

# コーチのアドバイスによるマルチエージェント環境への動的適応

## Dynamic Adaptation to Multi-Agent Environment by Coach's Advice

秋山英久  
Hidehisa Akiyama

安村禎明  
Yoshiaki Yasumura

新田克己  
Katsumi Nitta

### 東京工業大学大学院 知能システム科学専攻

Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology

In multi-agent environment, it is necessary for each agent to improve his cooperative strategy dynamically according to the situation. But in the environment where agents' sensory capability is greatly restricted, it is difficult to perform the exact situation judgment. In such an environment, the coach who overlooks the whole and gives advice to a team plays an important role. We use the RoboCup Soccer Simulator as experiment environment. In this report, we propose the method to gather the data required for coach's high-level decision-making using human's observation and instruction.

## 1. はじめに

マルチエージェント環境においては、各エージェントが状況に応じて協調戦略を動的に適応させ改善していく必要がある。しかし、エージェントの知覚能力やコミュニケーションが制限されている環境では、各エージェントが正確な状況判断を行うこと自体が困難であり、戦略改善のための意思疎通も望めない。これは人間の場合も同様であり、複数人数によるチームが構成されている場合は、チームを管理するコーチ(監督)が存在することが多い。

我々はこのようなコーチの役割を果たすエージェントに注目し、人間の観察と教示によって、コーチの高度な意思決定に必要なデータの蓄積を行うシステムを提案する。実験環境として、RoboCup サッカーシミュレーションリーグで使用されているサッカーシミュレータ[Noda 1998]を用いる。シミュレーションリーグでは実機リーグに比べて戦略的な部分での研究が進んでおり、より高度な戦略が求められている。このため、シミュレータにはコーチの機能が比較的初期の段階から既に実装されている。近年はコーチに関する研究も盛んに行われ始め、コーチ専門の競技も開催されている。本稿で提案するシステムと開発するコーチは、このコーチ競技への参加を想定したものである。

## 2. RoboCup サッカーシミュレータにおけるコーチ

### 2.1 エージェントの知覚能力

RoboCup サッカーシミュレータにおけるプレイヤーのセンサ情報、特に視覚情報に関しては初期の頃から厳しい制限が設定されている。プレイヤーの知覚能力では環境を正確に認識すること自体が難しいため、プレイヤー個人の判断で戦略を効果的に変更することは事実上不可能となっている。

一方、コーチが受け取る視覚情報にはプレイヤーのような制限は加えられておらず、常にノイズ無しの完全情報が得られる。このため、コーチによる試合の正確な分析が可能となっている。

### 2.2 コーチの役割

シミュレータには実験環境の整備などの目的に使用できるオフラインコーチの機能も存在するが、本稿では実際の試合で使用可能なオンラインコーチのみを取り扱う。以下、コーチという

表記は全てオンラインコーチのことを意味する。サッカーシミュレータにおけるコーチの役割には以下のものがある。

- 試合の分析
- プレイヤへのアドバイス
- ヘテロジニアスプレイヤーの役割配分と交代

本稿では、これらのうち試合分析とアドバイスを取り扱い、ヘテロジニアスプレイヤーに関しては扱わない。

コーチからプレイヤーへのアドバイスも全てシミュレータを介して行われ、アドバイス内容は聴覚情報としてプレイヤーへ与えられる。ただし、試合への過度の干渉を防止するためにコーチのアドバイス回数は制限されており、更に、アドバイスには専用のコーチ標準言語(Standard Coach Language: CLang)を用いなければならない。CLangは、プロダクションルールの形式で記述するアドバイス言語として設計されている。

## 3. アドバイスにおける課題

### 3.1 試合の分析

プレイヤーが意味のある一連の行動を実行するには、状況の予測とそれに適したプランのキューを積むことが必要になる。しかし、先読みの探索が深くなりすぎると計算資源が不足してしまう。また、プレイヤーの限られた知覚能力では、状況の変化を予測しきれない。敵チームの行動パターンが明らかであれば、探索量を減らすことと事前のプランニングに基づいた行動の成功確率の上昇が期待できる。明確な行動パターンの抽出や統計情報の収集といった試合状況の分析が、コーチのタスクとして不可欠である。

### 3.2 アドバイス生成

コーチのアドバイスには、そのチームに本来備わっていたバランスを崩してしまうという危険性を孕んでいる。実際に、2002年のRoboCup世界大会のコーチ競技においては、ほとんどアドバイスを行わなかったチームの方が良い成績を納めており、アドバイスによってパフォーマンスが改悪される可能性が高いことを示している。

観測によるモデリング結果だけから有効な行動ルールを導き出すことは難しく、事前のテストによる調整は不可欠である。このため、リアルタイムにルールを生成するのは現実的ではない。いくつかの有効な行動パターンの事例とそのパターンをCLangで記述した辞書を事前に用意しておき、試合ではその中から適切なものを選択するという手法が望ましいと考えられる。本稿では、この事例の収集に焦点を当てる。

連絡先: 東京工業大学大学院総合理工学研究科知能システム科学専攻, 〒226-8502 神奈川県横浜市緑区長津田町4259, Tel: 045-924-5214, akiyama@ntt.dis.titech.ac.jp

## 4. システムの構成

### 4.1 試合分析と事例蓄積

有効な事例の蓄積は、シミュレータが生成した過去の試合のログからデータを抽出することで行う。ログはシミュレーションサイクルごとの物体の位置・速度情報などで構成されており、コーチの視覚情報とほぼ同一のものが得られると考えてよい。従って、ログからのパターン抽出アルゴリズムは、コーチの試合分析アルゴリズムとして再利用可能になる。

ログからは、正事例と負事例の二種類のデータを自動抽出する。ボール所有チームが同一のままプレイが続行されている状態を正事例、ボール所有チームが入れ替わったときの行動を負事例と定義する。行動列の抽出アルゴリズムには、Kaminka が提案した行動列蓄積手法を用いる[Kaminka 2003]。この手法ではデータ構造に図1のような根付き木をベースとした Trie を採用し、効率的な行動列の蓄積を可能としている。



図 1 Trie 構造による行動列蓄積

### 4.2 標準言語への変換

抽出した事例は、人間の教師が GUI 操作のツールによって評価する。サッカーのような複雑な問題に対しては、まずは人間の高度な意思決定プロセスを分析し、模倣する必要がある。人間がサッカーシミュレータのプレイヤーとして参戦するシステムの提案はすでに行われており、人間の適応能力の高さが示されている[秋田 2001]。本稿では、人間の意思決定プロセスを具体的なデータとして蓄積することも目的としている。人間による教示結果は、最終的に CLang に変換される必要がある。プロダクションルールとの親和性の高さから、決定木を用いて CLang への変換を行う。本稿で提案するシステムは、図 2 のような構成になる。

同一とみなせる正事例が複数抽出された場合は、その行動パターンが一般的に有効なものであるかどうかの評価を行うことになる。シミュレーションリーグでは各チーム間でプレイヤーの基本能力差が大きく、プレイヤーによっては実行できない行動が多い。また、シミュレータのノイズやプレイヤーのミスにより、偶然成功した事例が含まれることもある。そのため、ある程度の基本能力があればほぼ確実に実行可能と判断できる事例のみを選別する必要がある。

負事例に関しては、有効と考えられる行動ルールに置き換える必要がある。負事例に対して適切な修正ルールを与えられない場合、その行動ルール自体を禁止するというアドバイスを与えることになる。また、正事例の行動列に対する守備の行動ルールも創発する必要がある。いずれの場合にも、新しいルールを入力した後に調整のための試合を行わせ、再びそのログの検証を行うというサイクルを繰り返すことになる。

蓄積した事例をアドバイスとして実際に利用する際には、対応する行動パターンが発生する位置の統計情報を考慮する必要がある。この統計情報は対戦相手によって異なるため、実際

のアドバイス時には、ルール構造は維持したまま位置のパラメータを調整することになる。

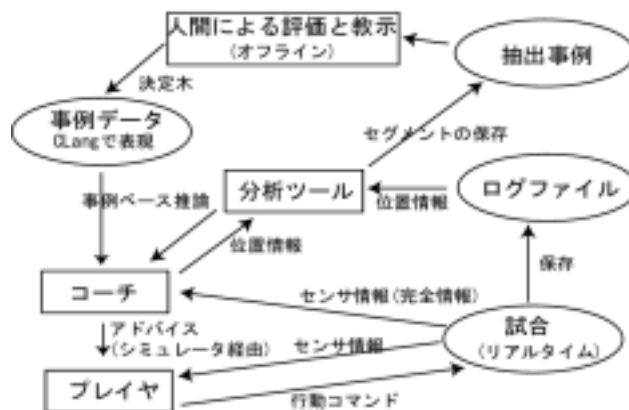


図 2 システム構成

## 5. まとめ

本稿では、より有効なアドバイスのための事例の蓄積手法を提案した。実際の試合では、事例ベース推論などにより状況に最も適した事例を選択し、パラメータを微調整してアドバイスするという方法を検討している。しかし、現状では、より汎用性のある行動ルールの抽出と蓄積を行っている段階であり、特に未知のチームに対してアドバイスを行う場合の有効性の検証が不十分である。コーチ競技自体が始まってからまだ日が浅く、アドバイスを受けるプレイヤー側もまだ必要なアーキテクチャの研究段階である[Carpenter 2003]。今後しばらくは、より汎用的なアドバイスのデータベース化とそれを理解できるプレイヤーの基本設計が必要になるだろう。

また、現在の CLang はある程度の拡張性を持った共通言語でしかなく、チーム戦略として意味のある単語などは用意されていない。言語仕様に関しては競技参加者からの提案[Reis 2002]と議論が行われている。より汎用的なアドバイスの蓄積は、言語仕様の拡張の補助となるだろう。

### 参考文献

- [秋田 2001] 秋田純一, 西野順二 et al.: RoboCup シミュレーションリーグ人間参戦システム OZ-RP の提案, 人工知能学会第 12 回 Sig-Challenge 研究会資料, 2001.
- [Carpenter 2003] Paul Carpenter et al.: Integration of Advice in an Action-Selection Architecture, RoboCup2002: The Sixth RoboCup Competitions and Conferences, Springer Verlag, 2003.
- [Kaminka 2003] Gal A. Kaminka et al.: Learning the Sequential Behavior of Teams from Observations, RoboCup2002: The Sixth RoboCup Competitions and Conferences, Springer Verlag, 2003.
- [Noda 1998] Itsuki Noda et al.: Soccer Server: A Tool for Research on Multiagent Systems, Applied Artificial Intelligence, 1998.
- [Reis 2002] Luis P. Reis et al.: COACH UNILANG – A Standard Language for Coaching a (Robo)Soccer Team, RoboCup2001: Robot Soccer World Cup V, Springer Verlag, 2003.