解 説

シミュレーションリーグとインフラ技術の技術的課題と展望

Research Issues on RoboCup Simulation League

野 田 五十樹* *産業技術総合研究所サイバーアシスト研究センター

Itsuki Noda* *Cyber Assist Research Center, National Institute of Advanced Industrial Science and Technology (AIST)

1. は じ め に

ロボカップシミュレーションリーグは、計算機内につくられた仮想的なサッカー場で、プログラム同士がサッカーを行う競技である。ロボカップのほかの実機リーグとは異なり、ロボットを実世界で動かす難しさから解放されるため、より高度な戦術に集中できるところが、シミュレーションリーグの特徴となっている。そのため、戦術の発展は非常に速く、それに伴い、研究の課題も初期に比べ徐々に変化している。本稿では、それらの変化を踏まえ、最近のシミュレーションリーグにおける技術課題について考察していく。

2. シミュレーションリーグの概要

シミュレーションリーグでは、Soccer Server (以下、サーバと略す) というサッカーのシミュレーションシステムが提供する仮想フィールドで競技が行われる。各参加チームは、11 人分のプレーヤープログラムを作成する。これらのプログラムはネットワークを介してクライアントとしてサーバに接続される (以下、プレーヤープログラムをクライアントと書く)。

各クライアントは、ネットワークを通して、仮想フィールド上の一人のプレーヤーだけを操作する。また、クライアント同士は直接通信を行ってはならないという規則がある。つまり、11 個の独立したクライアントによる分散制御でチームプレーを行わせるというのが、このリーグの趣旨である。

プレーヤーの操作のために、各クライアントは以下にあげる基本的な動作コマンドを使用できる。

- (**turn Angle**) : 体を *Angle* だけ回転させる。
- (**turn_neck Angle**) : 頭を *Angle* だけ回転させる。
- (**dash Power**) : 前方に *Power* の力で加速する。
- (**kick Power Dir**) : *Dir* の方向に *Power* の力でボールを蹴る。

ルを蹴る。

- (**catch Dir**) : ペナルティエリア内で、手の届く範囲 (2 [m] 以内) のボールを手でつかむ (キーパのみ)。

これらの動作コマンドは、**turn_neck** を除き、1 シミュレーションサイクル (100 ミリ秒) に一つだけ処理される。

一方、各クライアントは、以下のようなセンサ情報をサーバから得ることができる。

- (**see Time ObjInfo ObjInfo ...**) : 視覚情報。自分の視野 (前方 90 度) に入っている他のプレーヤーやボール、フィールド上のランドマーク (フラッグやゴール) などの相対位置情報 (*ObjInfo*) が得られる。標準では 150 ミリ秒に 1 回送られてくる。
- (**sense.body Time ...**) : スタミナや移動スピードの絶対値、動作コマンドの実行回数など、身体に関する情報が得られる。各シミュレーションサイクル (100 ミリ秒) ごとに送られてくる。

よって、クライアントは、これらのセンサ情報を元にフィールド上の状態を把握しながら、適切な動作コマンドを選択してサーバに送る必要がある。

サッカーは 11 人のチーム同士の対戦型スポーツであるので、チームワーク・協調が重要な要素となる。視野に入っていれば、お互いの位置関係から協調的な戦術を取るようプログラムすることもできるが、さらに細かい協調を行うためには、明示的なコミュニケーション手段が必要になってくる。シミュレーションリーグでは、クライアント同士の直接通信を禁じている代わりに、**say** コマンドと **hear** センサ情報を用意している。各クライアントは (**say Message**) により、回りにいるプレーヤーに対し *Message* を送ることができる。受け取る側には (**hear Time Dir Message**) という形式で情報が届けられる。ただし、*Message* の送り先を指定することはできず、周囲 50 [m] 以内のプレーヤー全員に伝えられ、受信側も送信者の方向 (*Dir*) のみ分かるようになっている。また、コミュニケーションに上限をもうけるため、各クライアントは、2 シミュレーションサイクルに、おのおののチームから一つずつしかメッセージを受け取ることができないようになっている。よって、効果的に **say** を用いなければ、肝心なときにコミュニケーション

原稿受付 2001 年 10 月 2 日

キーワード: RoboCup, Simulation, Multi-Agent, Learning, Opponent Modeling

*〒 135-0064 江東区青海 2 丁目 41-6

*Koto-ku, Tokyo

ンがとれないという事態となる。

フィールドの状態のシミュレーションは、離散的に行われている。1サイクルは100ミリ秒で、その間に受け取った動作コマンドを元に、ボールやプレーヤーの速度や位置、スタミナ状態などを計算し、状態を更新する。これらの計算では、実世界の不確実さを反映するために雑音成分を重畳している。また、速く走っているプレーヤーは曲がりにくい、スタミナが切れてくると、同じパワーをかけても速く走れない、など、アドホックな（人間的な？）規則も導入している。

3. 研究トピック

シミュレーションリーグでは実機リーグと異なり、実環境でロボットを動かす難しさから解放されている。このため、実機で問題となる実世界とのインターフェースよりも、エージェントにいかにか柔軟な振る舞いをさせるか、どのように協調的な大域的戦術にのっとりた行動をとらせるか、という点に研究の焦点が当たっている。ここでは、機械学習などの創発的手法、マルチエージェントプログラミング、エージェントモデリングの三つのテーマに関して、その問題点・現状および今後の課題について論じる。

3.1 機械学習および創発的手法

サッカーに必要なとされる行動は複雑であり、かつ微妙な調整を要求される。このため、行動を学習や遺伝的手法を用いて創発させることは、当然の要求として起こってくる。

学習・創発の対象は、問題の複雑さに応じて、動作調整、行動選択、ポジション選択、チーム戦術に分類することができる。以下では、この分類に対して、おのおの問題点と手法について論じる。

3.1.1 動作調整

動作調整では、動作コマンドに指定する *Power* などの値を求める関数の獲得が目的である。基本的には、センサ情報などを元にしたワールドモデルを入力とし、パラメータを出力とする関数の獲得であり、課題としては定式化しやすい。実際、多くのチームがこの課題に取り組んでおり、ニューラルネットやGAなどが用いられている[3][11]。

ただ、この課題の入出力は本質的に時系列データであり、時間をどのように扱うかが大きな問題となる。例えばドリブルをする場合、(1) ボールを小さく蹴る、(2) 必要に応じて方向を修正、(3) ダッシュ、が基本的なサイクルとなるが、おのおのどのタイミングでコマンドを発するか、あるいは何回繰り返すべきかなど調整すべき時間的要素は多くある。サッカーに限らず実世界を対象とする場合、時系列の扱いは避けて通れない問題であるが、現在のところ、これを一般的に解く有効な手法はまだなく、多くのチームは、タイミングや回数を固定するなど問題を単純化する方法をとっている。しかし、環境に柔軟に対応できるエージェントを作成するためにも、学習における時系列の扱いは、今

後の重要な課題となっていくと思われる。

3.1.2 行動選択

行動選択では、特にボールを保持しているとき、どのような行動（パス、ドリブル、シュートなど）を選択すべきかの判断ルールの獲得が目的である。

動作調整と異なり事象としては単発であるため、時系列の問題はあまり重要ではないが、フィールドの広い範囲の状態を入力とする必要があるため、広大な入力空間をどう扱うかが問題となる。文献[12]は、CMACと呼ぶ方法で、入力空間を少しずつずらしながらグリッド状に分割する方法を採用しており、これにより汎化能力のある強化学習を実現している。文献[10]は、実際に経験した状態や行動に必要な属性を元に、状態空間を関係の深い部分に縮退させる方法を用いている。また、文献[5]は行動決定のための多くのパラメータをGPを用いて修正する方法を用いている。

ただ、これらの手法では入力空間を人間の手でかなりアドホックに絞り込んでおり、スケーラビリティや汎用性での問題は多い。今後は、この部分の改良が望まれる。

3.1.3 ポジション選択

ボールを保持していないときにフィールド上のどこにいくべきかは、サッカーでは非常に重要な要素である。これも行動選択の一つではあるが、学習における報酬の与え方という観点では、より難しい問題になっている。つまり、行動選択であげた課題では、報酬が直接的に与えやすい（例えばパスが成功すれば報酬が与えられる）のに対し、ポジションの選択は、報酬が間接的にしか与えられない場合が多いことである。つまり、選手があるフィールド上にあるポジションを取る場合、それはパスを直接もらうだけでなく、その次のパスがもらいやすい、敵を引き付ける、味方が失敗した場合のカバーになっている、など、ボールに直接関与しない多くの効果がある。そして、これらの効果に対する評価は難しいため、どのように報酬を与えるかが問題となる。

文献[4]はこの問題に対し、観察によってほかでパスがうまく通った際に、自分のポジションが良かったからとして、勝手に自分に報酬を与えるという方法を提案し、より効果的にポジションの学習が行えることを示している。また、文献[6]らは、ポジションを含め、エージェントの行動をチーム単位ですべてGPで進化させることで、パス戦術に効果的なポジショニングが獲得できることを示している。

ただ、これらはいずれも、フィールドの絶対的な位置のポジションの学習であり、フィールドの状態によって動的にポジションを変化させることに成功しているチームはまだない。今後は、動的なポジション選択の獲得が主な課題となるであろう。

3.1.4 チーム戦術

対戦型のゲームでは、相手に合わせて戦術を変える必要

がある。特に、サッカーなどマルチエージェントのゲームの場合、サイドアタックや中央突破など、チーム全体としてとる戦術を変更することは、人間にとってもかなり高度な技術である。近年のシミュレーションリーグでは、上位チームのほとんどは守備が固く、これらを攻略するためにはチーム戦術を柔軟に切り替える機能が必要になってくると考えられる。

文献[8]はパスの成功率などから、あらかじめ用意されたいくつかのチーム戦術を動的に変更する方法を用いている。この方法では、位置よるパスの通りやすさを各エージェントが観察により求め、それに基づいてパス成功率が高くなるような戦術とポジションをとるようになる。このような方法は、適応性のマルチエージェントシステムをつくる上では有用であると考えられるが、サッカーのように動的な対象を扱う上では、少ない事例からいかに学習できるかなど、解決すべき課題は多い。

3.2 マルチエージェントプログラミング

自律的なエージェントの集団に協調的な行動をとらせるときに、どのように協調行動を記述し、どのようにその行動を発現させるかが問題となる。

文献[7]や文献[15]などは、BDI (Belief, Desire and Intention) モデルを用いてチーム戦術を記述し、それをおのおのエージェントの行動決定ルールに変換する方法を用いている。BDI モデルでは、各エージェントの信念を明示的に扱うことができ、各エージェントが部分情報しか得られないシミュレーションリーグの設定には適している。また、文献[13]などは、協調行動の記述方法として STEAM という枠組みを提案し、そのなかで、様々な協調行動の共通部分と領域知識を分けて記述する方法を提案している。

このように各チームは、協調行動のための枠組みに工夫をこらしているものの、現状では記述される協調行動は比較的静的なレベルにとどまっている。本来ならば、チーム全体で記述された行動ルールをエージェント単位で実行可能な行動ルールに変換する際、情報の部分性などから選択に曖昧さが生じる。人間の場合には、この曖昧さを解消するために、合図などのコミュニケーションや、相互のプレーの癖や嗜好などを考えにいれて動的に選択方法を変更している。ロボカップでは、試合中に選択できる協調戦術のバリエーションはまだ少ないためか、そのような高度なコミュニケーションや動的なプランニングを用いているところはない。しかしながら、徐々にレベルが高くなってきている守備を突破するためにも、多様な戦術とそれを試合中に柔軟に選択できる枠組みが、今後重要な課題となってくるであろう。

3.3 エージェントモデリングとコーチ

シミュレーションリーグでは 1999 年から、コーチエージェント (以下、コーチと略す) をチームに含めることを認

めるようになった。このコーチは、試合には直接作用することはできないが、フリーキックのときなどプレーが中断したときに選手に対しメッセージを送ることができる。また、ほかの選手と異なり、フィールド全体の視覚情報 (全プレーヤーとボールの位置) を知ることができるため、相手チームの特徴や、試合の流れなどを解析するのに容易な設定になっている。よって、コーチは、試合経過から味方や相手の特徴や弱点を探し出し、それに対する指示を選手に与えるという役割を演じることになる。

コーチの導入に伴い、相手チームをどのように解析するかというエージェントモデリングが、各チームで重要になってきている。上でも述べているように、近年の上位チームのプレーのレベルは非常に高く、得点するのは非常に難しくなっている。そうはいつても、どのチームにも弱点は存在し、そこをつかれると得点を許してしまう。よって、相手チームを既知のモデルに当てはめ、それに適した対応策を指示できるモデリング技術がコーチを実装する際に求められるようになってきている。

また、個々のプレーヤーにとっても、モデリングは重要である。2001 年に優勝した Tsinghuaeolus (清華大学、中国) では、相手チームのゴールキーパに対しフェイントを仕掛ける場面が見られた。これはある意味驚くべきことで、ゴールキーパがちゃんと、相手のプレーの先を予測して動いてこそ、フェイントの意味が出てくる。つまり、各エージェントがおのおの敵のモデリングを行っており、それにより、フェイントをかけたり、それに引っ掛かったりするようになってきている。現状ではまだ、これらの行動はハンドコーディングのレベルであると考えられるが、今後、このような場面が増えてくると考えられ、より一般的なモデリングが有効な手段になると思われる。

また 2001 年からは、Coach Competition という、コーチに対する評価セッションが行われるようになった。これは、コーチを採用しているチームが集まり、異なるチームのエージェントをコーチして、どこまで適切な指示を出してチームの能力を引き出せるかが競われる。もちろんこの場合には、コーチには共通言語が定められ、異なるチームでも指示がちゃんと伝わるように設定されている。初回であった 2001 年の大会では、まだ異なるチームを適切に指示できるコーチは現れなかったが、今後経験が積み重なるにつれ、各チーム共通の土台ができてくれば、マルチエージェントシステムとしておもしろいチャレンジになってくるだろう。

4. 評価セッション

シミュレーションリーグでは、通常のトーナメントとは別に、特定の技術課題に注目して各チームの能力を測る評価セッション (evaluation session) がもうけられている。この

技術課題は毎年変更されるが、例えば2001年には、フィールドの一部で、ボールやプレーヤーの動きやすさが変化する、という設定で、その変化にどれくらい各プレーヤーがオンラインで適応できるかということが計測された。実際、各チームがどのように適応したかはまだ解析中であるが、このような評価セッションを通じて、単に強いチームだけでなく、特定の技術的課題にチャレンジしているチームを奨励したり、マルチエージェントシステムの評価方法を探索する試みが行われている。

また、Soccer Serverのシミュレーション結果をどのように効果的に表現するか、という技術を競う、表示技術セッションもある。通常、Soccer Serverは二次元のフィールドしか表示しないが、それを三次元的に視覚化したり、あるいは選手の視野や身体状態をうまく画面に表現して、試合の様子をより分かりやすく表現する技術が評価される。特に三次元表示の場合、二次元では問題にならなかったカメラアングルの問題や、離散的な動作をいかにスムーズな動きに結び付けられるか、ということが課題となってくる。このほか、試合の様子を実況中継のように音声で説明するシステム[14]や、試合後に各チームの特徴を解析して評論したり、それを元に次の試合結果を予想するシステム[9]など、様々な試みがなされている。

5. お わ り に

シミュレーションリーグは計算機の中だけで行われるため、ロボット研究では当然の課題である実世界問題は直接は扱うことができない。しかしながら、複数プレーヤーの対戦ゲームであるサッカー本来のエッセンスは十分に反映されており、上にあげたように研究課題として非常に多くの要素を包含している。さらに、プレーヤー用プロトコルやモニター用プロトコルも公開されているため、エージェントプログラム以外に、仮想現実感や自動分析解説プログラムの研究など、周辺の技術も巻き込む形で、裾野が広がっている。

ただ、多くの研究要素を包含している分、おのおの研究課題に焦点を当てにくいという問題もあるのは確かである。しかしながら、ロボカップ自身が、トイプロBLEM用ではない総合的な技術の確立のための方法論を目指して提唱されたものであり、個々の技術課題を含めたシステム論も、今後重要になってくるだろう。

純粋にソフトウェアだけの技術競争であるシミュレーションリーグは、成果の共有が容易であり、多くの情報がWWWなどにより公開されている。実際に、大会に参加したチームのプログラムや開発用のライブラリ、開発環境、ドキュメントなどが整備されており、新規参入が容易になっている。これらの詳しい情報については、

<http://sserver.sourceforge.net/>を参照してほしい。

参 考 文 献

- [1] M. Veloso, E. Pagello, and H. Kitano (ed.): RoboCup-99: Robot Soccer World Cup III. Springer, 2000.
- [2] P. Stone, T. Balch, and G. Kraetzschmar (ed.): RoboCup 2000: Robot Soccer World Cup IV. Springer, 2001.
- [3] H. Akiyama: "A team," RoboCup 2000: Robot Soccer World Cup IV. pp.401-404, Springer, 2001.
- [4] T. Andou: "Andhill-98: A robocup team which reinforces positioning with observation," M. Asada and H. Kitano (ed.), RoboCup-98: Robot Soccer World Cup II. pp.338-345, Springer, 1999.
- [5] H. Iizuka, M. Hiramoto, H. Kawamura, M. Yamamoto, and A. Ohuchu: "Harmony team," RoboCup 2000: Robot Soccer World Cup IV. pp.449-452, Springer, 2001.
- [6] S. Luke, C. Horn, J. Farris, G. Jackson, and J. Hendler: "Co-evolving soccer softbot team coordination with genetic programming," H. Kitano (ed.), Proc. of The First International Workshop on RoboCup, pp.115-118, Aug. 1997.
- [7] J. Murray, O. Obst, and F. Stolzenburg: "Robolog koblenz," pp.628-631, RoboCup-99: Robot Soccer World Cup III. Springer, 2000.
- [8] M. Ohta: "Gemini in robocup-200," RoboCup 2000: Robot Soccer World Cup IV. pp.477-480, Springer, 2001.
- [9] T. Raines, M. Tambe, and S. Marsella: "Towards automated team analysis: A machine learning approach," Proc. of the RoboCup Workshop at the International Joint Conference on Artificial Intelligence (IJCAI'99), Aug. 1999.
- [10] M. Riedmiller, A. Merke, D. Meier, A. Hoffmann, A. Sinner, and O. Thate: "Karlsruhe brainstormers 2000 team description," RoboCup 2000: Robot Soccer World Cup IV. pp.485-488, Springer, 2001.
- [11] P. Stone: Layered Learning in Multiagent Systems: A Winning Approach to Robotic Soccer. MIT Press, 2000.
- [12] P. Stone, R.S. Sutton, and S. Singh: "Reinforcement learning for 3 vs. 2 keepaway," RoboCup 2000: Robot Soccer World Cup IV. pp.249-258, Springer, 2001.
- [13] M. Tambe, J. Adibi, Y. Al-Onaizan, A. Erdem, G.A. Kaminka, S.C. Marsella, I. Muslea, B. Salemi, and M. Tallis: "Towards flexible teamwork and learning in robocup agent: A short description of isi synthetic," H. Kitano (ed.): Proc. of The First International Workshop on RoboCup, pp.111-114, Aug. 1997.
- [14] K. Tanaka-Ishii, I. Noda, I. Frank, H. Nakashima, K. Hasida, and H. Matsubara: "MIKE: An automatic commentary system for soccer," Y. Demazeau (ed.): Proc. of Third International Conference on Multi-Agent Systems, pp.285-292, July 1998.
- [15] J. Wendler, M. Hannebauer, H.-D. Burkhard, H. Myritz, G. Sander, and T. Meinert: "Bdi design principles and cooperative implementation in robocup," RoboCup-99: Robot Soccer World Cup III. pp.531-541, Springer, 2000.

野田五十樹 (Itsuki Noda)

1992年京都大学大学院工学研究科電気工学第2専攻博士課程修了。博士(工学)。産業技術総合研究所サイバーアシスト研究センター主任研究員。神経回路網によるシンボル・パターン統合処理、学習、マルチエージェントシステム、分散シミュレーションの研究に従事。人工知能学会、日本神経回路学会員。RoboCup Federation Simulation League co-chair。