

協働群における行動戦略最適化～競合期待値を最小化する行動戦略決定

Strategy optimization of collaborative agents — Strategy selection to avoid conflict

藤原 一毅 正 倉林 大輔 沖田 邦夫 東京工業大学

Ikki Fujiwara, Daisuke Kurabayashi and Kunio Okita, Tokyo Institute of Technology

This paper introduces a method for efficient collaboration on the model of Multi-agent Traveling Salesmen Problems (m-TSP). Many heuristic algorithms have been proposed to get semi-optimal solution on TSP. The method we introduce is not to tune specific algorithm but to optimize mixed-strategy, the probabilities to choose each algorithm. Suppose that others' mixed-strategies can be acquired by estimation. The method we propose then computes probabilities for each target to be chosen by other agents, and optimize own mixed-strategy to minimize probability to conflict with others. We evaluated the efficiency by simulation in three different ways: (1) all agents have a common algorithm, (2) a half of agents have an algorithm and the other half have another, and (3) all agents have two algorithms and optimize their mixed-strategies by our method. The results showed that the agents performed collaboration most efficiently in case (3).

Keywords — Multi-agent Traveling Salesmen Problems, Collaborative agents.

1 はじめに

災害現場における救助ロボットの活躍が期待されている。現場は他のロボットや人間の救助隊など異種エージェント群の混在環境となるため、互いに明示的な通信を行わず、それでいて適応的かつ協調的に行動する手法が望まれる。本研究は災害救助活動をマルチエージェント巡回セールスマン問題 (mTSP = Multi-agent Traveling Salesman Problem) としてモデル化し、総巡回時間の最小化を目指す。巡回セールスマン問題は NP 困難であり、準最適解を求めるアルゴリズム = 純戦略はこれまでに数多く研究されてきた [4]。本研究は個々の純戦略を追究するのではなく、複数の純戦略を同時に用い、それらの選択確率である混合戦略を最適化しようとするものである。

戦略ベース mTSP ではエージェント同士が同じ目標に集中すること = 競合による効率低下がしばしば問題となる。倉林らは異なる行動戦略を持つエージェント群を混在させることで効率が向上することを示した [1]。また、沖田らは他機の行動を観測し行動戦略を推定する方法を提案している [2]。これらの知見を踏まえ、本研究では他機の行動戦略から目標選択確率を求め、他機と同じ目標へ向かう可能性 = 競合期待値が最小となるよう、自機の戦略を最適化する手法を提案する。

2 問題設定

2.1 環境

災害現場として 2 次元のフィールドを想定する。被災者の存在を示すものとして n 個のオブジェクトがフィールド上に存在する。オブジェクトは座標で表され、大きさは考慮しない。あるエージェ

ントが現在向かっているオブジェクトを特に目標と呼ぶ。

救援者として l 台のエージェントがフィールド上に存在する。エージェントは所定の位置から一斉に出発し、オブジェクトを巡回する。エージェントは座標で表され、衝突回避等は考慮しない。

オブジェクトは最初にいずれかのエージェントが到達した時点で消滅する。被災者の救助にかかる作業時間は考慮しない。オブジェクトが 1 個消滅するごとに n は 1 ずつ減少し、 $n = 0$ となった時点をもって作業完了とする。

エージェント間の情報共有手段として黒板型情報システム [3] の存在を仮定する。これは全エージェントを一元的に管理するのではなく、現場に設置された共有メモリに対し各エージェントが情報の書き込み / 読み出しを行うことで全体として情報の共有を図るものである。エージェントは (1) 残存するオブジェクトの座標、(2) 自機の座標、(3) 他機の座標、(4) 他機の混合戦略を随時取得できるものとする。(4) については後述する。

2.2 純戦略

1 台のエージェントは m 個の純戦略を持つ。純戦略は現在のフィールドの状況 (自機ならびに残存する全オブジェクトの座標) を入力とし、次

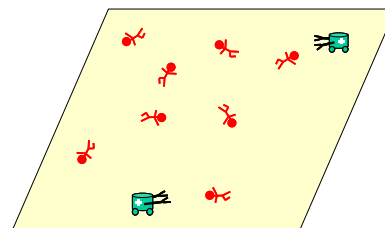


図 1: 想定する作業環境

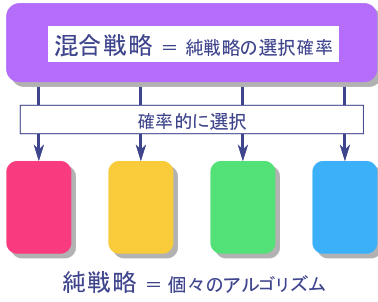


図 2: 純戦略と混合戦略

表 1: 純戦略の分類

	決定的	確率的
局所的	貪欲戦略	近傍優先戦略
大域的	計画戦略	ランダム戦略

に向かうべきオブジェクトの選択確率ベクトル \mathbf{b} ($n \times 1$) を出力するアルゴリズムである。衝突回避等を考慮しない本モデルでは純戦略が最下位の行動アルゴリズムとなる。

$$\left. \begin{array}{l} \text{自機の座標} \\ \text{オブジェクトの座標} \end{array} \right\} \text{純戦略} \rightarrow \mathbf{b}$$

$$\mathbf{b} = [b_1 \dots b_n]^T \quad (1)$$

ただし $|\mathbf{b}| = b_1 + \dots + b_n = 1$ である。 m 個の純戦略によってそれぞれ与えられる \mathbf{b} を並べた行列を \mathbf{B} ($n \times m$) とする。

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_m \end{bmatrix} = \begin{bmatrix} b_{11} & \dots & b_{1m} \\ \vdots & \dots & \vdots \\ b_{n1} & \dots & b_{nm} \end{bmatrix} \quad (2)$$

2.3 混合戦略

エージェントは目標選択に際し確率的に 1 つの純戦略を選ぶ。純戦略の選択確率を混合戦略と呼び、 \mathbf{a} ($m \times 1$) で表す。混合戦略は純戦略より上位の行動アルゴリズムとなる。

$$\mathbf{a} = [a_1 \dots a_m]^T \quad (3)$$

ただし $|\mathbf{a}| = a_1 + \dots + a_m = 1$ である。結局、あるエージェントが各オブジェクトへ向かう確率は

$$\mathbf{p} = \mathbf{B}\mathbf{a} \quad (4)$$

で表される。 \mathbf{p} を目標選択確率ベクトルと呼ぶ。

2.4 本研究で用いる純戦略

本研究では 4 つの純戦略を定義した。

計画戦略は 1 台ですべてのオブジェクトを巡回して元の位置に戻る準最適順路を CCAO 法 [4] により計画し、その順路に沿ってオブジェクトを巡

表 2: パラメータ

フィールドの大きさ	100 × 100 [m]
エージェントの速さ	0.5 [m / step]
オブジェクト数 n	60 [個]
エージェント数 l	2, 4, 10, 24, 60 [台]

回する。エージェントが 1 台のときは最も合理的な戦略である。

貪欲戦略は現在位置から最も近いオブジェクトへ向かう。局所的な合理性に基づく戦略である。

近傍優先戦略は現在位置からの距離の正規分布に従う確率で次の目標を選択する。オブジェクト i の選択確率 b_i は次式によって定める。貪欲戦略とランダム戦略 (後述) の中間的な性質を持つ。

$$b_i = \frac{e^{-10.0d_i/d_{\max}}}{\sum_{i=1}^n b_i} \quad (5)$$

ここで d_i は現在位置からオブジェクト i までの距離、 d_{\max} はフィールドの対角線長さである。

ランダム戦略はすべてのオブジェクトから等確率で次の目標を選択する。

計画戦略と貪欲戦略は次の目標が一意に定まる決定的な純戦略である。一方、近傍優先戦略とランダム戦略はランダム性を持つ確率的な純戦略である。

2.5 戦略構成

本研究では提案手法を含む 3 通りの戦略構成手法について性能評価を行う。

単一手法はすべてのエージェントが同じ 1 つの純戦略を持つ。既存手法である。

混在手法は全エージェントの半数がある 1 つの純戦略を、半数が別の 1 つの純戦略を持ち、フィールド内に混在する。既存手法である。

最適化手法は各エージェントが 2 つの純戦略を持ち、最適化された混合戦略に基づき戦略の切り替えを行う。提案手法である。

2.6 評価関数

台数効率 U [%] はエージェント 1 台あたりの稼働効率を表し、次式で定義する。

$$U = \frac{T_{\text{stg}=i}(1)}{Tl} \times 100 \quad (6)$$

ここで $T_{\text{stg}=i}(1)$ は CCAO 法で得られた準最適順路を 1 台で巡回した場合の所要時間、 l はエージェント数である。 l 台のエージェントを投入したことで所要時間が $1/l$ になれば $U = 100$ [%] であるが、実際には台数が増えるほど低下する。

2.7 シミュレーション条件

シミュレーションにおける諸条件を表 2 に示す。オブジェクトの位置は乱数を用いてランダムに決

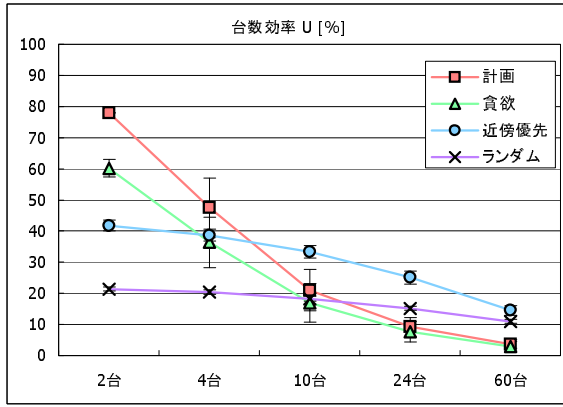


図 3: 単一手法の台数効率

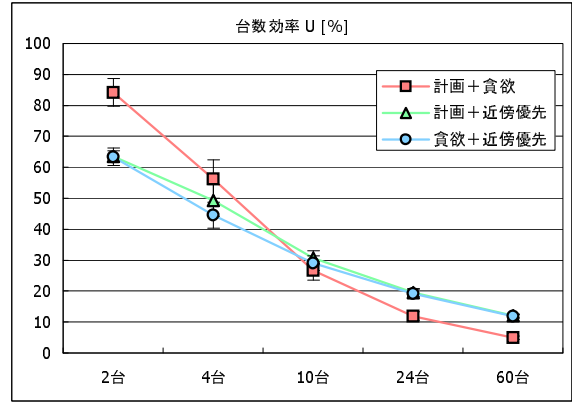


図 4: 混在手法の台数効率

定する。本稿では同一条件に対し乱数の種を変えて 100 回のシミュレーションを行い、それらの平均と標準偏差を示した。エージェントの初期位置はフィールド中央、半径 2.5 [m] の円周上で、エージェント同士は $2\pi/l$ [rad] の等間隔に配置されるものとした。

3 既存手法の評価

はじめに既存手法の性能を評価する。これは後に提案手法の性能を評価する際の比較対象となる。

3.1 単一手法

図 3 は単一手法のシミュレーション結果である。決定的な純戦略は台数が多いとき効率が悪く、確率的な純戦略は逆に台数が少ないとき効率が悪いことがわかる。そこで、台数の変化に対し安定的に良好な性能を得る手法が求められる。

3.2 混在手法

図 4 は 2 つの純戦略を用い混在手法を適用した場合のシミュレーション結果である。貪欲戦略と近傍優先戦略を混在させた場合、台数が増えても競合率が比較的低く、台数効率が極端に悪くならない。決定的な純戦略と確率的な純戦略の混在が競合回避に有効であることがわかる。しかし、混在手法では各エージェントの戦略は作業開始から終了まで変化せず、必ずしも最適な戦略が選択されているわけではない。本研究では戦略選択の動的な最適化を行う。

4 戦略最適化手法

本研究で提案する手法を述べる。エージェントは目標へ到達するたびに以下のアルゴリズムに従い自機の戦略を更新する。

4.1 競合期待値

2 台以上のエージェントが同時に 1 つのオブジェクトへ向かっている状態を競合と呼ぶ。各オブジェクトで競合の発生する確率を求め、その総

和を競合期待値とする。

他機 r の混合戦略 a は推定により得られたものとする。すべてのエージェントおよびオブジェクトの座標は既知であり、各エージェントが持つ純戦略は共通であるから、他機 r の目標選択確率ベクトル $p = B^T a$ がわかる。各オブジェクトにいずれかの他機が向かう確率は、すべての他機の目標選択確率ベクトルの算術平均として得られ、 l 台のエージェント ($r=0$: 自機、 $1 \leq r \leq l-1$: 他機) に対し

$$\bar{p} = [\bar{p}_1 \dots \bar{p}_n] = \frac{1}{l-1} \sum_{r=1}^{l-1} r p \quad (7)$$

で表される。一方、自機が各オブジェクトへ向かう確率は目標選択確率ベクトル p に等しい。したがって競合期待値は $J = \bar{p}^T p$ として得られる。

4.2 線形計画法の適用

競合期待値 J を最小化する混合戦略 a を求める。 $g^T = \bar{p}^T B$ とおくと $J = \bar{p}^T p = \bar{p}^T B a = g^T a$ を得る。 a は確率ベクトルであるから、 $|a| = 1$ かつ $a \geq 0$ が常に成り立つ。 g は既知の係数ベクトルである。したがって、(8) ~ (10) を線形計画問題として解くことにより J の最小値とそれを導く混合戦略 a が得られる。

$$\text{minimize} \quad J = g^T a \quad (8)$$

$$\text{subject to} \quad a \geq 0 \quad (9)$$

$$|a| = a_1 + \dots + a_m = 1 \quad (10)$$

4.3 求解

求解にはシンプレックス法など線形計画法の一般的な解法が適用できるが、この場合はより単純な方法で十分である。

$$J = g^T a = g_1 a_1 + \dots + g_s a_s + \dots + g_m a_m \quad (11)$$

であるから、

$$a_s = \begin{cases} 1 & \text{if } g_s = \min g \\ 0 & \text{else} \end{cases}$$

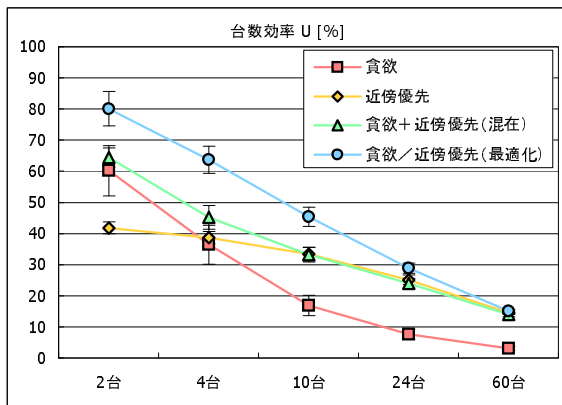


図 5: 単一・混在・最適化手法の比較

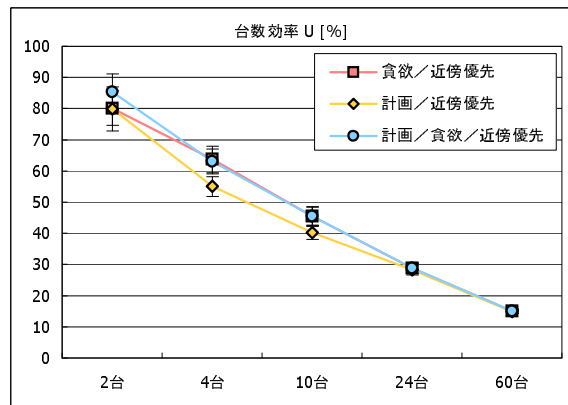


図 7: 2 つまたは 3 つの純戦略からの最適化

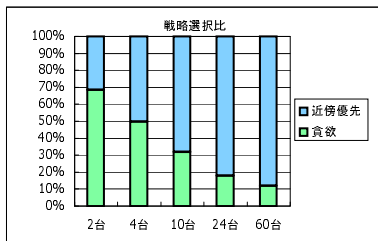


図 6: 最適化手法の戦略選択比率

とすれば、 J は明らかに最小となる。複数の純戦略 s について $g_s = 0$ となり解が一意に定まらない場合は 計画 > 貪欲 > 近傍優先 > ランダム の順に優先とした。

4.4 推定誤差の影響

他機の混合戦略 a は推定によって得られるものとした。自機の戦略選択を決定するのは g の中で最小の要素であるから、推定に伴う誤差が a に含まれていたとしても、それにより g の最小要素が交代しない限り自機の戦略選択は影響を受けない。

5 提案手法の評価

5.1 単一手法・混在手法・最適化手法の比較

図 5 は貪欲戦略と近傍優先戦略を用い、単一手法、混在手法、最適化手法を適用した場合のシミュレーション結果である。最適化手法の台数効率は混在手法に対し 2 台のとき 24 [%]、4 台のとき 41 [%]、10 台のとき 37 [%]、24 台のとき 20 [%]、60 台のとき 8 [%]、それぞれ向上している。戦略選択比率を図 6 に示す。台数が多くなるにつれて確率的な純戦略の選択比率が高くなっている。このことから、提案手法による自律的な最適化の効果が明らかになった。

5.2 すべての純戦略からの最適化

2 つの純戦略を用いるとき、望ましい純戦略の組合せが明らかであるとは限らない。図 7 は計画・貪欲・近傍優先の中から 2 つまたは全部の純戦略

を用い、最適化手法を適用した場合のシミュレーション結果である。全部の純戦略を用いた場合の台数効率は全体として最も良好となっている。このことから、あらかじめ 2 つの純戦略を選ばなくとも、可能なすべての純戦略に対し最適化手法を適用することで自律的に最適化できることが確かめられた。

6 おわりに

本研究では災害救助活動をマルチエージェント巡回セールスマン問題としてモデル化し、競合期待値を最小化する混合戦略最適化手法を提案した。

シミュレーション結果より (1) 単一手法では台数の変化により効率悪化が避けられないこと、(2) 決定的な純戦略と確率的な純戦略の併用が望ましいこと、(3) 提案した混合戦略最適化手法を適用することで既存手法よりも高い効率が見られることが明らかになった。また、提案手法による最適化は望ましい純戦略の組合せが明らかでない場合にも有用であることを確かめた。

参考文献

- [1] 倉林 大輔, 太田 順, 新井 民夫, 野口 克行: 不均質な行動原理を有する移動ロボット群の協調作業, 計測自動制御学会論文集, Vol.35, No.11, pp.1377-1384 (1999).
- [2] 沖田 邦夫, 倉林 大輔: 協働群における行動戦略選択推定, 東京工業大学制御システム工学科卒業論文 (2004), 日本機械学会 ロボティクス・メカトロニクス講演会 (2004) 発表予定.
- [3] Daniel D. Corkill: Blackboard Systems, AI Expert, Vol.6, No.9, pp.40-47 (1991).
- [4] E.L.Lawler, J.K.Lensta, A.H.G.Rinnooy Kan and D.B.Shmoys (Edited): "The TRAVELING SALESMAN PROBLEM — A Guided Tour of Combinatorial Optimization", Wiley (1985).