

## マルチエージェントにおける情報交換ルールの自動獲得に関する一考察

4 J-2

鈴木佐俊 鈴木誠 平澤茂一

早稲田大学理工学部経営システム工学科

suzuki@hirasa.management.waseda.ac.jp

## 1 はじめに

複数の自律したエージェントの相互作用としてモデル化されたマルチエージェントシステムでは、エージェント間の協調作業、分担作業が重要である。エージェントに協調作業や分担作業を行わせるアプローチとしては集中管理方式のようなトップダウン型とエージェント間通信によるボトムアップ型が考えられる。負荷分散、耐故障性、拡張性という点では後者が優れている[2]。また、エージェントシステムの形態はエージェントが共通のルールで動作する均一系と個々のエージェントが別々のルールで動作する非均一系に分類することができる。本研究ではボトムアップ型の均一のエージェント系を扱う。

従来研究[1]では、環状に接続された5エージェントの情報交換のルールが強化学習によって獲得できることと、学習時にノイズが混入した場合においても誤り訂正能力を持つ情報交換ルールが獲得できることが示された。しかし、エージェント数が5個に固定されていたために、問題の難易度と通信量の関係は明確にされていなかった。

本研究では、従来研究で提案されたモデルを利用して情報交換ルールを強化学習により自動獲得する際のエージェント数と必要な通信量の関係をシミュレーションにより調べる。その結果、エージェント数が増加するに従い、当然通信量も増大し、ベンチマークとして選んだ問題ではルールを

獲得するための通信量はエージェント数に対し対数的に増大することを示す。

## 2 問題設定

## 2.1 エージェント

仮想問題としてタスクの割り当て問題を考える[1]。エージェント数と同じだけの問題が用意されている。エージェントは通信の後、問題を選択する。このとき、それぞれのエージェントが重複なく問題を選ぶようにしたい。ここで本研究で扱うエージェントには以下の仮定をおく。

- (1) 全てのエージェントは同じルール、同じアルゴリズムで動作する。
- (2) エージェントは環状に接続されている。
- (3) 全てのエージェントは自分に接続されたエージェント（左右のエージェント）と通信 ( $n$  ビット) できる。

エージェントは自分の発言のほかに左右に接続されているエージェントの発言 ( $3n$  ビット) を受信することで局所的情報を得ることができる。全エージェントの発言をならべたもの  $s \in \{0,1\}^{3n}$  を状態と呼ぶ。各エージェントは左のエージェントの発言、自分の発言、右のエージェントの発言を元に共通の「発言ルール」にしたがって情報交換し次の発言 ( $n$  ビット) を決める。これによって状態が遷移する。エージェントは「選択ルール」を用いて担当するタスクを選び出す。このとき異なるタスクを選択するためには全てのエージェントの入力が異なるような状態になっている必要がある。逆にすべてのエージェントが同じ入力となる状態のとき、すべてのエージェントが同じ発言ルール、同じ選択ルールを使用するので、タスクの割り当ては失敗する。

## 2.2 学習の方法

学習の方法として強化学習を使用する。実際に問題の割り当てを行い、第  $i$  の試行でエージェントによって選ばれた問題の数  $k_i$  を報酬としてフィードバックする。報酬は学習時に選ばれた発言ルールと選択ルールとに配分される。(例えば  $a=5$  エージェントが 1,2,5,3,3 と選択したら  $k_i=4$  点) エージェントは配分された得点に基づき、ルーレット選択を行う。

## 3 シミュレーション

この問題設定で難易度を決定するのはエージェント数  $a$ 、エージェントの発言の符号長  $n$  である。ここで、エージェント数  $a$  と符号長  $n$  を変化させることにより、報酬にどのような影響を及ぼすのかをシミュレーションによって示す。

### 3.1 シミュレーションの条件

1 回の通信を行い、発言ルールを適用後、再度通信を行う。選択ルールを用いて各エージェントにタスクの選択を行わせ、得点に従い報酬を配分する。ここで得点と報酬は一致する。この試行を繰り返すと平均点  $\hat{k}$  は収束し、学習は終了する。

### 3.2 シミュレーションの結果

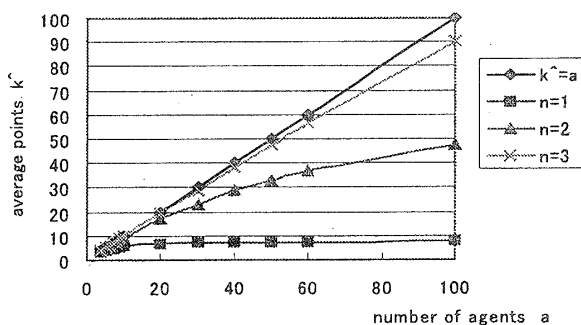


図1: エージェント数と平均得点

図1はエージェント数  $a$  (=タスクの数) と符号長  $n$ 、平均得点  $\hat{k}$  の関係を示したものである。なお、今回は通信路にノイズは混入しないモデルを扱った。

図1で、タスクの分割を行うためには、 $k=a$  に十分接近している必要がある。しかしすべてのエージェントの入力が同じ状態になり、システムが膠着状態に陥り、タスクの割り当てに失敗する場合

があるので  $\hat{k}$  と  $a$  は完全には一致しない。

## 4 考察

実験の結果、以下のことが明らかになった。符号長を固定した場合、正しくタスクの選択が行われるエージェント数には限界がある。また  $\hat{k} \leq a \leq 2^{3n}$  である。これは選択ルールの入力パターン数 ( $\|s\| = 2^{3n}$ ) が出力パターン数 ( $a$ ) を上回っていないなければならないという制限によるものである。しかし、図2で示されるように実際はも

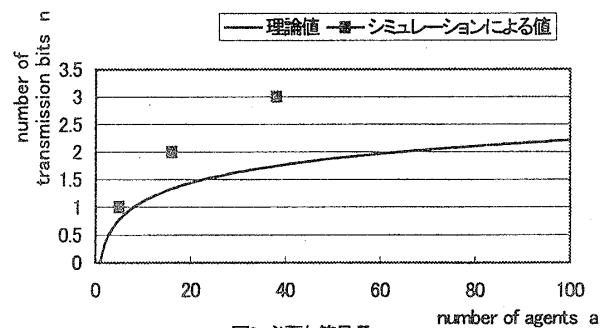


図2: 必要な符号長

っと早く限界がくる。

## 5 まとめ

本研究では情報交換ルールを強化学習により自動獲得する際のエージェント数と必要な通信量の関係をシミュレーションにより調べた。その結果、エージェント数  $a$  が増大してもそれに必要な発言の符号長  $n$  はエージェント数の対数のオーダーで与えられることが分かった。

今後の課題としてより現実の問題に即したモデルを構築する必要がある。情報理論的な定式化を行い、通信容量やレートをノイズの影響を含めて考慮したい。また、ルールを保持するテーブルの大きさを小さくするような工夫も必要である。

## 参考文献

- [1] 宇田川建文, 村川正宏, 平岡和幸, 森永聡, 吉沢修治, "分担作業における情報交換ルールの強化学習による自動獲得", 信学技報, AI99-9, pp.61-68.
- [2] 荒井幸代, 宮崎和光, 小林重信, "マルチエージェント強化学習の方法論 - Q-Learning と Profit Sharing による接近 -", 人工知能学会誌, Vol.13, No.4, pp.609-618.