ を決定し、5.3 項の方法により協調行動を評価する。
- (4) 切断・接合/突然変異: 切断・接合に用いる個体を交叉率にしたがってランダムに選択し、ランダムに決定した切断箇所以降を入れ替え接合する。また、突然変異率にしたがって突然変異を施す。
- (5) 一定世代 n_G の間最適行動が更新されなければ終了。そうでなければ (3) へ戻る。

5.5 状況に応じた協調行動の自律計画

2 対 1 協調サッカー行動の最適実現問題を考え VGA を適用する。ここでは、Fig.8 の状況を状況 1 とし、これを基準に \mathbf{p}_B^0 のみを 50mm ずつ R_1 の方へ変化させた状況 2~4 を考える。個体数 10、 $n_G = 10$ として VGA を適用したところ、いずれも 10 世代程度で状況に応じた最適行動が求められた。

状況 1~状況 4 に対して得られた最適行動を Fig.14 に示す。図中の×印は \mathbf{p}_B^0 を示している。状況 1,2 では R_2 の方が R_1 よりボールに近いため R_2 が R_1 へパスするセンタリングシュート行動が得られた。状況 3 では R_1 の方が R_2 よりボールに近くなるため R_1 が R_2 へパスするセンタリングシュート行動が得られた。更に、状況 4 になると R_1 がよりボールに近く単独でゴールできる状況で、かつ、評価関数 (8) で要素動作数の少ない方を高く評価しているため単独のシュートが得られた。

6. おわりに

複数ロボットによる協調サッカー問題を例に、要素行動最適化問題を多段階遺伝的アルゴリズムで解いて得た行動知能を保存し、最適動作合成を繰返すことで複雑協調行動を自律的かつ効率的に実現する手法を示した。比較的少数の典型的な状況に対してのみの動作学習にもかかわらず、得られた行動知能は広範な状況に対して適用可能である。しかも、必要に応じて追加学習を行うことでこの適用可能範囲を徐々に拡張できる。ロボットが自

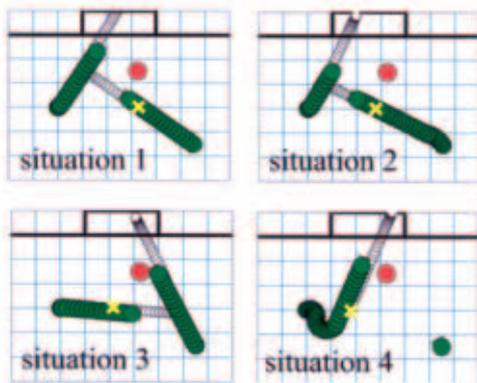


Fig. 14 Cooperative actions planned by VGA

律的に状況に応じて多様な協調行動を計画できるようにするため、可変染色体長遺伝的アルゴリズムにより多様な協調を計画する方法を提案し、MGA と結合することにより適切なロボット協調行動を実現できることを示した。

今後の課題としては、攻守の切換えの判断まで含めたシュートだけを目的とするのではない協調行動への拡張や、ロボットのセンサ能力・記憶容量・通信能力等に応じて目的とする協調行動を高い成功率で実現するのに最適な要素動作の学習スケジュールの検討などが挙げられる。

参考文献

- 1) 中村, 山崎: 反射行動の重ね合わせ理論とその多指ハンドの反射的把握動作への応用, J.RSJ, Vol.15, No.3, pp.136-147, 1997
- 2) J. Morimoto and K. Doya: Acquisition of stand-up behavior by a real robot using hierarchical reinforcement learning, Robotics and Autonomous Systems, Vol.36, pp.37-51, 2001.
- 3) M. Matsushima, T. Hashimoto, F. Miyazaki: Learning to the robot table tennis task -ball control & rally with a human, IEEE Int. Conf. SMC 2003, Vol. 3, pp.2962-2969, 2003
- 4) M. Suzuki: Sub-optimal robot motion for inexperienced situations through Intelligent Composite Motion Control -method of knowledge array-, Int. J. Syst. Science, Vol.30, No.1, pp.87-94, 1999
- 5) M. Suzuki, K.-U. Scholl, R. Dillmann: A method for learning complex and dexterous behaviors through knowledge array network, Proc. IROS2001, Maui, pp.532-538, 2001
- 6) M. Suzuki: Evolutionary Acquisition of Complex Behaviors through Intelligent Composite Motion Control, Proc. 6th IEEE Int. Symposium on Computational Intelligence in Robotics and Automation, Espoo, 2005
- 7) 鈴木, 淵: 多段階遺伝的アルゴリズムによる大規模最適化問題の解法, 計測自動制御学会論文集, Vol.35, No.11, pp.1385-1393, 1999
- 8) M. Suzuki, Y. Hiyama, H. Yamada: An efficient solution for compositional design problems by Multi-stage Genetic Algorithm, Proc. 2007 IEEE Multi-conference on Systems and Control (22nd IEEE International Symposium on Intelligent Control), pp.626-633, Singapore, 2007
- 9) M. Suzuki, Y. Takada: Complex Cooperative Robot Actions based on Recursive Optimal Motion Composition, Proc. SICE Annual Conference 2008, Chofu, 2008
- 10) 高木, 高田, 鈴木: 遺伝的プログラミングを用いたロボットの行動計画, 電気学会産業計測制御研究会資料, IIC-07-11, pp.51-56, 2007