

# RoboCup サッカーエージェントに対する強化学習を用いた 行動選択手法の提案

## Proposal of Action Selection Using Reinforcement Learning in RoboCup Soccer Agents

笹岡 久行† 村木 俊介†  
Hisayuki Sasaoka Shunsuke Muraki

### 1. はじめに

機械学習の一手法である強化学習[1]はマルチエージェントシステム(MAS)における追跡問題の解決等に適用され、その有効性が確認されている[2,3,4].

MAS に関する研究テーマの 1 つとして RoboCup サッカーシミュレーションリーグが提案されている[5,6]. これは人工知能の研究の観点から、実世界の問題を多く兼ね備えた良いテストベットとなっている。なぜなら、ここでは仮定の広大なサッカーフィールド上に 22 体の自律エージェント(自チーム 11 体 vs. 相手チーム 11 体)とボールが存在するという複雑な状況に対して、実時間で対応が必要であるためである。また、エージェント同士のチームプレイやコミュニケーションが存在し、さらに視覚情報や聴覚情報の欠落などの不確実な要素を処理しなくてはならない。先行研究では、エージェントのドリブル等のような基本技術と呼ばれているものを獲得するため強化学習等が利用され、その有効性が報告されている[6]. 一方、チーム戦略の獲得を目的として進化型計算を適用し、1 試合当たりの得点に対する評価から有効な戦略を得ることができたことを報告している[8]. また、縮小されたフィールドにおける少人数でのミニゲームにおいて、フィールドの状態に対する条件付き確率を利用し、エージェント毎の基本技術やチームプレイの獲得に対する有効性が報告されている[9].

本研究では、RoboCup サッカーシミュレーションリーグにおけるサッカーエージェントにおいて、戦術の選択のために強化学習を用いる。既に、エージェント単体における強化学習等を用いた基本技術の獲得手法について、その有効性は確認されている。しかし、「ポストプレー」や「サイド攻撃」の様なさらにもう一段高度なサッカーの戦術[9]の獲得手法が必要である。だが、文献[4]あるいは[8]において述べられているように、機械学習は適用するタスクによっては、探索領域の広さあるいは学習結果が収束するまでの試行回数多さ等が大きな問題となる。ここで、これらの問題の解決を目指し、人手で作成した静的な規則を組み込んだ強化学習の手法を提案する。ここで、人手で作成した静的な規則とは人間が行うサッカーの試合あるいはサッカーエージェント同士の試合のログ等に基づいて、予め人手により抽出した静的な行動規則のことである。

### 2. 基本的な考え方

人間は優れた学習能力を有し、それを利用して、例えばサッカーでは人間の選手は素晴らしいプレーを実現してい

ると考えている。我々の研究の最終的な目標は、人間が有する学習機能能力の解明とその工学的な応用にある。

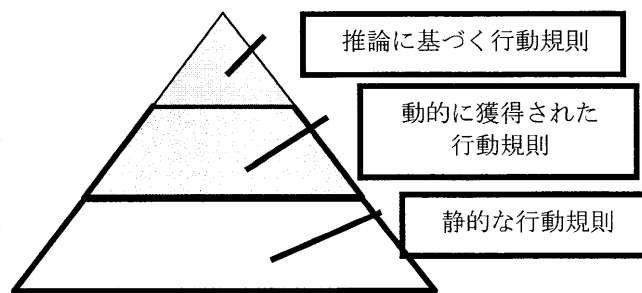


図 1: 行動規則の分類

図 1 に示すように、「推論された行動規則」、「動的に獲得された行動規則」、「静的な行動規則」のように行動規則を分類している。このように考えたのは、日常生活の様々な場面で、種々の方法により、人間は知識を獲得していると内省されるからである。例えば、サッカーに関する知識では、サッカーの入門書や指導者からプレーのやり方を直接的な形で学んだり、あるいは、プロのサッカー選手がプレーする試合を観戦しそれを模倣することから始め、プレーのコツをつかんだりすると思われる。つまり、図 1 における最下層の「静的な行動規則」とは直接的にサッカーエージェントに組み込まれる行動規則であり、人間が本などから直接的に獲得する知識に相当する。また、「動的に獲得された行動規則」とはサッカーエージェント同士の試合のログから帰納的学習を用いて抽出される行動規則である[10]. さらに、「推論に基づく行動規則」とは、上記 2 つの行動規則において対応が困難なケースにおいて利用されるものであり、その獲得方法に関しては今後検討する予定である。

### 3. 得点に至る経緯に対する分析

攻撃についての静的な規則を考えるために、本研究ではエージェント同士の試合における得点場面を分析した。その対象として、2004 年 5 月に大阪において開催された「RoboCup Japan Open 2004」における RoboCup サッカーシミュレーションリーグの試合のログ[11]を利用して頂いた。この大会では、1 次予選 50 試合、2 次予選 30 試合、決勝トーナメント (3 位決定戦を含む) 8 試合が行われた。この中で、1 次予選には実力差が大きいチーム同士の試合があり、一方のチームが 10 点以上の得失点差を付けて勝利した試合が 15 試合 (30.0%) もあった。これ

† 旭川工業高等専門学校専攻科 生産システム工学専攻

らの試合ではエージェントの基本技術の差が大きいと考えられ、静的な行動規則を取得するのが困難であると考えられる。さらに、2次リーグの試合において、リーグ戦であるため、得点が無い0対0の引き分けの試合が5試合あった。そこで、これ以外の試合の中から20試合(全59得点)を選択し、分析を行った。

表1: 得点へと至った経緯に基づく分類

得点への経緯	回数
(a)サイドの選手からのセンタリング	26(44.1%)
(b)ショートパスの交換	21(35.6%)
(c)ドリブルからシュート	5(8.4%)
(d)ミドルシュート	0(0.0%)
(e)その他	7(11.9%)
合計	59(100.0%)

上記の表1に得点へと至った経緯に基づいて分類を行った。これによると、経緯(a)のサッカーのフィールドにおいてサイドに配置された選手が中央にいる選手へパス(サイドの選手からのセンタリング)し、得点になった場合が最も多く、サイドに配置された選手が重要であることが分かる。その次に多いものが、経緯(b)ショートパス交換から得点に至ったケースであった。これも、その内容は中央に配置された選手とサイドに配置された選手間のパス交換から得点に至ったものが15回もあり、サイドの選手の重要性が考察される。一方、経緯(c)~(e)の割合は少なかった。また、経緯(e)となったものは、相手チームのパスミスや奪ったもの、あるいはゴール付近からのフリーキックによるもの等であった。

これらのことを元にして、さらに優勝チームであるYowAI[11]において、サイドに配置されている選手が行った得点の直前の動作を分析した。YowAIの試合の中から、これまでと同じ基準で7試合(14得点)について調べ、その結果を表2にまとめた。この表2の地域1~3とは相手チームのゴールがあるフィールドをサイドラインに垂直に交わる方向に3分割し、図2のように名付けている。これにより、サイドに配置された選手はボールを保持すると地域1ではドリブルという行動を選択する割合が大きいことが分かった。また、地域2では、ドリブル、センタリング、前方向へのパスの順で行動を選択し易く、最後に、地域3では、センタリング、ドリブルの順で行動を選択し易いことが考察された。これらを基にして、初期状態のエージェントに静的な規則として与え、その後、強化学習を行い、その成否によりその報酬を与える。

#### 4.おわりに

本稿では、本研究に対する我々の基本的な考え方や行動規則に対する分類を述べた。そして、実際の試合ログにおける得点をあげた経緯に基づく分類およびサイドに配置された選手に対する考察から攻撃における静的な行動規則の例を挙げて説明した。今後は、これらの静的な行動規則と既存のサッカーエージェント(例えば、[13])を利用し、強化学習を行う実験システムを作成し、それらを用いて評価実験を行い、我々の手法の有効性を確認する予定である。

表2: サイドに配置された選手の動き

内容	地域1	地域2	地域3	計
(a)ドリブル	8	8	1	17
(b)前方向へパス	0	1	0	1
(c)横方向へパス(センタリング)	0	6	6	12
(d)後ろ方向へパス	0	0	0	0
合計	8	15	7	30

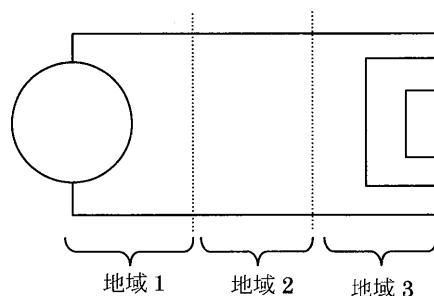


図2: 地域の分類

#### 参考文献

- [1] Richard S. Sutton, Andrew G. Barto (三上 貞芳, 皆川 雅章 共訳): 強化学習, 森北出版, 2000.
- [2] 高玉 圭: マルチエージェント学習, コロナ社, 2003.
- [3] 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, 1998.
- [4] 片山 謙吾, 奥石 尚宏, 成久 洋之: 強化学習エージェントへの階層化意志決定法の導入 - 追跡問題を例に -, 人工知能学会誌, Vol.19, No. 4, pp. 279 - 291, 2004.
- [5] RoboCup Soccer Simulation league 公式ホームページ: <http://sserver.sourceforge.net/>
- [6] 秋山 英久: サッカーシミュレーションリーグの動向, 人工知能学会誌, Vol. 17, No. 6, pp. 708 - 713, 2002.
- [7] 中島 智晴, 高谷 将裕, 有働 昌代, 石渕 久生: 進化型計算を用いたチーム戦略獲得手法の一提案, 人工知能学会, SIG-CHALLENGE 研究会, 2004年5月.
- [8] 熊田 陽一郎, 植田一博: 予測能力を持つサッカーエージェントによる協調戦術の獲得, 人工知能学会誌, Vol. 16, No. 1, pp. 120 - 127, 2001.
- [9] FC 東京: ゴールを目指せ! 攻撃サッカー, 池田書店.
- [10] Hisayuki Sasaoka, Shunsuke Muraki and Kenji Araki: "Soccer Agents Using Inductive Learning with Hand-Coded Rules", In Proceedings of 2003 IEEE International Conference on Systems, Man & Cybernetics, pp. 46 - 51, 2003.
- [11] RoboCup Soccer Japan Open 2004 公式ホームページ: <http://www.fair.or.jp/robocup/>
- [12] 電気通信大学情報工学科 竹内 郁雄研究室: <http://ne.cs.uec.ac.jp/>
- [13] Remco de Boer and Jelle R. Kok: "The Incremental Development of a Synthetic Multi-Agent System: The Uva Trilearn 2001 Robotic Soccer Simulation Team", Master's thesis, University of Amsterdam, 2002.