

多様な協調行動獲得のための段階的行動決定

* 森田晋作

奈良工業高等専門学校 電子情報工学専攻科

〒639-1080 奈良県大和郡山市 矢田町22番地

morita@info.nara-k.ac.jp

Abstract: 本研究は, マルチエージェント環境において, 状況によって一意に決まる多様な協調行動をエージェントに獲得させることが最終的な目的である. これにより, 目的を柔軟に達成していくエージェントを実現することを目指す. 実験には, 近年マルチエージェント環境の具体例として注目されている RoboCupSoccerSimulation で用いられる SoccerServer 上で動作するエージェントを扱う.

1. はじめに

現実世界のような常に变化する環境において, 集団で目的を達成しようと行動する場合, どのような方法が最適であるかを特定するのは難しい. 人工知能分野では, こうした集団行動について, マルチエージェント[1] 環境での問題として扱う.

マルチエージェントで重要視されるのはエージェント間の協調行動である. 近年では RoboCupSoccerSimulation[2] というソフトウェアロボットチームによる競技大会が, マルチエージェントの環境として注目されている. RoboCupSoccerSimulation では様々な制限があるなか, いかんにしてエージェントに協調行動を最適に獲得させるかの研究が進められてきた.

本研究ではエージェントに最適な協調行動を獲得させるのではなく, 様々な動きを見せる協調行動の獲得を目指している. RoboCupSoccerSimulation において有効な方法はプランニング[3] であるが, アクションパターンを設計者が考えて与えなければならないため, パターンの多様性は設計者に依存し, また, その作業量も膨大になると予測される. この問題に対し, 2階層のモデルで段階的に行動を獲得させて解決を試みる. 上位階層において協調行動を示す行動の組合せを学習させ, 下位階層において各行動の適切なパラメータの付与を行い, 人間によるアクションパターンの設定をせず, 多様な協調行動を実現する. これらの詳細を次に述べる. なお, 実際に SoccerServer 上でこの様な段階的行動決定により獲得した行動を実行するエージェントはまだ完成していない.

2. 研究方法

本研究で用いる段階的な行動決定での2階層モデルは, 上位階層を抽象化レベル, 下位階層を連続値レベルと呼ぶ. 2階層モデルイメージを図1に示す. 抽象化レベルでは, サッカープレーを抽象化し, なるべく小さな状態空間の中でどのような協調行動パターンがありうるかを調べる. そして, 連続値レベルでは, 抽象化レベルで得られた行動の組合せを実際に SoccerServer 上で実行する.

2.1 抽象化レベル

このレベルでは, kick, dash など SoccerServer で用意されているアクションを, どう組合せれば協調行動になるのか, とした各プレイヤーがとる行動の組合せを各プレイヤーに学習させる.

学習は完全観測MDP(Markov Decision Processes) を用いた強化学習により行う. MDPとは状態, 行動の組合せに対し, 遷移する状態の確率を表したものであり, 完全観測MDPとはフィールド内のプレイヤーの状態遷移が全て把握できるMDPを指す. MDPの状態, 行動はサッカープレーを抽象化したものであり, 状態を全プレイヤーとボールの位置関係で表し, 行動を全プレイヤーの行動の組合せで表す. 例えば図2のように, ボールとプレイヤー1が同じ位置にいる状態を P1 keep と表現し, 行動は (Player1's action, Player2's action) と表す.

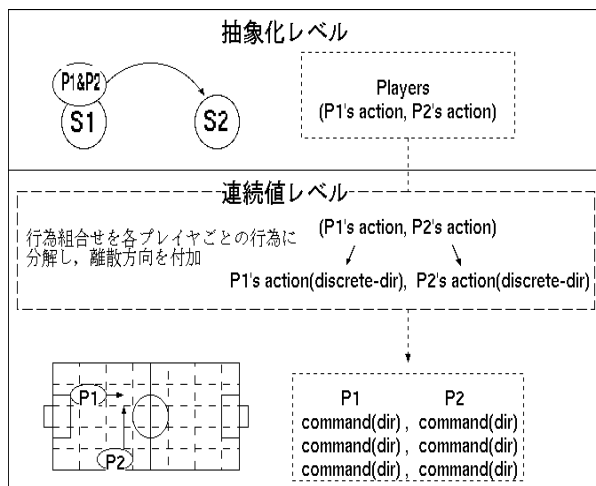


図1 2階層モデル

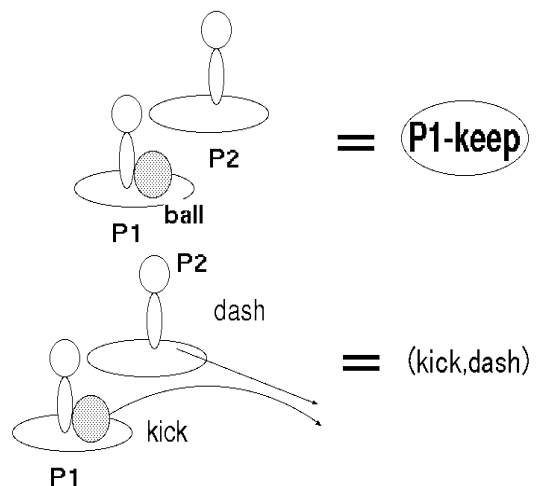


図2 抽象化レベルでの行動と状態の表現

2.2 連続値レベル

連続値レベルでは、フィールドをグリッド状に分割し、プレイヤーがどのセルに居るかを区別できるようにする。また、連続値レベルではアクションのパラメータを決定する機能が必要である。そこで探索空間を減らすために連続値のパラメータではなく離散パラメータを導入する。本研究では方向のパラメータは360°を8方向程度に分割する。SoccerServerが直接実行可能な命令群を組合せ、抽象的で人間にとって理解が容易なアクションを設計する。

3. おわりに

本研究は多様な協調行動の実現にむけ、SoccerServer上のエージェントに段階的な行動決定により協調行動を学習させることを試みるものである。これまでの研究で我々は、Player1とPlayer2が同時にある方向にボールを蹴るまたは走る、というようにある瞬間に共起する協調行動を扱ってきた。しかし、サッカーのように協調行動が連続することでプレーがなりたつ環境では、協調行動を行うにいたるまでの各行動も重要である。

ある瞬間に共起する協調を単純な協調とし、連続する行動系列による協調行動では単調な協調が複数回起き、それらの協調行動は依存関係を持つと考える。単純な協調がn回起きる行動系列による協調をn重協調と定義し、今後、本研究では連続する依存関係を持つ行動列によるn重協調行動の獲得を目指す。これを達成するために、先に述べた2階層モデルによる段階的行動決定によるエージェントの同時共起的協調を実現する。

4. 参考文献・URL

- [1] 山口智浩, 堀口正一: マルチエージェント強化学習における強化信号の役割. 計測自動制御学会 システム・情報部門学術講演会, 2002
- [2] RoboCup日本委員会公式ホームページ, <http://www.symbio.jst.go.jp/~yuki/RoboCupJNC/official.html>
- [3] 木下修平, 山本喜一: エージェントの戦略的プランニング機構, システム生命を有する知的システムの構築, 1997