

RoboCup 小型リーグにおける相手行動の学習に関する研究

安井 興太郎

指導教員：成瀬 正

1 はじめに

RoboCup 小型ロボットリーグ (Small Size League, SSL) では、ロボットの移動速度が年々高速化してきており、2012 年の優勝チームは最高速度 3.5 m/s で制御されている。また、味方ロボットへパスを行うときのボールの速度は 4 m/s を超えることも珍しくない。こうした環境下、相手の行動を予測することは非常に重要な課題になってきた。

昨今の SSL では、主に「ボールの位置・速度」「味方ロボットの位置・速度」「相手ロボットの位置・速度」「Referee Box からの信号」の 4 つの情報をもとに、自チームの戦略を決定している。一方、人間のサッカーでは、各プレイヤーはこれらの 4 つの情報に加え、「相手プレイヤーの過去の行動情報」を認知した上で、次行動を決定している。ロボットのサッカーでも、人間と同じように相手が過去に行った戦略を認識し、それによって優位な行動を選択することができれば、より効果的にゲームを支配していくことができると考えられる。

本研究では、この課題を解決するための第一ステップとして、相手の戦略パターンを分類する問題を検討する。具体的には、2 つの試合局面間の非類似度関数を提案し、この非類似度関数を用いて過去に実行された複数の試合局面のクラスター分析を行い、戦略パターンの分類を試みる。また、分類結果を利用して、相手が実行している戦略を予測し、リアルタイムに味方に優位な行動を選択させる手法について考察する。

2 戦略の比較

原点 O をフィールド中心に取り、 x 軸を自ゴール中心から相手ゴール中心の向きに取る。そして、 x 軸を反時計回りに $\pi/2$ 回転させた軸を y 軸とする。時刻 t_1 における i 番目の相手ロボットの座標を $R_i(t_1)$ 、時刻 t_2 における j 番目の相手ロボットの座標を $R_j(t_2)$ とする。時刻 t_1 における局面と、時刻 t_2 における局面の非類似度 d を次のように定義する。

$$d(t_1, t_2) = \min_{U \in \{U_1, U_2\}} \left\{ \min_{\sigma \in S_6} \sqrt{\text{trace}(F(U)P_\sigma)} \right\} \quad (1)$$

$$U_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, U_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$F(U) = [f_{ij}], f_{ij} = \|UR_i(t_1) - R_j(t_2)\|^2$$

S_6 は 6 次の対称群、 P_σ は置換 σ の置換行列であり、 U_2 は相手ロボットの y 座標を符号反転させるための x 軸に対する鏡映変換である。 F は、時刻 t_1 におけるそれぞれの相手ロボット座標と時刻 t_2 におけるそれぞれの相手ロボット座標間の全ての組み合わせのユークリッド平方距離を成分とする 6×6 行列である。

式 (1) は、直感的には、時刻 t_1 におけるそれぞれの相手ロボットと、時刻 t_2 におけるそれぞれの相手ロボットを、ユークリッド平方距離が最も小さくなるように対応付けた時の、それぞれの距離の総和を表している。対応付けを行うことにより、時間による各ロボットへの戦術の割り当てられ方の違いを吸収できると考えられる。また、全ての戦略は x 軸に関して線対称な形が存在するので、対称形を含めて 1 つの戦略とするために、 U を用いて非類似度 d を計算している。

時刻 t_1 における局面と、時刻 T_s から時刻 T_e における連続局面の非類似度 d_1 、および、時刻 $T_s^{(i)}$ から $T_e^{(i)}$ における連続局面 X_i と、時刻 $T_s^{(j)}$ から $T_e^{(j)}$ における連続局面 X_j の非類似度 d_2 を、式 (1) を拡張して次のように定義する。

$$d_1(t_1, T_s, T_e) = \min_{t \in [T_s, T_e]} \{d(t_1, t)\} \quad (2)$$

$$d_2(T_s^{(i)}, T_e^{(i)}, T_s^{(j)}, T_e^{(j)}) = \min_{t \in [T_s^{(i)}, T_e^{(i)}]} \{d_1(t, T_s^{(j)}, T_e^{(j)})\} \quad (3)$$

相手が戦略を実行している複数の連続局面に、式 (3) を適用すれば、戦略の比較を行うことができる。

3 クラスター分析

戦略を実行している N 個の連続局面に対して式 (3) を適用すると、各戦略間の非類似度を要素とする $N \times N$ の非類似度行列が生成される。この行列を距離行列と見てクラスター分析を行うことで、戦略の分類が可能になる。本研究では、階層的クラスタリングの一種である、群平均法を用いてクラスター分析を行う。

相手チームの戦略の数は分からないため、何らかの形でクラスター数 K を決定する必要がある。本研究では、各クラスターが生成する空間の大きさに注目し、式 (4) を用いて、クラスター数を決定する。

$$W(K) = \sum_{i=1}^K \sum_{X_k \in C_i} \sum_{X_l \in C_i} d_2(T_s^{(k)}, T_e^{(k)}, T_s^{(l)}, T_e^{(l)}) \quad (4)$$

式 (4) は、各クラスター内の任意の要素対の距離の総和を、全てのクラスターに対して合計したものである。とくに、 $K=1$ の場合、 $W(1)$ は N 個の全ての要素対の距離の総和となり、クラスタリング空間の大きさを表す。式 (4) は、 K が小さくなるに連れて単調増加する。本研究では、類似性の低いクラスターが結合された場合に、 $W(K)$ が急増することを利用する。ただし、クラスタリング空間の大きさが、分析する相手チームの戦略や要素数 N (戦略の実行回数) によって異なるため、

$$W'(K) = \frac{W(K)}{W(1)} \quad (5)$$

として値が 0 から 1 の範囲をとるように規格化し、式 (6) を満たす K を、最終的に用いるクラスター数とする。

$$\arg \max_{1 \leq K \leq N} (W'(K) \leq h) \quad (6)$$

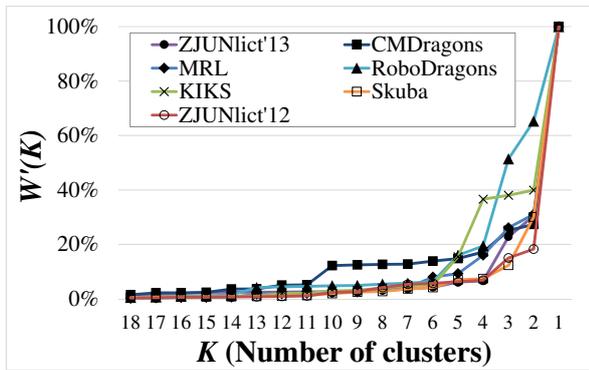
ここで、 h は閾値として与える定数である。この手法では、閾値として小さめの h を用いることで、類似性の低い要素のクラスター分離性が保証される。

4 実験

昨今の SSL での主要な得点シーンはセットプレーであり、各チーム多種多様な戦略が用いられている。そこで、以降の実験では各チームのセットプレー時における攻撃戦略に対して、提案手法を次のように適用し、戦略パターンの分類を行う。

X_i を i 回目 ($1 \leq i \leq N$) のセットプレーとし、 X_i の開始時刻を $T_r^{(i)}$ (Referee Box から指示コマンドを受け取った時刻)、 X_i においてキッカーがキックした時刻を $T_e^{(i)}$ とする。

$$T_s^{(i)} = \max(T_e^{(i)} - T_{\text{behavior}}, T_r^{(i)})$$

図1 クラスタ数 K に対する式 (5) の値

とし、式 (3) より、非類似度 $d_2(T_s^{(i)}, T_e^{(i)}, T_s^{(j)}, T_e^{(j)})$, ($1 \leq i \leq N, 1 \leq j \leq N$) を計算する。そして、生成された $N \times N$ の非類似度行列をもとに、群平均法によってクラスタの階層構造を得る。最後に、式 (6) によってクラスタ数 K を決定し、 K 個のクラスタに分類する。 T_{behavior} は、キック直前にとった行動によって戦略を分類するために必要な定数で、適用しているチームのロボットが一つの行動にかかる時間を指定する。本研究では、 $T_{\text{behavior}} = 1.0\text{s}$ を使用する。

手法の評価は、正解データのクラスタリングに対する、提案手法によるクラスタリングの適合率 P 、再現率 R を計算し、2つのクラスタリングを比較することで行う。適合率 P は異なる要素が異なるクラスタに分類されているか否かを示す評価尺度であり、再現率 R は、類似の要素が同一のクラスタに分類されているか否かを示す評価尺度である。つまり、一般に良いクラスタリングとは、適合率と再現率が共に高いクラスタリングのことを指す。しかし、両者はトレードオフの関係にあり、クラスタ数の減少に連れて適合率は低下し、再現率は向上する傾向が強い。相手戦略の学習という観点からは、異なる戦略が同一のクラスタに分類されることは、同一の戦略が異なるクラスタに分類されることよりも悪いため、本研究では適合率を重視して評価を行う。

RoboCup 2012, 2013 の試合ログデータのうち、7チームのセットプレー戦略に対して提案手法を適用し、クラスタ分析を行った。図1から、どのチームも $W'(K)$ が約5%を超えるまではほぼ横ばいに推移しており、類似性の高いクラスタが結合されていることが分かる。そして、約5%を超えると加速度的に割合が増えていくことから、徐々に類似性の低いクラスタが結合されていくことが分かる。これは、 $W'(K) = 0.05$ となる近傍に、良いクラスタリングを行う点があることを意味している。この結果に基づいて、式 (6) で与える閾値 h を、0.03 から 0.07 の間で変化させてクラスタ数 K を決定し、分類を行った。正解データと提案手法によるクラスタリングの比較結果 (表1) から、閾値 h が大きくなるに連れて適合率が平均的に低下し、再現率が平均的に向上していくことが分かる。 $h = 0.03$ から 0.06 への変化を見ると、平均適合率をわずか 2.2% の減少で抑えつつ、平均再現率を 25.8% も向上できていることが分かる。一方で $h = 0.06$ から 0.07 への変化を見ると、平均再現率は 2.1% ほど上昇するが、平均適合率は 3.8% ほど減少してしまう。以上のことから、相手戦略の学習という観点からは、式 (6) で与える閾値 h を 0.06 とすることで適合率を保証し、再現率も維持できる、良いクラスタリングを行うことができることが分かる。また、この時平均適合率は 0.928、平均再現率は 0.921 となり、提

表1 正解データと提案手法によるクラスタリングの比較

| 対象チーム | $h = 0.03$ | | $h = 0.06$ | | $h = 0.07$ | |
|-------------|------------|-------|------------|-------|------------|-------|
| | 適合率 | 再現率 | 適合率 | 再現率 | 適合率 | 再現率 |
| ZJUNlict'13 | 1.000 | 0.951 | 0.943 | 1.000 | 0.845 | 1.000 |
| CMDragons | 0.929 | 0.556 | 0.950 | 0.966 | 0.950 | 0.966 |
| MRL | 0.907 | 0.796 | 0.849 | 0.918 | 0.849 | 0.918 |
| RoboDragons | 1.000 | 0.386 | 0.960 | 0.867 | 0.923 | 0.867 |
| KIKS | 1.000 | 0.747 | 0.987 | 1.000 | 0.987 | 1.000 |
| Skuba | 0.812 | 0.636 | 0.808 | 0.783 | 0.677 | 0.845 |
| ZJUNlict'12 | 1.000 | 0.573 | 1.000 | 0.913 | 1.000 | 1.000 |
| チーム平均値 | 0.950 | 0.663 | 0.928 | 0.921 | 0.890 | 0.942 |

案手法の有効性が確認できる。

5 リアルタイム学習への応用

クラスタ C_i に含まれる複数の連続した試合局面 X_j と、現在時刻 t における試合局面の非類似度 d_3 を、次のように定義する。

$$d_3(t, C_i) = \frac{1}{|C_i|} \sum_{X_j \in C_i} d_1(t, T_r^{(j)}, T_e^{(j)}) \quad (7)$$

現在時刻 t における試合局面と、クラスタ C_i に含まれる X_j の非類似度の平均値を求めている。提案手法によって分類された全てのクラスタに対して式 (7) を計算することで、相手が現在実行している戦略が、それまでに実行したどの戦略に近いかをリアルタイムに求めることができる。

図2は、一例としてSkubaの37回目のセットプレー X_{37} に対し、式 (7) を計算したものである。クラスタ C_3 のみ、式 (7) の値が常に低い値で推移しており、式 (7) を用いれば過去に実行されたクラスタ C_3 に対応する戦略が、今もう一度実行されそうであることが予測できる。

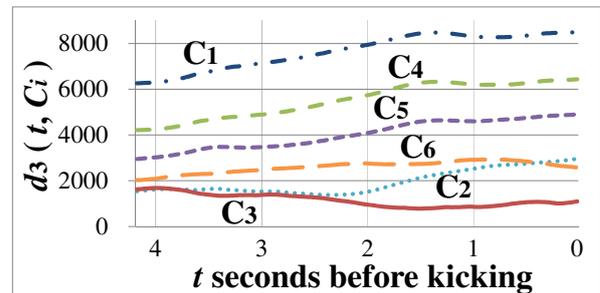


図2 各クラスタごとの式 (7) の値: Referee Box によるセットプレー開始信号を受け取ってから、キッカーによるキックが完了するまでの推移。

6 おわりに

本研究では、2つの試合局面間の非類似度 d を定義し、 d を用いて複数の局面のクラスタ分析を行うことで、相手の戦略を分類する手法について提案した。そして、提案手法を各チームの戦略へ適用した。その結果、チーム平均値で適合率 92.8%、再現率 92.1% の精度でクラスタリングを行うことができ、手法の有効性を確認した。また、分類結果から相手の戦略をリアルタイムに学習する手法についても考察した。今後は、戦略分類のさらなる高精度化、分類結果を用いた学習手法の詳しい検討、および、インプレー中の戦略分類の検討を進める予定である。

参考文献

- [1] K. Yasui, K. Kobayashi, K. Murakami, and T. Naruse, "Analyzing and Learning an Opponent's Strategies in the RoboCup Small Size League," LNAI, Springer, 掲載予定