

# サッカーエージェントにおける局面評価関数の強化学習

田川諒<sup>†</sup> 五十嵐治一<sup>†</sup>

**概要** : RoboCup サッカーシミュレーションリーグ 2D は、コンピュータ内の仮想フィールド上でエージェント同士がサッカーの試合を行う競技であり、人工知能研究のテストベッドとしても利用されている。本研究では、プレイヤーエージェントがボールを保持した際の行動決定に用いる局面評価関数の重みを、人間の主観評価を報酬として用いる強化学習の方法を提案する。

**キーワード** : RoboCup, agent2d, チェーンアクション, 局面評価関数, 強化学習

## Reinforcement Learning of State Evaluation Functions in Soccer Agents

RYO TAGAWA<sup>†</sup> HARUKAZU IGARASHI<sup>†</sup>

**Abstract**. RoboCup Soccer Simulation League 2D is a game competition where independently moving software players (agents) play soccer on a virtual field inside a computer. This league is used as a testbed in research on multi-agent systems. We propose and develop a learning system that learns weight parameters in the state evaluation function of an agent that holds the ball. In our system, spectators evaluate actions and states observed in simulation soccer games and give rewards for learning the policy of agents.

**Keywords**: RoboCup, agent2d, Chain Action, State Evaluation Function, Reinforcement Learning

### 1. はじめに

RoboCup サッカーシミュレーションリーグ 2D[1]は、多くのチームがベースとなるプログラムとして「agent2d」[2]を採用している。agent2d は近年、ボール保持者の行動を、ゲーム探索木と局面評価関数を用いて決定する「チェーンアクション」と呼ばれる手法を実装している。

しかし、agent2d に組み込まれている評価関数は極めて単純であった。そこで、我々はサッカーの局面評価に関するヒューリスティクスを用いて、複数の評価項の線形和により構成された評価関数を考案し、強化学習による重みの学習を行った[3]。この時、報酬はボールの移動距離に応じて自動的に与えていたが、学習後のチームが agent2d に勝ち越すことは出来なかった。そこで、質的に優れると思われる人間の主観評価を用いて評価関数中の重み係数を決定することを試みた。先行研究[4]では、試合後に局面の静止画を被験者に見せ、局面の評価値と評価関数中の各項の値との相関強度から重みを決定していた。しかし、これは何らかの学習理論に基づく方法ではなく、処理もバッチ処理的な方法であった。

本研究では、強化学習理論に基づいたオンライン的な学習方法を用いることを試みた。すなわち、試合中に観戦者が自由に局面の優劣を評価し、その情報を報酬とするオンライン強化学習システムを作成した。強化学習としては報酬にマルコフ性を必要としない方策勾配法を用いた。

### 2. サッカーシミュレーションリーグ 2D

RoboCup サッカーシミュレーション 2D リーグの公式シミュレータは、UDP/IP によるクライアント・サーバ方式で構成されている[5]。試合中各エージェントは、サーバからボールやプレイヤーの座標等の環境情報を受け取り、これを基に行動の決定を行う。そして各エージェントは、選択した行動を「kick」や「dash」といった制御コマンドとしてサーバに送信し、サーバはそれを基に環境情報を更新する。

### 3. agent2d とチェーンアクション

#### 3.1 チェーンアクション

本研究で使用した agent2d は、RoboCup2010 世界大会の優勝チーム「HELIOS」を基にサンプルプログラムとして作成され、公開されているチームプログラムである。基本的な行動戦略等は既実装されており、チーム開発を支援するライブラリも豊富に用意されている。

agent2d の ver3.1.0 以降では、チェーンアクションと呼ばれる手法により、ボール保持時における行動の決定を行う[6]。すなわち、ボール保持者はパスやドリブル、シュートといった行動を枝、行動の結果生じる予測状態をノードとする探索木を生成する。各ノードは、評価関数によって優劣が点数化される。agent2d では、点数が最も高いノード(必ずしも葉ノードとは限らない)を最良優先探索によって探索し、そのノードへ至る行動を決定論的に選択する(max 戦略)。しかし、本研究では評価関数の学習を行うため、次のボルツマン分布で定義される確率的方策を使用した。

<sup>†</sup> 芝浦工業大学工学部情報工学科  
Shibaura Institute of Technology

$$\pi(a|s;\omega) = \frac{e^{E_a(a,s;\omega)/T}}{\sum_x e^{E_a(x,s;\omega)/T}} \quad (1)$$

ここで、 $E_a(a, s; \omega)$ は局面 $s$ における行動 $a$ の評価値である。先行研究[3]では、行動 $a$ の実行で得られる局面以下のノードで局面評価値が最も高い局面 $s_a$ の局面評価値 $E_s(s_a; \omega)$ で、次のように置き換えている。

$$\pi(a|s;\omega) \approx \frac{e^{E_s(s_a;\omega)/T}}{\sum_x e^{E_s(s_x;\omega)/T}} \quad (2)$$

図1は現在局面 $s$ からの行動を枝とする探索木の例を示しており、数字はその局面の局面評価関数 $E_s(s; \omega)$ の値を表す。

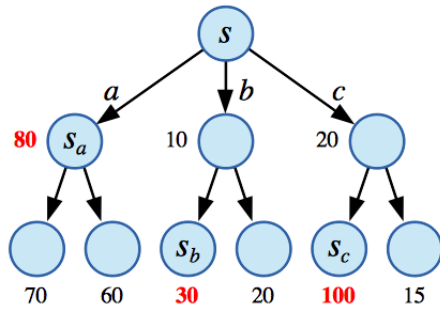


図1 探索木と局面評価関数の値の例

Figure 1 Example of the search tree and values of the state evaluation function.

この例では、行動 $a, b, c$ の行動評価関数 $E_a(a, s; \omega)$ の値はそれぞれ80, 30, 100となるが、これらの値は局面 $s_a, s_b, s_c$ を局面評価関数 $E_s$ により評価した値である。

### 3.2 局面評価関数

agent2dにおける局面評価関数は、ボールと敵ゴールの距離のみに依存した単純な関数である。そこで、我々は次の(3)のような評価関数をこれまでに提案してきた[3]。

$$E(s;\omega) = U_0(s)[\omega_1 U_1(s) + \omega_2 U_2(s)] + \sum_{i=3}^5 \omega_i U_i(s) \quad (3)$$

ここで、 $U_i(s) (i = 0, 1, \dots, 5)$ は、局面 $s$ を評価する際に有用だと思われる6つのヒューリスティクス(先見的知識)を表した関数であり、 $\omega_i (i = 1, 2, \dots, 5)$ は重み係数である。(3)の評価関数 $E(s; \omega)$ は、正/負で絶対値が大きいかほど自身の所属チームが優勢/劣勢な局面を表すよう定義している。

(3)の各項の内容を簡単に述べると、 $U_0 (\in \{-1, 0, 1\})$ は、ボール保持者の所属チームを表している。 $U_1$ はボール保持者と敵の距離、 $U_2$ はボール保持者から見て敵ゴール側にいる敵と味方の人数比、 $U_3$ はボールと両ゴールの距離、 $U_4$ はボールと両チームのプレイヤーの距離、 $U_5$ はボール周辺のプレイヤーの分布を、それぞれ表している。 $U_1, U_2$ は $[0, 10]$ 、 $U_3$ から $U_5$ は $[-10, 10]$ の区間内の値を取るよう正規化している。

## 4. 方策勾配法による評価関数の学習

### 4.1 学習則

エピソードを定義し、エピソード終了後にエピソード全体を評価して報酬を与える。この報酬の期待値を最大にしたい。(3)の方策中のパラメータ $\omega$ の学習則は、強化学習の一種である方策勾配法[7]によると、次のように表される。

$$\Delta\omega = \varepsilon \cdot r \sum_{t=1}^L e_\omega(t) \quad (4)$$

$$e_\omega(t) = \partial \ln \pi(a(t)|s(t);\omega) / \partial \omega \quad (5)$$

ただし、 $s(t)$ は時刻 $t$ における局面、 $a(t)$ は選択された行動、 $L$ はエピソード長、 $\varepsilon$ は学習係数である。(5)に(2)を代入すると、 $e_\omega(t)$ は次のように表される。

$$e_\omega(t) = (1/T) \left[ \partial E_s(s_{a(t)}; \omega) / \partial \omega - \sum_x \pi(x|s(t); \omega) \partial E_s(s_x; \omega) / \partial \omega \right] \quad (6)$$

(4)~(6)では、大きな報酬 $r$ が与えられたエピソードにおいて、実際に選択された行動 $a$ の選択確率が大きくなるように行動 $a$ の値 $E_s(s_{a(t)}; \omega)$ を大きくする勾配方向に $\omega$ を更新している。

### 4.2 学習システムの作成

本研究では、図2に示すような人間による報酬を用いた強化学習システムを作成した。図2において、試合観戦者はモニタプログラムを通じて試合を観戦する。その間、投票画面の操作を通じて、評価対象チームのプレーの評価を行う。学習プログラムは、試合観戦者による投票から報酬値を計算し、重みの学習に用いる。重みの更新は、(4)~(6)の学習則に従って試合中の各エピソードの終了時に行われる。本研究では、コーチエージェントを利用して、学習計算に必要な情報を各プレイヤーエージェントから収集し、各エージェントのパラメータ更新も試合中にリアルタイムで実行する。なお、局面評価関数は全エージェントで同一のものを利用すると仮定している。

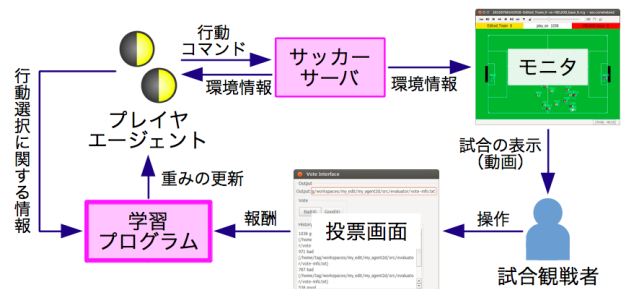


図2 学習システムの構成

Figure 2 Architecture of the proposed learning system.

## 5. 実験

### 5.1 状態 $s$ と行動 $a$

各プレイヤーが得た、敵と味方を合わせた計22人のプレイヤー

とボールの位置情報を状態 $s$ と定義する。また、agent2dのチェーンアクションでは、シュート、パス(ダイレクトパス、リードパス、スルーパス、クロスパス)、ドリブル(ショートドリブル、ロングドリブル、シンプルドリブル)、ホールド(ボールの保持)の行動に対応するクラスが実装されており、今回はこの9種類を行動 $a$ として用いた。

また、agent2dではボールのX座標のみを考慮した評価関数を使用しており、敵にボールを奪われると予測した局面の枝刈りを行うことで局面の安全性(敵へのボールの奪われにくさ)を考慮しているが、安全性を考慮しない評価関数に合わせて枝刈りが厳しく設定されているため、ボールを奪われるリスクを伴う有効な行動が排除される恐れがある。実際に、agent2dの枝刈りではパスの安全性を重要視しすぎて、後方へのパスや無駄なパスのやり取りが多い傾向がある。そこで本研究では、安全性を考慮するヒューリスティクスを含む(3)式に示した評価関数を実装し、パス(リードパス、スルーパス)とドリブル(ショートドリブル、ロングドリブル)について、味方や敵との位置関係による枝刈りを緩和するよう変更した。図3に、枝刈りの緩和による生成数の変化の一例を示す。

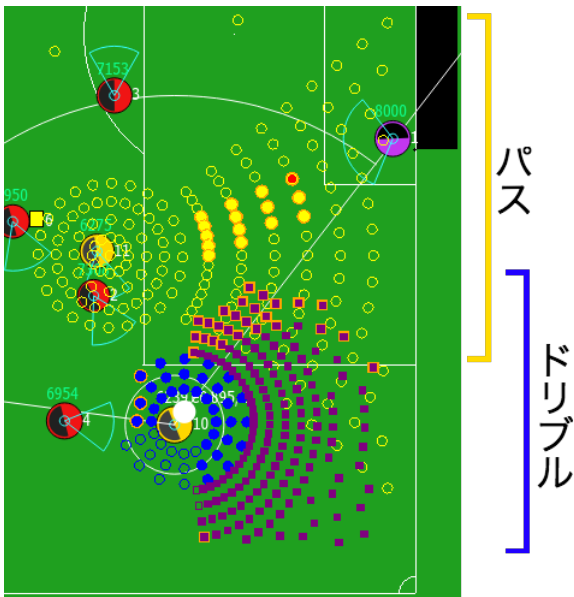


図3 枝刈りの緩和による生成数の変化

Figure 3 Change of nodes after relaxing the pruning constraints.

図3において、黄色の円、青色の円、紫色の四角はそれぞれ、パス(ダイレクトパス、リードパス、スルーパス)、ショートドリブル、ロングドリブルのボールを送る先の目的地点の候補(いずれも探索の深さは1段目)を表している。周囲がオレンジ色の候補は、agent2dのデフォルトの枝刈りでは生成されず、今回枝刈りの緩和によって新たに生成されるようになった候補である。また、内側が塗りつぶされていない箇所は枝刈りによって除かれた候補、赤で塗りつぶされた箇所はチェーンアクションで実際に選択した候補である。図3に示した局面

では、チェーンアクションを用いた行動決定の結果、デフォルトの枝刈りの設定では生成されなかった候補(スルーパス)を実際に選択し、その後得点を挙げた。今回は、このようにagent2dのパス先候補の枝刈り条件を緩めることにより、スルーパスやリターンパスなどの攻撃的で、かつ、協調的なパス回しを実現することを試みた。

## 5.2 エピソードと報酬 $r$

今回の学習実験では、チェーンアクションの探索の深さを2段に設定した。したがって、連続した2回の行動列からなる協調行動をボール保持者は計画できる。その計画が実現されたかどうかを評価するには、少なくとも2回の行動結果を観測する必要がある。つまり、試合観戦者には、投票の直前の行動だけでなく、探索した深さの分だけ前の行動からの一連の行動列を評価して報酬を与えさせる方針とした。そこで、エピソードは試合観戦者が投票した時点から直前2回の行動と定義する。投票は、自チームのプレイヤーがボールを保持してから、敵プレイヤーに奪われるか、ファール等で審判によりプレーが中断されてから10サイクル以内に行われたものを有効とする。

報酬 $r$ は、エピソード中における試合観戦者の投票によって決定する。投票は「Good」と「Bad」の2種類を用意した。「Good」/「Bad」の投票を1回行うごとに報酬を10加算/減算する。なお、投票後10サイクル以内であれば連続で投票を行うことで、報酬を上乗せする事が可能である。図5に試合観戦者が操作する画面を示す。

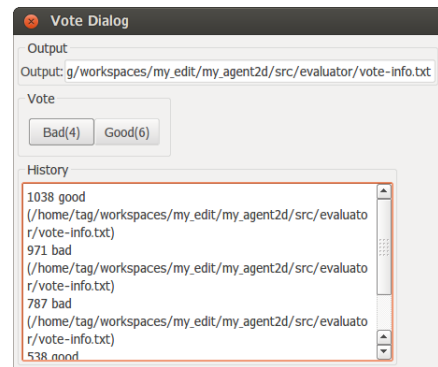


図5 投票画面

Figure 5 Voting system.

試合観戦者へは、エピソード中に評価対象のチームが得点に繋がる行動を取ったと感じたら「Good」を、逆に得点に繋がらない無駄な行動を取ったと感じたら「Bad」を投票するよう指示した。投票時間終了時に報酬を与え、重みを更新する。

その他、学習中のチームが得点を挙げた際には報酬を100だけ自動的に与えた。また、今回の実験では全エージェントが同一の評価関数を持ち、ボール保持時の行動決定にチェーンアクションを用いるとした。

### 5.3 学習実験の条件

学習対象のエージェントは11人全員とし、評価関数の重みの初期値は全て5とした。試合はagent2d-3.1.1を相手に10試合行い、被験者1名が試合の観戦および投票を行った。チェーンアクションにおける探索について、深さは2段、探索ノード数の上限は1000とした。

### 5.4 学習実験の結果

表1に、学習実験における試合の内容を示す。「得失点」の欄は、左が学習チームの得点、右が失点(対戦相手の得点)を表す。「+r」と「-r」はそれぞれ、正の報酬および負の報酬を与えた回数(1エピソード中に20以上または-20以下の報酬を与えた場合も1回とする)を表す。また、表2に学習実験終了時の重みを、図5に学習実験における重みと報酬の推移を、それぞれ示す。図5において、左の縦軸は重み、右の縦軸は報酬、横軸はエピソードを、それぞれ表す。また、報酬rは80エピソードごとに平均化している。

表1 学習実験における戦績

試合	1	2	3	4	5	6	7	8	9	10
得失点	0-1	0-1	0-2	0-2	0-0	0-2	0-5	0-0	0-4	0-0
投票数	94	88	92	84	78	75	80	81	49	58
+r	24	23	28	29	23	25	26	22	14	26
-r	70	65	64	55	55	50	54	59	35	32

表2 学習後の重み

$\omega_1$	$\omega_2$	$\omega_3$	$\omega_4$	$\omega_5$
3.68	4.17	12.71	5.30	0.08

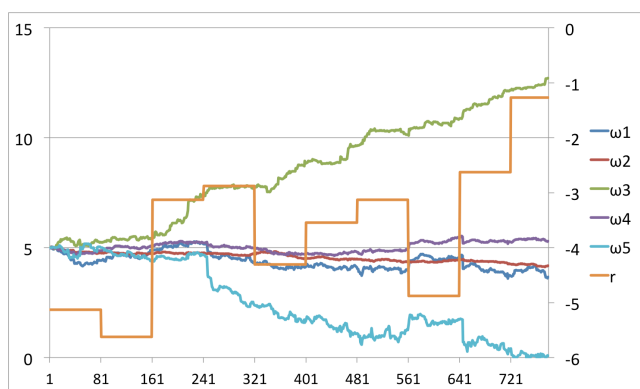


図5 学習実験における重みと報酬の推移

Figure 5 Change of weights and rewards in the learning experiments.

今回の学習実験では、学習チームが試合中に得点を挙げることは無く、勝敗は10試合全てで負けあるいは引き分けという結果となったにも関わらず、投票による報酬は右肩上がりになった。重みは最終的に $\omega_3$ が特に大きくなり、 $\omega_5$ がほぼ0となった。

### 5.5 評価実験

学習後のチームの強さを確認するため、未学習のチームと300試合行った。実験条件は基本的に学習実験と同等の条件だが、本実験では方策にGreedy選択を使用した。その結果を表3に示す。表3において、①は学習後のチームと未学習のチーム、②は未学習のチーム同士、③は学習後のチームとデフォルトのagent2dの試合をそれぞれ示す。「ボールX座標平均」は試合中のボールのX座標の平均値(X軸は敵ゴール方向が正で、コート中心を原点とする)を示す。

表3 評価実験における戦績

試合	勝-負-分 [%]	勝数/負数	平均得点-失点	ボールX座標平均
①	26.67-5.67-67.67	4.71	0.36-0.10	2.89
②	10.33-12.33-77.33	0.84	0.13-0.14	-0.21
③	4.33-79.00-16.67	0.05	0.37-1.97	-5.71

表3から、学習後のチームは未学習チームと比べて勝率が10.33%から26.67%に向上している。平均得点が増加し、ボールのX座標平均も未学習チームのゴール側に寄っていることが分かる。一方でagent2dとの対戦では、平均得点は未学習チームとの対戦と同程度の0.37点であるが、平均失点が1.97点と平均得点と比べて大きく、勝率も4.33%に留まっている。

## 6. 考察

### 6.1 重みの変化

今回の実験では、エピソードを投票直前の最大2回の行動と定義しているため、最大2回の行動決定に対してそれぞれ重みの変化量の計算を行っている。重みの変化量については、(3)~(6)式より、次の(7)式のように表せる。

$$\Delta\omega_i \propto r \left[ (1 - \pi(a|s))U_i(s_a) - \sum_{x \neq a} \pi(x|s)U_i(s_x) \right] \quad (7)$$

$r > 0$ の場合、角括弧内の値が大きいかほど $\omega_i$ は増加し、逆に角括弧内の値が小さいほど $\omega_i$ も小さくなる。これは、被験者が良いと感じた行動 $a$ によって遷移する予測局面 $s_a$ について、他の予測局面より $U_i$ の値 $U_i(s_a)$ が大きかった場合、行動 $a$ が選ばれたのは $U_i$ の寄与が大きかったためだと考えてその項の重み $\omega_i$ を増加させ、逆に $U_i(s_a)$ の値が小さかった場合には小さいから大きな報酬を得たと考えて $\omega_i$ を減少させると解釈できる。また、実際に選んだ行動 $a$ の選択率 $\pi(a|s)$ が小さく、かつ報酬 $r$ が大きい場合は、その行動の選択確率の値を高めるように $\omega_i$ を大きく増加させている。

例えば図5では、646番目のエピソードの終了後、 $\Delta\omega_5$ は-1.01と小さな値をとっている。このエピソードでは、相手陣地のコーナーをショートドリブルで攻め、その後ペナルティエリア内の味方へとリードパスを出した局面に対して20の報酬が与えられた(ただし、レシーブには失敗している)。このエピソードにおける投票直前の2行動に対応する学習用の

データを、表 4 に示す。

表 4 エピソード 646 における学習データ

行動 $a$	$\pi(a s)$	$(1 - \pi(a s))U_5(s_a)$	$\sum_{x \neq a} \pi(x s)U_5(s_x)$
ショートドリブル	0.007	4.591	5.768
リードパス	0.006	-8.999	5.968

このエピソードで特に $\Delta\omega_5$ の絶対値が大きくなったのはリードパスを選択した時である。探索において選択した予測局面では、 $U_5$ の値は-9.051 と小さく、またその予測局面を選ぶ確率も約 0.50%と比較的小さかった(最も選ばれる確率の高かった予測局面の $U_5$ の値は 7.616, 選択率は約 3.1%)にも関わらず、リードパスの実行後に多く報酬を得た。(7)式にこれらの値を代入すると、角括弧内の値は-14.968 と小さな値を取ったため、 $\omega_5$ が大幅に減少した。

また、217 番目のエピソードではショートドリブルからスルーパスが実行され、30 の報酬が与えられた結果、 $\Delta\omega_3$ は 0.37 とやや大きな値をとっている。このエピソードにおける学習用のデータを、表 5 に示す。

表 5 エピソード 217 における学習データ

行動 $a$	$\pi(a s)$	$(1 - \pi(a s))U_5(s_a)$	$\sum_{x \neq a} \pi(x s)U_5(s_x)$
ショートドリブル	0.015	0.533	0.616
スルーパス	0.104	7.587	1.220

このエピソードでは、スルーパスを選択した時の学習で $\Delta\omega_5$ の絶対値が大きくなっている。スルーパスの選択率自体は他の行動と比べても高かったが、他の行動による予測局面と比べて選択した予測局面における $U_3$ が他の予測局面に比べて特に大きかったため $\omega_3$ も上昇した。

一方で、他の重み(特に $\omega_2$ )は全体的にあまり大きな変化は見られなかった。次の表 6 は、全試合の学習に用いた予測局面のデータから、1 回の探索あたりの $U_i$ の分散を計算したものである。この値が大きいほど探索において各予測局面で $U_i$ の値にばらつきがあり、0 に近づくほどどの予測局面を選んでも値があまり変化しないという事を意味する。

表 6 学習データにおける $U_i$ の分散

$V[U_1]$	$V[U_2]$	$V[U_3]$	$V[U_4]$	$V[U_5]$
5.53	0.24	3.83	1.19	11.98

今回のエピソード設定では特に $U_2$ の分散が小さく、図 5 の傾向と一致していることから、 $U_2$ の値に各予測局面で差が少なかったことが、1 回の学習での $\omega_2$ の変化量が小さくなった

要因と考えられる。

## 6.2 勝率および平均得点の変化

学習実験で報酬を与えた局面について、20 点以上の報酬を与えた回数は 10 試合(計 779 エピソード)中 26 回で、ドリブルによる敵ディフェンスラインの突破やそれを促すパスを試みた局面が 13 回、スルーパスによる突破を試みた局面が 11 回だった。また、30 点以上の報酬を与えた回数は合計 4 回で、いずれもスルーパス実行直後の局面だった。そこで、評価実験における未学習チームとの試合(表 3 の①および②)で、学習後のチームおよび未学習チームがスルーパスを成功させた局面の数等を集計した。ただし、今回は手作業で集計を行ったため、各チームが 2 得点以上挙げた試合を集計の対象としている。なお、ここではスルーパスを受け取った時点で、「ボールまたはレシーバがレシーバに最近接の敵を X 座標で抜いている」、または「レシーバが敵のディフェンスラインやペナルティエリアへ進入している」ことのいずれかを満たしていれば「スルーパス成功」とした。表 7 に集計結果を示す。

表 7 評価実験での未学習チームに対するスルーパス成功数

チーム	スルーパス成功数	スルーパス後シュート成功数	累計得点	試合数	試合あたりスルーパス成功数
学習後	30	15	35	17	1.76
未学習	1	1	8	4	0.25

未学習チームと比べて学習後のチームは、1 試合当たりのスルーパス成功数が約 7 倍に増加した。学習後のチームについて、スルーパス成功数 30 回の内、スルーパスのレシーブ直後 100 サイクル以内にシュートを成功させた回数は 15 回(50.00%)、集計対象の試合における得点の累計得点 35 点中 15 点(42.86%)がスルーパス直後のシュートによるものである。以上より、学習後のチームが未学習のチームと比べて勝率や平均得点が向上したのは、スルーパスの成功率が向上したためと考えられる。

学習後のチームと agent2d の試合を見ると、スルーパス後に得点を挙げる場合は度々見られたが、枝刈りの緩和によって生成されたリスクの高い候補を選択し、その後ボールを奪われてしまうという局面が目立った。これが低い勝率の原因の一つである。問題はリスクを冒さずに済む局面で危険な行動を選択することであり、安全性に関する一部の項の重みがあまり変化していないことから、リスクとリターンのバランスが上手く取れていないのではないかと考えられる。また、全ての選手が重みも含めて同一の局面評価関数を持っていることも問題であろう。

## 7. 終わりに

本研究では、複数のヒューリスティクスで構成された評価関数において、パラメータを人間の主観評価を報酬として使

用する強化学習システムを構築した。本システムを用いて学習実験を行った結果、学習前のチームと比べてスループスの成功回数が増加し、勝率や平均得点が増加した。

しかし、今回の実験は被験者が 1 名のみであり、被験者の価値観によって結果が変わる可能性がある。また、人間による試合の投票では学習回数を重ねることが難しいという問題がある。今後は、学習に複数人の被験者による投票を用いることによってこれらの問題に対応する予定である。また、プレイヤーの役割ごとに局面評価関数の関数形や重みパラメータが異なるような自律分散型の学習法も試みる予定である。

**謝辞** 本研究はJSPS 科研費 26330419 の助成を受けました。ここに感謝の意を表します。

## 参考文献

- 1) 野田五十樹, “RoboCup におけるマルチエージェントシミュレーション”, セルオートマトン・シンポジウム講演論文集, pp. 63-68, 2011.
- 2) RoboCup tools / soccer simulation wiki  
<http://sourceforge.jp/projects/rctools/releases/>
- 3) 谷川俊策, 五十嵐治一, 石原聖司, “RoboCup サッカーシミュレーションリーグ 2D における局面評価関数”, 電子情報通信学会総合大会, D-8-5, 2013.
- 4) 田川諒, 谷川俊策, 五十嵐治一, “agent2d のチェーンアクションにおける局面評価関数の重み調整”, FIT(情報科学技術フォーラム)講演論文集, Vol. 13, No. 2, pp.285-288, 2014.
- 5) 秋山英久, “RoboCup サッカー2D シミュレーションリーグ解説:仕組みと環境構築”, 知能と情報, Vol. 23, No. 5, pp.714-720, 2011.
- 6) 秋山英久, “アクション連鎖探索によるオンライン戦術プランニング”, 人工知能学会研究会資料, SIG-Challenge-B101-6, pp. 23-28, 2011.
- 7) 五十嵐治一, 石原聖司, 木村昌臣, “非マルコフ決定過程における強化学習—特徴的適正度の統計的性質—”, 電子情報通信学会論文誌 D, Vol. J90-D, No. 9, pp.2271-2280, 2007