

マルチブロックエージェントの自律的

1N-11

振る舞いに関する基礎研究

木下正博 北海道大学工学部

川上 敬 北海道工業大学

嘉数侑昇 北海道大学工学部

1. はじめに

完全に自動で柔軟な生産システムあるいは工場を考えると、いくつかの困難な問題を解決する必要がある。最も困難な問題の一つはロボットナビゲーション問題である。現在まで多くのナビゲーションシステムでは、スーパーバイザ等がロボットのスタート地点から目標地点まで、障害物等との干渉を回避しながら移動する最適経路を決定していた。このようなナビゲーションシステムを内部にもつ移動ロボットは自律移動ロボットであると言える。特に複数の自律ロボットがナビゲーション領域に存在する場合、静的な障害物のみならず動的に変化する他のロボットにより形成される障害物の回避、すなわちロボット間のインタラクションを考慮したナビゲーション問題となり、一般にこれを解くことは困難である。今日まで、プロダクションルールやファジーコントロールをベースにした多くのアプローチが存在するが、これらはすべて問題そのもののルールや方法論を採用しているため自律的な行動獲得には不向きであった。このような背景から、複数のロボットが与えられ、複雑なナビゲーション領域が設定された場合、自律的にロボットのナビゲーション問題を解くことができる新たな学習機構が求められている。

このように、ナビゲーションシステムは自ら最適解を獲得することが可能でなければならない。このような学習システムのために多くはニューラルネットワークのような情報処理モデルが使用されるが、本研究では例からの効果的な学習技術として提案されているクラシファイアシステムを用いて、マルチエージェント環境のナビゲーション問題におけるエージェントの自律的振る舞いについて考察する。ここではいわゆるプロダクションライクなルールであるクラシファイア、いわゆるシステムの戦略を内部にもつ自律エージェントを用意し、任意の環境内において自律的柔軟的に移動ロボットのナビゲーション問題を解くことを試みる。ここでは複数のロボットに与えられたタスクとして、コンテナのような入れ物から外に出ることを想定し、それぞれの自律ロボットをブロック形状をした自律ブロックエージェントと呼ぶ。

本論文では、提案する方法論を基本として自律ブロックエージェントナビゲーションシミュレータを構成し、いくつかの数値実験と視覚的な実験を行った。これらの実験において複数のブロックエージェントは同じ環境内でその自律ナビゲーションシステムに従って同時に行動する。これらの実験結果からクラシファイアシステムをもつ自律ブロックエージェントの学習効果について議論する。

Autonomous Behavior of the Multiple Block Agents
Masahiro Kinoshita, Takashi Kawakami* and Yukinori Kakazu
Autonomous Engineering, Complex System Engineering
Faculty of Engineering, Hokkaido University
N-13, W-8, 060-8628, SAPPORO, JAPAN
kino@complex.eng.hokudai.ac.jp

*Dept. of Industrial Eng., Hokkaido Institute of Technology
7-15 Maeda, Teine-ku, 006-8585, SAPPORO, JAPAN
kawakami@hit.ac.jp

2. マルチエージェント問題としてのブロック取り出し問題

ここでは、取り出すあるいは抜け出る行為をブロックエージェントがなすタスクの一つとする。少なくとも一つの出口をもつコンテナ、コンテナ内にいくつかのブロックが与えられたとき、扱う問題はこれらブロック群の一部をコンテナから取り出すことである。幾何的な設定としてこれらエージェントおよび環境は2次元X-Y絶対座標系で表現する。ここで、ブロックエージェントの振る舞いをマルチエージェントナビゲーション問題として議論する。マルチエージェントナビゲーション問題(MANP)はブロックエージェント(BA)と作業環境(WE)から構成される。

$$MANP = (BA, WE) \quad (1)$$

$$BA = (M, S, A, D) \quad (2)$$

ここで、 M はブロックエージェントのモーフロジー、 S はセンシング情報、 A は行動パターン、 D は行動決定器である。

$$M = (m_j^1, m_j^2, m_j^3, \dots, m_j^n, j=1 \dots n_b) \quad (3)$$

m_j^1 : BAの形体(たとえば長方形)

$m_j^2, m_j^3, m_j^4, m_j^5$: 幾何的屬性(寸法, 重心等)

m_j^6, m_j^7, \dots : センサー屬性(赤外線, 接触等)

M はいくつかのモーフロジカルな項目の集合として表わされる。 BA の振る舞いの空間的制約はその形からの影響を受けるため BA にとって幾何的屬性は重要で、 BA のインタラクションを考慮した場合、 BA の動作の複雑性はその形状に起因している。

$$S = (s_1, s_2, s_3, \dots, s_k) \quad (4)$$

S はセンシング属性の集合でたとえば次のように設定することができる。

S_1 : ブロック重心がゴール位置にあるかどうか。

S_2 : ブロック重心とゴール間との距離。

S_3 : BA が他の BA あるいはコンテナの壁と干渉したかどうか。

S_4 : BA の位置と姿勢。

S_5 : 現在時刻。

$$A = (a_1, a_2, a_3, \dots, a_l) \quad (5)$$

ここで、 A は BA が1タイムステップにとることのできる行動、あるいはあるステップ幅の移動間にとることのできる行動の要素の集合とする。たとえば、

a_1 : BA を座標軸に直交して移動。

a_2 : BA の姿勢を変更する。

a_3 : 他のエージェントを飛び越す。

$$F: S \rightarrow A \quad (6)$$

関数 F はセンシング情報から行動へのマッピング関数であ

る。

$$D = (d_1, d_2, d_3, \dots, d_p) \quad (7)$$

D はセンサーアクトマッピング関数 F の変換機構の集合とする。たとえば、

d_1 : クラシファイアースystem

d_2 : ニューラルネットワーク

d_3 : 遺伝的アルゴリズム

また、作業環境を以下のように表現する。

$$WE = (DC, R) \quad (8)$$

$$DC = (me_1, me_2, me_3, \dots, me_q) \quad (9)$$

ここで、 WE は BA を含む作業環境、 DC はコンテナ、 R は環境からの報酬である。

BA は DC の制約を満たす間各タイムステップにおいて行動することができる。 BA が DC から取り出されるとき、 BA と DC の関係が重要となる。ナビゲーションにおけるコンテナ内部で出口付近では、出口とブロックの寸法、位置、姿勢などの空間的制約を考慮して干渉のない経路を決定しなければならない。

$$\exists BA: \text{both width and height} > \text{width of the exit} \quad (10)$$

コンテナ内部には幅と高さが共に出口の幅より大きい BA が存在すると仮定する。これらの種類のブロックは他の小さなブロックの経路を塞ぐことになり、このブロックの振る舞いが全体のタスクの完了に大きく影響する。すなわち、他のエージェントへ経路を開けるような社会的振る舞いが要求される。ブロックエージェントの自律的行動においては、エージェント間の移動順序に関する優先順位や移動方向などの事前の情報は存在しない。ここで、ブロック取り出し問題を下記のように定式化する。

$$*P = \operatorname{argmax} \exists BA \notin DC \quad (11)$$

* P は DC に含まれない BA の個数を最大化する問題であり、この種の問題は空間的制約を含む大域的最適化問題となる。最適経路の学習過程において、すべての自律的 BA は同時にコンテナ出口に集まろうとする。いたるところでデッドロックが発生するため小さな BA であっても取り出すことが困難となる。そこでこの問題において* P を増加させるアプローチを提案する。

3. クラシファイアースystemのインプリメント

ここで行うシミュレーションでは自律的ブロックエージェントは適当な経路を見つけるための簡単な機構をもつ。以下の知覚機構を持たせる。

- ・ 現在位置の座標, bp_i^t .
- ・ コンテナの出口を通りぬけられるか否か, bl_i と w
- ・ 目的地への到着. この情報は次の関数によって判定される.

$$f(bp_i^t) = \begin{cases} 1 & : \text{if } bp_i^t = dp_i \\ 0 & : \text{otherwise} \end{cases} \quad (12)$$

クラシファイアのストレングス部の更新を以下のように設定し bp_i^t の状況により環境からの報酬を変化させる。すなわち、強化学習のためのプランとして正の報酬、負の報酬を設定する。

$$P1) \text{if}(bl_i^{t+1} \text{ contacts with other blocks}) \text{ then } R < 0$$

$$P2) \text{if}(bp_i^{t+1} = bp_i^t \exists \tau \leq t \in T) \text{ then } R < 0$$

$$P3) \text{if}(d(bl_i^{t+1}) \geq d(bl_i^t)) \text{ then } R < 0$$

$$P4) \text{if}(bp_i^{t+1} = dp_i) \text{ then } R > 0$$

dp_i は目標地点の座標で関数 d はブロック重心と目標地点との距離を表わす。

4. 計算機実験

クラシファイアを内部にもつエージェントの取り出し問題における振る舞いを実験により解析した。学習の初期段階では効率(ブロック数)が増加するがしだいにデッドロックが発生し効率が減少する。しかし、しだいにデッドロックを解消し、効率を上げる現象がみられた(図2)。また、経路を妨げるブロックエージェントが存在する場合、経路を譲る振る舞いがみられた(図3)。

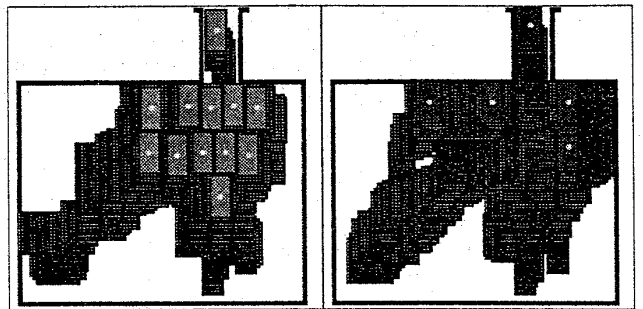


図1 ブロックエージェントの移動経路

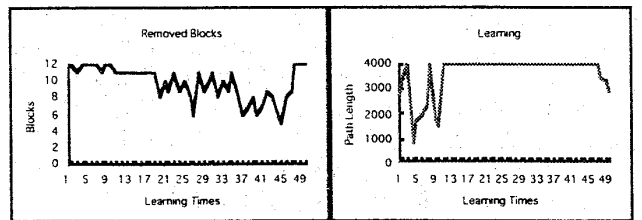


図2 学習曲線

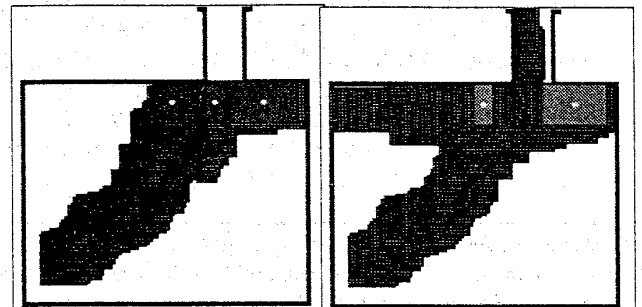


図3 譲り合いの振る舞い

5. おわりに

マルチエージェント系におけるナビゲーションに対し定式化を行い整理し、自律ブロックエージェントの振る舞いをクラシファイアースystemによるナビゲーションシステムによって考察し、解の獲得への可能性を示した。