

エージェント間提携形成に基づく協調追跡行動の進化的学習

4G-12

佐藤 秀樹

渥美 雅保

創価大学工学研究科情報システム学専攻

創価大学工学部情報システム学科

1 はじめに

複数のタスクをエージェントが協力して解決する場合、それぞれのタスクをあるエージェントのグループが合理的に解決することが要求される。この要求に対して、協力ゲームの理論 [1]は、あるタスクをエージェント間提携のもとで解決する場合のエージェント間の利得配分の合理的基準を与えている。

本論文では、複数のハンターエージェントが複数の獲物を追跡する拡張追跡タスクを例として、どの追跡タスクについて誰といかなる利得配分で提携を形成するかという問題と、提携のもとで追跡行動を選択する問題とを、遺伝的プログラミング[2]に基づき並行的に学習するアーキテクチャを構築し、シミュレーション実験を通じて本アーキテクチャのもとでのエージェントの振る舞いについて考察を加える。

2 拡張追跡タスク

拡張追跡タスクは、追跡タスク[3]を次の2点で拡張したものである。

- 1)無限に広がる2次元格子上に、2つの獲物エージェント（獲物）と6つのハンターエージェント（ハンター）がいる。
- 2)あるハンターがある獲物を捕らえたときに得られる利得は、ハンターによって異なり、かつ捕らえる獲物によっても異なる。

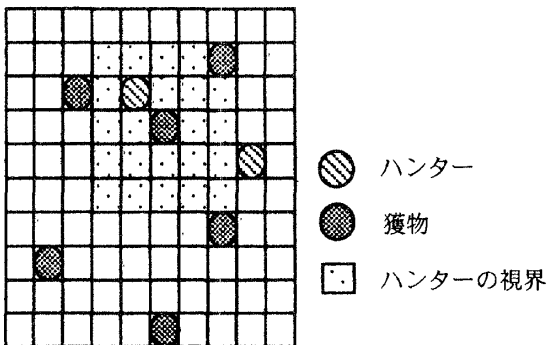


図1 拡張追跡タスク

各ハンターは一度に1グリッドずつ移動し、動かずに留まることも可能である。各グリッドには、同時

Evolutionary Learning of Cooperative Pursuit Behavior based on Coalition Formation

Hideki Satoh, Masayasu Atsumi Soka University
1-236 Tangicho, Hachiojishi, Tokyo 192, Japan

に複数のエージェントが入ることはできない。また、各ハンターには視界が等しく与えられており、獲物の4近傍を囲むことで獲物を獲得できる。また、獲物は常にランダムに行動する(図1)。

拡張追跡タスクの目的は、ハンターが協調して獲物を捕らえることによりできるだけ多くの利得を獲得することであり、そのためには、エージェント間提携の形成と追跡行動との学習が各エージェントに要求される。

3 提携形成に基づく協調追跡アーキテクチャ

3.1 提携

提携とは、1つの協力単位として行動することが可能なグループであり、形式的にはプレイヤー全体の集合の部分集合をいう。即ちn人のプレイヤーの集合 $N=\{1,2,\dots,n\}$ に対して、その部分集合 $S(\subseteq N)$ を提携と呼ぶ。

3.2 交渉集合型特性関数

提携形成においては、各エージェントがより多くの利得を得られる合理的な提携を形成することが重要である。ここで、合理的であるとは、その提携のもとでの各エージェントへの利得配分が、各エージェントにとって同意できる個人合理的利得構成であることをいう。このような個人合理的利得構成はM安定であるといわれる。このとき、各提携の提携値は、交渉集合型特性関数により決定される。

拡張追跡タスクにおける交渉集合型特性関数を次のように定式化する。ハンターの集合をN、許容提携を $B(\subseteq N)$ 、提携Bの持つ共同戦略の集合を $S_B=\{S_{p_1}, S_{p_2}\}$ （ただし、 p_1, p_2 は獲物で S_{p_i} はBのメンバーが獲物 p_i を追跡することを表す）、提携Bに属するプレイヤー h_i の利得関数を $f_{h_i}; S_B \rightarrow \mathbb{R}$ （実数）、 $f_{h_i}(S_{p_1})=v_{i1} \in \mathbb{R}, f_{h_i}(S_{p_2})=v_{i2} \in \mathbb{R}$ とする。このとき、交渉集合型特性関数を図2のように定義する。

$$v(B) = \max_{\substack{S_p \in S_B \\ p_j \in B}} \sum_{h_i \in B} f_{h_i}(S_{p_j}) = \begin{cases} 0 & (|B| \leq 3) \\ \max_{h_i \in B} \{ \sum_{h_i \in B} f_{h_i}(S_{p_1}), \sum_{h_i \in B} f_{h_i}(S_{p_2}) \} & (|B|=4) \\ \max_{B' \subseteq B, |B'|=4} \max_{h_i \in B'} \{ \sum_{h_i \in B'} f_{h_i}(S_{p_1}), \sum_{h_i \in B'} f_{h_i}(S_{p_2}) \} & (|B| \geq 5) \end{cases}$$

図2 交渉集合型特性関数

各ハンターエージェントは、この特性関数のもとで、できるだけ大きな提携値を与えてくれる提携を形成しM安定な利得配分を達成すべく交渉を行う。

3.3 提携公募アーキテクチャ

各エージェントは、交渉集合型特性関数と提携下での利得配分を学習することにより、合理的な提携形成をすることができる。しかし、このように複雑な関数並びにM安定な利得配分を直接学習することは非常に困難である。よって、本研究では、提携形成の学習を行うための提携公募アーキテクチャ(図3)を提案し、導入する。

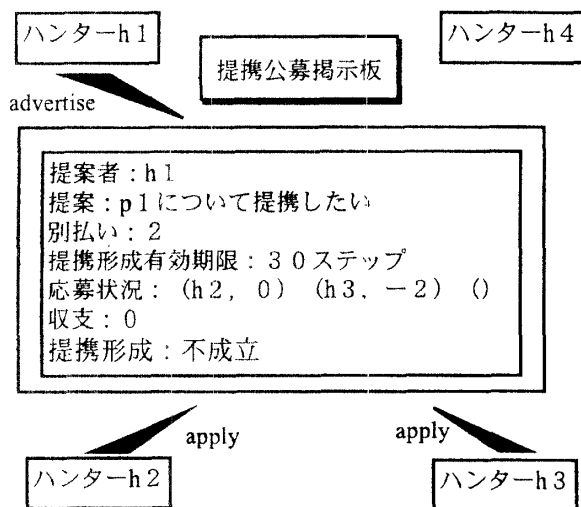


図3 提携公募アーキテクチャ

提携公募アーキテクチャにおいては、ハンターは捕獲したい獲物に対する提携形成を公募し、その公募に対して他のハンターが応募することにより提携形成を行う。公募や応募をするときには、自らが支払ったり、もしくは受け取りたい別払い額を同時に提示する。応募者が4人になり、かつ別払い額の収支がプラスであるとき、それら4人のハンターエージェント間に提携が成立する。

3.4 エージェントアーキテクチャ

追跡行動と提携形成を並行的に学習するエージェントのアーキテクチャを図4に示す。各エージェントは、提携形成のためのGP木(提携形成木)と追跡行動のためのGP木(追跡行動木)を持つ。

提携形成木は、環境状態と提携公募掲示板の公募・応募状況を観測し、ある獲物の捕獲に関する提携の公募または応募を提携公募掲示板に対して行う木である。提携形成木の進化は[4]と類似の方法によりSTGP(Strongly-Typed GP)により実行される。一方、追跡行動木は、環境状態の観測と提携公募掲示板に記されている提携が成立しているかどうかの情報に基づき、東・西・南・北のいずれかに移動する

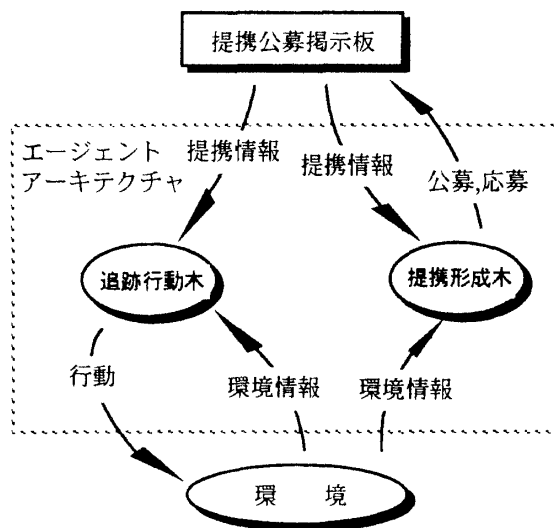


図4 エージェントアーキテクチャ

かその場にとどまるかの選択を行う木である。

4 考察

シミュレーション実験により、エージェント間に提携が形成され、その提携のもとで追跡行動が発現することを確認している。また、本アーキテクチャのもとでの追跡が追跡行動木だけで追跡タスクを遂行する場合に比べてよい結果を生むことも確かめている。しかし、エージェント間に形成されている提携が合理的なものであるかについては、さらにシミュレーションを重ねて検討をしていくことが必要である。

5 むすび

本論文では、拡張追跡タスクを例として、エージェント間での提携行動を交渉集合型特性関数を用いて特徴づけた。そして、合理的提携形成を学習するための提携公募アーキテクチャと、追跡行動と提携形成を並行的に学習するエージェントアーキテクチャを提案した。今後、シミュレーションを重ねて、本アーキテクチャの有効性を実験的に明らかにしていく。

参考文献

[1] 鈴木光男: "新ゲーム理論", 勁草書房 (1995)
 [2] Koza, J: Genetic Programming on the Programming of Computers by Natural Selection, MIT press(1994)
 [3] Benda, M., et al.: On Optimal Cooperation of Knowledge Sources, technical Report BCS-G2010-28, Boeing AI Center(1985)
 [4] Dworman, G, et al.: Bargaining by Artificial Agents in Two Coalition Games: A Study in Genetic Programming for Electronic Commerce, Proc. 1th Conf. on GP(1996)