法政大学学術機関リポジトリ
HOSEI UNIVERSITY REPOSITORY

PDF issue: 2024-09-01

サッカービデオゲームにおけるエージェントの強化学習とチームフォーメーションの関係

赤塚, 洋介 / AKATSUKA, Yousuke

(発行年 / Year)
2008-03-24
(学位授与年月日 / Date of Granted)
2008-03-24
(学位名 / Degree Name)
修士(理学)
(学位授与機関 / Degree Grantor)
法政大学 (Hosei University)

サッカービデオゲームにおける
エージェントの強化学習とチームフォーメーションの関係

Relationship between Agent's Reinforcement Learning and Team Formation in Soccer Video Games

赤塚　洋介
Yosuke Akatsuka
法政大学大学院情報科学研究科情報科学専攻
E-mail: yousuke.akatsuka.dn@gs-cis.hosei.ac.jp

Abstract

This paper reports on the result of investigating the influence that the formation of a team exerts on the reinforcement learning of agents in soccer video games. We have already suggested the method that agents automatically acquire algorithms by using the reinforcement learning by the event-driven classifier learning system. Moreover, we pay attention to reward values in reinforcement learning, we suggest the method for improving of the reinforcement learning efficiency by quasi-optimization of the reward allotment. As a result, we showed that there is possibility that the configuration of the reward allotment influences the team strategy, which is appeared as series of action of every player. In this paper, we verify the influence of team formation exerting on the reinforcement learning to consider the efficiency of learning and team strategy in detail. To analyze the difference of the learning efficiency depending on team formations, the experiment is carried out by making two teams fight repeatedly with applying several formations to them. Furthermore, we investigate the influence of team formation exerting on team strategy by verifying the decision making of each agents. As a result, it is shown that the efficiency of learning is greatly changed by the compatibility of formations. In addition, we show the possibility that the suitable team strategy for applied formation is generated by the learning.

Supervisor: Yuji Sato

1. はじめに

ビデオゲーム作成の現場におけるゲームアルゴリズムの開発は，人手によって行われるのが一般的である．またアルゴリズムの記述方法として，IF-THEN 形式のプロダクションルールを用いることも一般的になっている．IF-THEN 形式のルールを用いることの利点は，アルゴリズムの可読性が上がることや，デザインやメンテナンスが容易になるという点である．近年，インターネットの爆発的な普及によりビデオゲーム開発のおかれている状況に変化が現れている．引き起こされた問題の例として，ゲームユーザの爆発的な増加，ゲーム環境のオンライン化に伴うゲーム環境の複雑化やアルゴリズムの短寿命化が挙げられる．我々はゲームの一例としてサッカービデオゲームを取り上げ，上記の問題を解決する方法としてバケツリレーアルゴリズム[1-6]を用いたクラシファイアシステム[7-11]による強化学習をエージェントに適用することで，自動的に行動アルゴリズムを獲得する手法を提案した[12]．また効率的に学習を行うための案として，ゲーム内の役割によって報酬値に偏りを持たせる学習方法を提案した[13-14]．上記二つの提案を通して，各プレイヤーの行動内容に偏りが生じ，学習効率やチーム全体としての一連のアクション（チーム戦術）に影響を及ぼしている可能性があることを示した．具体的には，対戦相手の戦術によって獲得されたチーム戦術が異なっていたことから，個々のエージェントは対戦相手に適した意思決定を学習することが出来る可能性が示した．

本稿では，学習効率と獲得されるエージェントの意思決定について深く検証すべく，フォーメーションが学習に及ぼす影響について分析した結果を報告する．フォーメーションに依存した学習効率やチーム戦術の違いを，チームの勝率，各選手のさまざまなアクションの頻度，成功数，成功率から分析する．以下，2 章では，実験に用いるサッカーゲームの概要と従来研究について紹介する．3 章ではチームフォーメーションと学習効率の関係を検証するための具体的な内容を示す．4 章で，実験方法と実験結果を示し，最後に考察とまとめを述べる．

2. サッカービデオゲームの概要

2.1. サッカービデオゲームの概要

(1) サッカービデオゲーム

サッカービデオゲームとは，サッカーをテーマとした二つのチームでの点の取り合いをコンピュータ上に表現したゲームソフトウェアである．図 1 に，本研究で用いるサッカーゲームの概観を示す．サッカーチームは 11 人で構成されるため，ゲームを行うには 22 人の人間が必要となる．一般的に 22 名全員が揃わず，不足分をコンピュータが操作するプレイヤーが補うことによりゲームが行われる．本稿では学習を行う都合上，22 人全てのプレイヤーがコンピュータによって操作される．

(2) サッカーゲームに接続されるコンピュータエージェント

コンピュータによって操作される 22 人のエージェントは，ゲームプレイヤーの代理を務めるための能力を所持する．具体的には，味方ないし敵プレイヤーの位置，ボールの位置，

図 1. サッカーゲームの概観

現在の自分の位置等の情報を環境から取得し，状況に合った適切な判断をエージェントが自ら下すことが出来る能力が求められる．以上の能力を実現するため，各エージェントには認識器，意思決定部，行動器が実装されている．

・認識器

エージェントは環境から表 1 に示す 19 種類の内容を認識器によって判断することが出来る．各認識内容をエージェントは YES か NO かで判断し，結果を 19 ビットのビット列に変換する．変換されたビット列は意思決定部に渡される．

・意思決定部

本研究におけるエージェントの意思決定部には，イベント駆動型ハイブリッドクラシファイア学習システムを採用している．イベント駆動型ハイブリッドクラシファイア学習システムの詳細については後節で記述する．

表 1. エージェントが認識できる環境状態

認識 No	認識内容
1	自分がチーム内でボールに最も近いか？
2	味方が攻撃中か？
3	自分の前にドリブルスペースがあるか？
4	視界 30M 内に，パスを出せる味方が存在するか？
5	視界 30M～60M 内に，パスを出せる味方が存在するか？
6	自分の周りに敵プレイヤーがいないか？
7	自分は敵陣方向を向いているか？
8	自分はボールをコントロール出来るか？
9	自分は敵陣地内にいるか？
10	自分はゴールライン際にいるか？
11	自分はサイドエリアにいるか？
12	ボールキーパーは敵陣にいるか？
13	ボールキーパーはゴールライン寄りにいるか？
14	ボールキーパーはサイドエリアにいるか？
15	自分は周りの味方と距離を保てているか？
16	自分はチーム内でボールに 2 番目に近いか？
17	自分は予想ボール停止地点に一番近いか？
18	ゴール方向にスペースがあるか？
19	敵チームにボールを長時間キープされているか？

・行動器

エージェントは意思決定部での決定に応じて表 2 に示す 8 種類の行動を起こすことが出来る．エージェントが行動を起こした結果，環境に変化が生じる．ゲームは以上の認識，意思決定，行動を 22 人のプレイヤーが順次行うことで進行する．

2.2. 従来研究

一般的な機械学習の方法として，Q-Learning[15]，ニューラルネットワーク，遺伝的アルゴリズム等を用いるものがある．クラシファイアシステムによる機械学習も一般的な学習方法のひとつである．本研究では，バケツリレーアルゴリズムを用いたクラシファイアシステムによる学習方法を採用している．クラシファイアシステムを用いる利点として，IF-THEN で構成されるルールへの適用が比較的容易なことがあげられる．

2.2.1. クラシファイアシステムによるエージェントの強化学習

2.2.1.1. クラシファイアシステム

クラシファイアシステムでは，クラシファイアと呼ばれるルールを扱う．ルールは IF(Condition)-THEN(Action)の形式で現されており，エキスパートシステムでしばしば用いられるプロダクションルールと同じ形をしている．クラシファイアシステムは，

1. 実行部（Performance Component）
2. 強化部（Reinforcement Component）
3. 発見部（Discovery Component）

表 2. エージェントが実行可能な行動の一覧

行動名	詳細
No.1　待機	一定時間，その場で待機する
No.2　パス	味方選手に対してボールを渡す．
No.3　ドリブル	自分の正面に向かってドリブルを行う．
No.4　シュート	敵ゴールに向けて強くボールをキックする．
No.5　ターン	自分が向いている方向を変化させる．
No.6　移動	移動する．移動目標は，ボールまたは基本守備位置の 2 種である．
No.7　チェック	ボールを所持する選手に向かっていき，プレッシャーをかけに行く．スペースを消しに行く，ボールを奪いに行く等の行動．
No.8　クリア	敵ゴール方向にボールを大きく蹴り出す．

の 3 つの要素から構成されるのが一般的である．本研究で扱うクラシファイアシステムについても上記の 3 要素を有する．

本研究で扱われている各エージェントは IF-THEN の形式で記述されたルールの集合を保持している．認識器から渡されたビット列は各エージェントの IF 節と照合され，マッチする IF 節を持つルールが全て選出される．各ルールには信頼度が設定されており，信頼度の値に比例した確率に依ってひとつのルールが選び出される．選び出したルールの THEN 節部分を行動器に渡す．エージェントは行動器を通して環境へ出力を行い，報酬もしくは罰を受ける．報酬ないし罰の値によってエージェントが保持するクラシファイアシステムの信頼度を増減させることにより，強化学習を実現する．

2.2.1.2. バケツリレーアルゴリズム

本システムでは，バケツリレーアルゴリズムを用いたクラシファイア間での信頼度の受け渡しを行っている．バケツリレーアルゴリズムを用いることの利点は，行動の流れを学習することが出来る点である．具体的には，「シュートを打つ」という行動が環境からの報酬によって信頼度が上昇されることにより，直前の行動である「ドリブルをする」「パスをする」などの行動も合わせて信頼度が上昇する仕組みである．バケツリレーアルゴリズムを用いることで，エージェントはより成果を挙げやすい行動の流れを学習することが出来る．図 2 にバケツリレーアルゴリズムの概要を示す．バケツリレーアルゴリズムの欠点は，十分な学習を完了するために多大な時間を要するという点である．

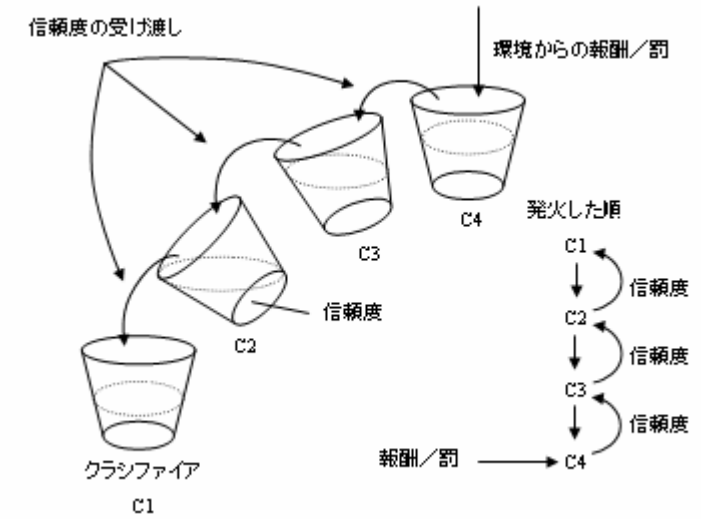

図 2. バケツリレーアルゴリズムの概要

2.2.2. イベント駆動型ハイブリッドクラシファイアシステム

ハイブリッド型意思決定とは，従来の人の手によって作成されるアルゴリズムと，クラシファイアシステムに代表される機械学習の仕組みを合わせたものを呼ぶ．具体的には，各エージェントに 2 種類のクラシファイアを保持させることでハイブリッド型クラシファイアシステムを実現する．1 種類目は，あらかじめ製作者が手作業で記述されたクラシファイアである（以下アルゴリズム部と記す）．2 種類目は，実際に各エージェントが試合を通して学習獲得するクラシファイアである．アルゴリズム部の特徴は，学習によって更新されないことと，学習によって得られるクラシファイアより優先して選択されることである．本研究で用いられるハイブリッドクラシファイアシステムのアルゴリズム部には，サッカープレイとして必要最低限なアルゴリズムが記述されている．エージェントにアルゴリズム部を保持させることにより，エージェントは学習開始当初からある程度のパフォーマンスを発揮することが可能になる．さらにアルゴリズム部で記述されている部分に関しては学習の必要がないため，探索領域を大幅に削減することが出来る．イベント依存型ハイブリッド型意思決定システムの構成図[12]を図 3 に示す．

本稿で扱うクラシファイアシステムはイベント分析機能を有する．イベント分析機能の具体的な機能とは，エージェントが実際に認識した回数の多いイベントに対応した行動を学習することが出来るしくみを指す．本サッカーエージェントが保持するクラシファイアの IF 節は，長さ 19 ビットのバイナリ列で表現されている．つまり，エージェントは 19 種類の情報を環境から取得することが出来るように設計されている．エージェントが保持するクラシファイアの IF 説には，各ビットについて 0，1，#（ドントケア）のいずれかをセットすることが出来るため，3^{19} 通りの IF 説を作成可能である．一方，実際に試合中に認識するイベントに対応するクラシファイアは非常に限定されたものになる．学習の過程において IF 節をランダムに生成する設計方法を採用した場合，実際に試合で使われることのない，不要なクラシファイアが生成されてしまう可能性が非常に高い．この問題を解決する

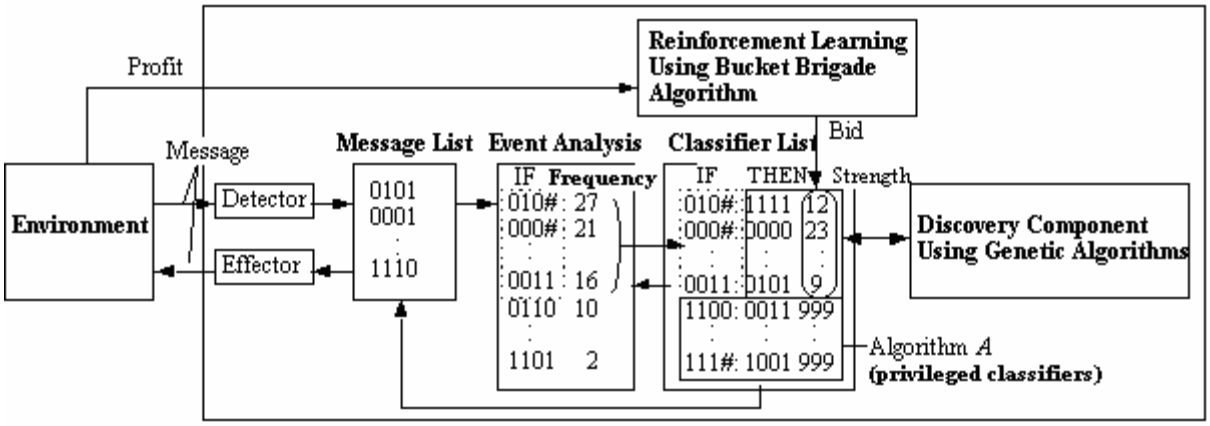

図 3. イベント依存型ハイブリッドクラシファイアシステムの構成図（文献[12] より引用）

ために，エージェントに多くのクラシファイアを持たせることで有効なクラシファイアを持つ可能性を高める手法が存在する．一方，エージェントがクラシファイアを多く抱えると，環境からの情報認識に要す時間が肥大化する．サッカーゲームはリアルタイムで進行するため，エージェントの処理時間が長くなるとゲームスピードの低下をまねくことに繋がり，クラシファイアの量によってはゲームが成り立たなくなる可能性もある．従って，一定時間内にエージェントの学習，行動の双方を完了させるためには学習獲得するクラシファイアの数を制限する必要があると言える．イベント分析機能は，実際に現れたイベントを，そのまま IF 節としてエージェントに適用する．先の研究において，この手法はクラシファイアの数に制限がある中で，エージェントに有効なルールを保持させることに有効的であることを示した[12].

2.2.3. 役割分担を考慮した報酬割り当て

ハイブリッド型意思決定による強化学習の弱点として，バケツリレーアルゴリズムを採用した学習方法に起因する，学習時間の肥大化がある．もうひとつの欠点として，強化学習に用いられる報酬値に決定指標が存在せず，報酬値は開発者の判断で決定せざるを得ないという問題がある．報酬値が学習に適さない値に設定されてしまうと，強化学習の効率が下がり学習時間のさらなる肥大化や学習の品質低下を招く原因となる．学習時間の肥大化の問題に対する解決案のひとつとして，我々は役割分担を考慮した報酬設定を行うことによる学習の効率化手法を提案した[13]．サッカーにおけるポジションの役割を考慮し，各エージェントの役割に即した行動の学習を促進させることを目標としている．エージェントに持たせる役割は，大まかにフォワード（FW），ミッドフィールダー（MF），ディフェンダー（DF），ゴールキーパー（GK）の 4 種類に大別した．基本的には，FW の役割を持ったエージェントは攻撃を成功させることにより大きな報酬値を得られるように設定した．MF の役割を持ったエージェントはボールの繋ぎ役と設定し，パス，ドリブル等といったボール運びの行動に対して大きな報酬値を得られるように設定した．DF の役割を持ったエージェントは，相手からボールを奪う等といった対戦相手の攻撃を防ぐ行動に対し，大きな報酬値を得られるように設定した．この手法を用いることにより，FW の役割を設定されたエージェントは，役割分担を考慮しない報酬設定を行う場合と比べてシュートの成功率，ドリブルの試行回数を上昇させた．同様に MF の役割を設定したエージェントについても，ドリブル，パスの成功率，試行回数の上昇に成功した．DF の役割を持ったエージェントに関しても，ボールカット数の上昇等の効果を得た．役割分担を考慮した報酬割り当ては，以上の結果と合わせて大きく分けて 2 つの効果を上げられることが示された．一つ目は学習収束の早期化が可能であること，二つ目はより勝率を高く保つことが出来るアルゴリズムの獲得が出来るようになることである．我々は上記 2 つの有効性から，エージェントの学習時間短縮に，役割分担を考慮した報酬設定を用いることが有効である可能性を示した．

2.2.3. 従来研究の考察

前節に記した 2 つの報告（イベント駆動型ハイブリッドクラシファイアシステムと役割分担を考慮した報酬設定）を通して，報酬値の変化が各エージェントの学習結果に大きく影響することが示された．具体的な検証の結果として，1）報酬値を変化させることに伴って，各エージェントが獲得する意思決定の内容が変化している可能性が示された．さらに，2）各エージェントが学習によって獲得する意思決定は，学習相手（対戦相手）のアルゴリズムに依って変化することが示された．また，3）各エージェントが獲得した意思決定は対戦相手に依存して変化していた一方，どの対戦相手でも定性的に学習の効果を上げることが出来た．以上の 3 点から，役割分担を考慮して報酬値の設定を行うことはエージェントの意思決定の内容に変化を及ぼすと同時に，サッカーチーム全体としての一連のアクション（チーム戦術）に影響を及ぼしている可能性があると考察された．

3. プレイヤーの強化学習とチームフォーメーションの関係

3.1. フォーメーションの概要

サッカーのフォーメーションとは，サッカーにおける選手の配置を意味する．システムという言葉も同じ意味で用いられる．11 人全員の配置隊形を示すこともあれば，特定の局面における数人の配置隊形を示すこともある．フォーメーションを議論する時には，しばしば「4-4-2」や「3-5-2」といった数字の羅列が使用される．これは，DF-MF-FW の順番でそれぞれのポジションの人数を表したものである．すなわち「4-4-2」だと，DF4 人，MF4 人，FW2 人という意味である．「3-5-2」では DF3 人，MF5 人，FW2 人という意味になる．「4-2-3-1」や「4-3-1-2」という表記がされる場合もある．これは，MF を守備的 MF と攻撃的 MF に分けて表記しているもので，「4-3-1-2」であれば DF4 人，守備的 MF3 人，攻撃的 MF1 人，FW2 人という意味である．

3.2. 現実世界におけるサッカーの戦術とフォーメーションの関係

現実世界におけるサッカーではさまざまな戦術が取り入れられている．大きく分けると，攻撃的，守備的，バランスを取るといった戦術が挙げられる．具体的な戦術として，サイド攻撃を行う，中央突破を狙う，カウンターを狙う等がある．戦術を実行するに際して，重要な要素となるのがチームのフォーメーションである．フォーメーションは，11 人のフィールドプレイヤーをどのような役割で，どのように配置するかを示し，サッカーにおける戦術の基本となる．例えば，守備を重視する戦術を採用する場合なら「DF の選手を多く配置する」「MF の選手を通常より後退させて配置する」等が有効であると言える．攻撃を重視する戦術を採用する場合ならば，「FW の選手を多く配置する」ことや，「攻撃的な MF を配置する」等が有効である．

上記から，現実世界のサッカーにおける戦術と，チームフォーメーションの間には密接な関係があると言える．一般的に，チームのフォーメーションは実行される戦術に依って決定される[16].

3.3. フォーメーションが学習に与える影響

先の研究において，各エージェントの学習に用いられる報酬値の設定が，チーム戦術に影響を及ぼしている可能性が示された．一方，学習によって獲得した意思決定とチームのフォーメーションの関連については考察を行っていないため，役割分担を考慮した報酬設定を適用することによるチーム戦術への影響の検証は不十分であると言える．既に 3.1.で述べたように，チームの戦術とチームフォーメーションの間には密接な関係があると考えるのが一般的である．そのため，チームフォーメーションの変化に伴い，エージェントが獲得する意思決定の内容も変化していく可能性が考えられる．またフォーメーションによって各ポジションが担う役割も変化するため，役割分担を考慮した報酬割り当てについて，

従来研究で用いた報酬値は効率の良い学習を行える設定ではなくなる可能性がある．エージェントの学習とチームフォーメーションに関する考察を深めるためには，様々なフォーメーションを適用してエージェントに学習を行わせ，学習効率やチーム戦術への影響について検証を行う必要があると考えられる．

3.3. フォーメーションが学習に与える影響の検証

様々なフォーメーションを適用して実験を行うことで，フォーメーションの変化がエージェントの学習に及ぼす影響について検証を行う．本研究では，各ポジションの人数が異なる 4 つのフォーメーション間で比較を行う．実験に用いるフォーメーションのプレイヤー配置を図 4 から図 7 に示す．何れも現実のサッカーにおいて一般的に用いられるフォーメーションで，FW の数，MF の数，DF の数にバリエーションの異なるものを選択した．

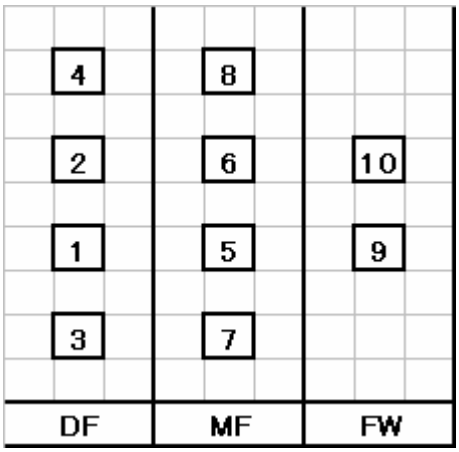

図 4. フォーメーション 1

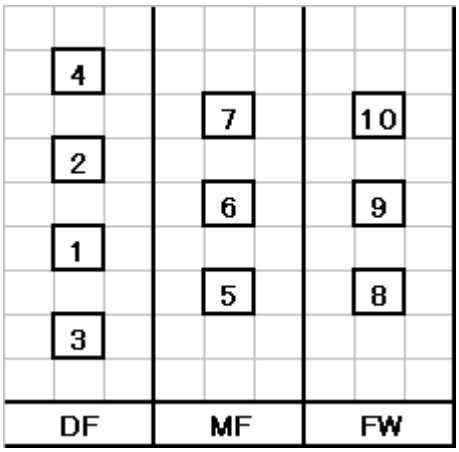

図 5. フォーメーション 2

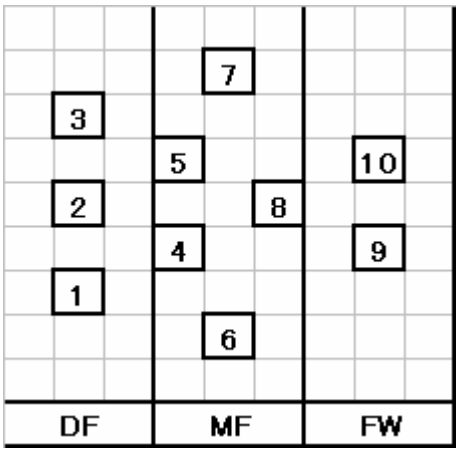

図 6. フォーメーション 3

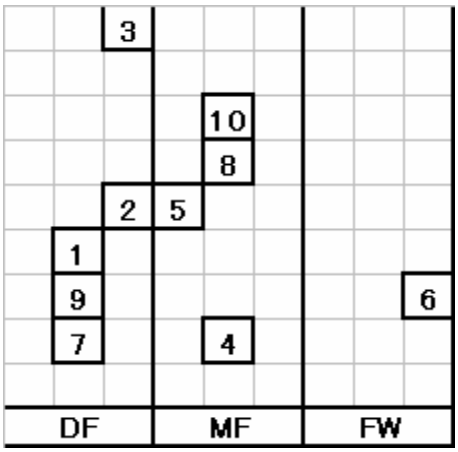

図 7. フォーメーション 4

また実験ではフォーメーションの違いが学習に及ぼす影響を顕著なものにするために，片方のチームのみ，エージェントの学習を適用する．以下本稿では，学習を行うチームを自チーム，学習を行わないチームを相手チームと定義し，フォーメーションの違いによる学習効率への影響とチーム戦術への影響を調査・検討する

(1) 学習効率への影響

フォーメーションを変更することによる，学習効率の違いを調査する．具体的には，各フォーメーションを適用して実験を行った際の，自チームの勝率の推移について比較を行うことで学習効率を検証する．また対戦相手のフォーメーションを変更した際の学習効率についても同様に検証し，フォーメーションの相性，戦術の相性について言及する．

(2) チーム戦術への影響

フォーメーションを変更することによる，各プレイヤーの行動内容の変化について調査する．具体的には，プレイヤーの各行動の頻度，成功数，成功回数を比較することでプレイ内容の変化を考察し，どのような意思決定が獲得されているかを検証する．またプレイの一連の流れとして現れるチーム戦術について，各プレイヤーのプレイ内容から考察を行う．

4. 評価実験

4.1. 実験方法

対戦相手として，先の研究[12-14]で扱ったコンピュータチームを用いる．自チームに学習を行うハイブリッド型意思決定システムを埋め込み，相手チームには学習を行わないアルゴリズム型の意思決定システムを埋め込む．相手チームのアルゴリズムには，アルゴリズム A，アルゴリズム B，アルゴリズム C の三種を採用する（各アルゴリズムの概要については後述）．また，フォーメーションについては，それぞれ異なる特徴を持つ 4 種を採用した．両者に各種フォーメーションを適用して 60 回対戦を行い，勝敗を観測する．また同時に各プレイヤーの各プレイの頻度，成功回数，成功率について観測する．

4.1.1. 相手チームに適用されるアルゴリズム

(1) アルゴリズム A

他のアルゴリズムの基準となるアルゴリズムである．本研究で用いられているハイブリッド型クラシファイアシステムの，アルゴリズム部として採用されている．バランス良く攻撃と守備を行う．

(2) アルゴリズム B

アルゴリズム A に比べて攻撃的なアルゴリズムである．フィールドのサイドに位置するプレイヤーを中心に攻撃を展開する．

(3) アルゴリズム C

アルゴリズム A に比べて守備的なアルゴリズムである．

4.1.2. 適用されるフォーメーション

実験は自チームに図 4 から図 7 に示したフォーメーションを適用して対戦を行い，勝率と各エージェントのアクションの成功回数，試行回数について観測を行うことで成される．フォーメーション 4 に関しては遺伝的アルゴリズム（以下 GA）を用いたフォーメーションの探索を行って作成したものであるため，他のフォーメーションに比べて整然としていない．また，以下に各フォーメーションの一般的な特徴を示す[16].

(1) フォーメーション 1

システムは 4-4-2．4-4-2 は 4 バックを主体とした中で最も普遍的で，DF が 4 人，MF が 4 人，FW が 2 人というフォーメーションである．MF の 4 人はセントラル MF が 2 人，サイド MF が 2 人という構成が一般的である．サイド MF はサイドを攻略するだけでなく，攻撃

の組み立てに貢献することも期待されている．サイドバックが中盤を追い越すプレーをする事が戦術として組み込まれていることが多い．理論上フィールドプレイヤーが満遍なくピッチをカバーできるため，このフォーメーションは 4 バックの基本フォーメーションとも言える．

(2) フォーメーション 2

システムは 4-3-3．4-3-3 は DF が 4 人，MF が 3 人，FW が 3 人のフォーメーションである．4-4-2 のフォーメーションと比べて FW が一人多く，攻撃を重視するフォーメーションであると言える．

(3) フォーメーション 3

システムは 3-5-2．3-5-2 は DF が 3 人，MF が 5 人，FW が 2 人というフォーメーションである．一般的な MF の配置方法として，バランス重視型，攻撃重視型の 2 種類が用いられる．バランス重視型の場合はフィールドの左右に MF を 1 人ずつ，守備的 MF が 2 人，FW の後ろに攻撃的 MF を 1 人置く配置（3-4-1-2）になる．攻撃重視型の場合，守備的 MF が 1 人，FW の後ろに攻撃的 MF を 2 人置く配置になる（3-1-4-2）．本研究では，バランス重視型の配置に則ったものを採用している．

(4) フォーメーション 4

システムは 5-4-1．DF が 5 人，MF が 4 人，FW が 1 人というフォーメーションである．FW の数が一人であるため，他のフォーメーションに比べて攻撃力が低い．そのため，攻撃に際しては中盤のプレイヤーないし DF のプレイヤーが攻撃参加するのが一般的である．一方，MF，DF のプレイヤー数が多く，守備に多くの人数を使うことが出来るという利点がある．

また，フォーメーション 4 は遺伝的アルゴリズムを用いて作成されている．遺伝的アルゴリズムを用いたフォーメーションの探索の詳細については，巻末に記載する．

4.2. 実験結果

4.2.1. 学習効率への影響

4.2.1.1. 対アルゴリズム A

図 8, 図 9 に, 対戦相手にアルゴリズム A を適用した際に得られた実験結果を示す. 図 8 は, 相手チームにフォーメーション 1 を適用した際に得られた実験結果である. 図 8 より, 自チームにフォーメーション 2 を適用した場合に，他のフォーメーションを適用したチームに比べて勝率が早期に高い値を示し，また安定して高い値を保っていることが分かる．

一方，フォーメーション 3 を適用したチームは勝率の立ち上がりが遅く，収束先も低い値になっている．

図 9 は, 相手チームにフォーメーション 2 を適用した際に得られた実験結果である. 図 9 より，自チームにフォーメーション 4 を設定した際に，早期に高い勝率を示し，また安定して高い値を保っていることが分かる．一方，自チームにフォーメーション 1 を適用した場合には勝率の立ち上がりが遅いという結果が得られた．

4.2.1.2. 対アルゴリズム B

図 10，図 11 に，対戦相手にアルゴリズム B を適用した際に得られた実験結果を示す．図 10 は, 相手チームにフォーメーション 1 を適用した際に得られた実験結果である. 図 10 より，自チームにフォーメーション 4 を適用した場合に，他のフォーメーションを適用した場合に比べて勝率が高い値に収束していることが分かる．一方，自チームにフォーメーション 2 を適用した場合には勝率の立ち上がりが遅くなり，低い値に収束しているという結果を得た．フォーメーション 1，フォーメーション 3 を自チームに適用した場合，勝率の立ち上がりに関してはほぼ同様の結果を得た．また，フォーメーション 1 を適用した場合のほうがフォーメーション 3 を適用した場合に比べて 5%ほど高い値に収束していることが分かる．

図 11 は，相手チームにフォーメーション 2 を適用した際に得られた実験結果である．図 11 より，フォーメーション 2 ないしフォーメーション 4 を自チームに適用した際に勝率が高い値に収束していることが分かる．一方，フォーメーション 3 を適用した場合には，他のフォーメーションを適用した場合と比べて勝率が低い値に収束している．フォーメーション 1 を適用した場合の特徴は, 比較的早い段階で勝率が高い値を示していることである．

4.2.1.3. 対アルゴリズム C

図 12，図 13 に，対戦相手にアルゴリズム C を適用した際に得られた実験結果を示す．図 12 は, 相手チームにフォーメーション 1 を適用した際に得られた実験結果である. 図 12 に表れている特徴として, 上記に示した図 8 から図 11 と比べて勝率の収束先に大きな差が生じていることが分かる．また，対戦開始当初の数試合について，どのフォーメーションを適用した際にも勝率が非常に低い値を示していることが分かる．具体的な実験結果について，自チームにフォーメーション 2 を適用した際に，勝率が最も高い値に収束していることが分かる．一方，フォーメーション 4 を適用した場合は最も低い値に勝率が収束している．また，フォーメーション 1，フォーメーション 3 を適用した場合には，勝率はほぼ同じ値(約 60%)に収束している．一方，フォーメーション 3 のほうが早期に高い勝率を示していることが分かる．どのフォーメーションを適用した場合にも表れる共通の特徴として，早い段階で勝率が収束しているという点が挙げられる．

図 13 は，相手チームにフォーメーション 2 を適用した際に得られた実験結果である．図

13 より，自チームにフォーメーション 4 を適用した場合に，勝率が最も高い値に収束していることが分かる．一方，自チームにフォーメーション 2 を適用した場合は勝率の立ち上がりが遅く，最も低い値に収束している．フォーメーション 1，フォーメーション 3 を適用した場合には，早い段階で一度高い勝率を示す一方，フォーメーション 4 を適用した場合に得られた結果よりも低い値に収束していることが分かる．また，フォーメーション 1 とフォーメーション 3 の勝率を比較した場合，フォーメーション 1 を適用した場合のほうが 5%ほど高い値に収束していることが分かる．

4.2.2. チーム戦術への影響

チーム戦術に関する実験結果として，自チームに各フォーメーションを適用して実験を行った際に得られた実験結果の中から，(1)シュート成功回数，(2)シュート試行回数，(3)パス成功回数，(4)パス試行回数，(5)ドリブル成功回数，(6)ドリブル試行回数，(7)ボール奪取回数，の 7 つの要素に着目した．以下に対戦相手にアルゴリズム A，B，C を適用し，かつフォーメーション 1，2 をそれぞれ適用した際に得られた実験結果を示す．得られたデータは以下の手順によってそれぞれのアクション別に 5 段階評価に写像される．

(シュート試行回数の例)

1. シュート試行回数に関する全実験データの中から，最大の値を選び出す（max 値）
2. シュート試行回数に関する全実験データの中から，最少の値を選び出す（min 値）
3. シュート試行回数に関する，任意のデータを取り出す（data 値）．data 値の評価は以下の(1)式によって導出される．

$$Fitness(data) = \frac{4(data - \min)}{\max - \min} + 1 \qquad (1)$$

他のアクションについても，同様にして評価点を算出する．

4.2.2.1. 対アルゴリズム A

図 14 に，対戦相手のアルゴリズムにアルゴリズム A を適用し，フォーメーションにフォーメーション 1 を適用して実験を行った際に得られた実験結果を示す．図 14 より，自チームにフォーメーション 1，もしくは 3 を適用した場合に，ボール奪取回数について他のフォーメーションを適用した場合よりも大きい値を示していることが分かる．またフォーメーション 3 を適用した場合にはパス試行回数が他のフォーメーションを適用した場合に比べて大きくなることが分かる．一方，パス成功回数についてはパス試行回数と比較して小さい値になっていると言える．自チームにフォーメーション 4 を適用した際には，ドリブル成功回数，ドリブル試行回数が他のフォーメーションを適用した場合に比べて大きい値

を示していることが分かる．またシュートの成功回数についてもフォーメーション 4 を適用した場合に比較的大きい値を示している．一方，ボール奪取回数については，自チームにフォーメーション 4 を適用すると他のフォーメーションを適用した場合に比べて小さい値を示していることが分かる．またフォーメーション 4 を自チームに適用した場合，パス試行回数も少なくなっていることが分かる．自チームにフォーメーション 2 を適用した場合は，シュート成功回数，シュート試行回数ついて比較的小さい値になるという結果を得た．

図 15 に，対戦相手のアルゴリズムにアルゴリズム A を適用し，フォーメーションにフォーメーション 2 を適用して実験を行った際に得られた実験結果を示す．図 15 より，自チームにフォーメーション 4 を適用した際に，シュート成功回数，ドリブル成功回数，ドリブル試行回数の 3 要素について，自チームに他のフォーメーションを適用した場合よりも大きな値を示していることが分かる．一方，自チームにフォーメーション 4 を適用した場合にはボール奪取回数が最低となった．またパス試行回数についても同様で，自チームにフォーメーション 4 を適用した場合，他のフォーメーションを適用した際に比べて低い値を示している．自チームにフォーメーション 3 を適用した場合，シュート試行回数，パス試行回数，ボール奪取回数について比較的大きな値を示していることが分かる．なお，ボール奪取回数については，自チームにフォーメーション 1 を適用した場合とほぼ同数となった．自チームにフォーメーション 2 を適用した場合は，シュート成功回数，シュート試行回数について低い値を示していることが分かる．またドリブル試行回数，成功回数について，自チームにフォーメーション 4 を適用した場合のみ，大きい値になることが示された．

4.2.2.2. 対アルゴリズム B

図 16 に，対戦相手のアルゴリズムにアルゴリズム B を，フォーメーションにフォーメーション 1 を適用して実験を行った際に得られた実験結果を示す．図 16 より，自チームにフォーメーション 4 を適用すると，他のフォーメーションを適用した場合に比べてシュート試行回数，ドリブル成功回数，ドリブル試行回数の 3 要素が大きな値を示すことが分かる．パス成功回数，パス試行回数については大きな差は表れなかった．ドリブルについては，フォーメーション 4 を除いた 3 つのフォーメーション間に大きな差は表れなかった．ボール奪取回数については，フォーメーション 2 とフォーメーション 4 を自チームに適用した際にほぼ同数の値が得られた．一方，フォーメーション 3 を適用した場合にはボール奪取回数が著しく低くなることが示された．

図 17 に，対戦相手のアルゴリズムにアルゴリズム B を，フォーメーションにフォーメーション 2 を適用して実験を行った際に得られた実験結果を示す．図 17 より，シュート成功回数，シュート試行回数について自チームにフォーメーション 2 を適用した場合に，他の

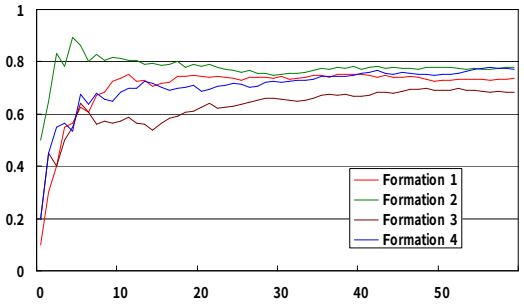

図 8. 各フォーメーションを適用した際の勝率の推移（対戦相手: フォーメーション 1, アルゴリズム A）

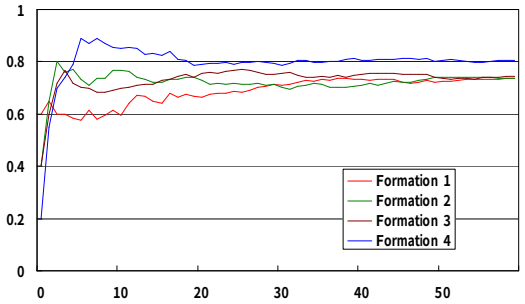

図 9. 各フォーメーションを適用した際の勝率の推移（対戦相手: フォーメーション 2, アルゴリズム A）

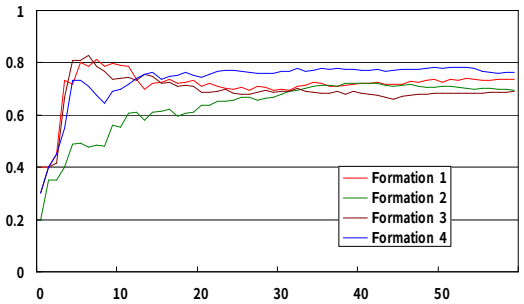

図 10. 各フォーメーションを適用した際の勝率の推移（対戦相手: フォーメーション 1, アルゴリズム B）

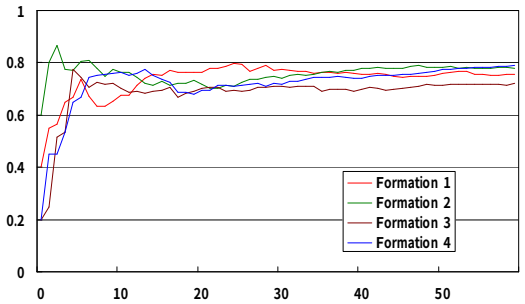

図 11. 各フォーメーションを適用した際の勝率の推移（対戦相手: フォーメーション 2, アルゴリズム B）

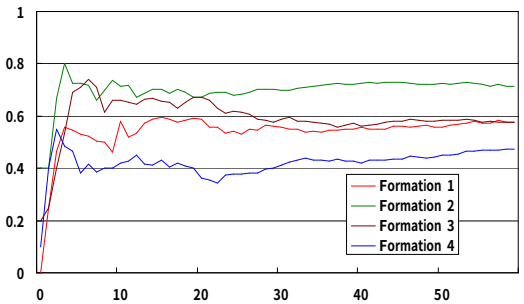

図 12. 各フォーメーションを適用した際の勝率の推移（対戦相手: フォーメーション 1, アルゴリズム C）

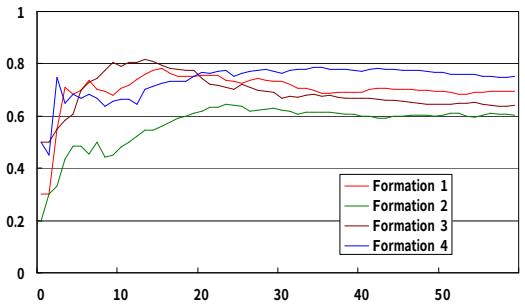

図 13. 各フォーメーションを適用した際の勝率の推移（対戦相手: フォーメーション 2, アルゴリズム C）

フォーメーションに比べて小さい値を示していることが分かる．他の 3 種のフォーメーションを適用した場合には，シュート成功回数，シュート試行回数，共にほぼ同数を示した．パス試行回数については，どのフォーメーションを適用した場合もほぼ同数を示していることが分かる．パス試行回数について，自チームにフォーメーション 3 を適用すると比較的大きな値を示すことが分かる．一方，フォーメーション 4 を適用した場合にはパス試行回数が大きく減少していることが分かる．ドリブル試行回数，ドリブル成功回数に関しては，フォーメーション 4 を自チームに適用した際に比較的大きな値を示している．次いでフォーメーション 1 を適用した場合に大きい値を示し，フォーメーション 2，フォーメーション 3 を適用した場合は比較的小さい値を示した．ボール奪取回数については，自チームにフォーメーション 1，もしくはフォーメーション 3 を適用した場合に比較的大きな値を示していることが分かる．一方，自チームにフォーメーション 4 を適用すると非常に小さい値になることが見て取れる．

4.2.2.3. 対アルゴリズム C

図 18 に，対戦相手のアルゴリズムにアルゴリズム C を，フォーメーションにフォーメーション 1 を適用した際に得られた実験結果を示す．図 18 より，自チームにフォーメーション 2 を適用した場合に，シュート試行回数，パス成功回数，パス試行回数の 3 要素について高い値を示すことが分かる．ボール奪取回数については，自チームにフォーメーション 3 を適用した場合に高い値を示している．フォーメーション 3 を適用した際に表れるもう一つの特徴として，パス成功回数が他のフォーメーションを適用した場合に比べて低い値を示しているという点が挙げられる．また，どのフォーメーションを適用した場合にも表れる特徴として，シュート成功回数，ドリブル成功回数，ドリブル試行回数の 3 要素の値が低くなるという点が挙げられる．

図 19 に，対戦相手のアルゴリズムにアルゴリズム C を，フォーメーションにフォーメーション 2 を適用した際に得られた実験結果を示す．図 19 より，自チームにフォーメーション 1，もしくはフォーメーション 3 を適用した際に，シュート試行回数の値が非常に大きくなっていることが分かる．シュート試行回数については次いでフォーメーション 3 を適用した場合に大きい値を示している．フォーメーション 2 を適用した場合には小さい値を示しており，他の 3 者を適用した場合に比べて大きな差が出ていることが分かる．シュート成功回数については，どのフォーメーションを適用した場合にも大きな差は表れなかった．パス試行回数については，フォーメーション 2，もしくはフォーメーション 3 を適用した際に大きな値を示している．次いでフォーメーション 1 を適用した場合に大きい値を示している．フォーメーション 4 を適用した場合は，他のフォーメーションを適用した場合に比べると，非常に小さい値を示していることが分かる．パス成功回数については，どのフォーメーションを適用しても大きな差は表れていないと言える．ドリブルに関しては成功回数，試行回数，共にフォーメーション 4 を適用した場合に大きな値を示していることが分

かる．一方，フォーメーション 1，フォーメーション 2，フォーメーション 3 を適用した場合には非常に小さい値を示している．ボール奪取回数については，フォーメーション 1，もしくはフォーメーション 2 を適用した場合に非常に大きい値を示している．次いでフォーメーション 2 を適用した場合に大きい値を示している．一方，フォーメーション 4 を適用した場合には著しくボール奪取回数が減少していることが分かる．

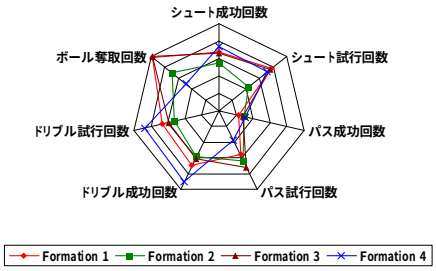

図 14. 各フォーメーションを適用した際の各アクション数の比較（対戦相手: フォーメーション 1，アルゴリズム A）

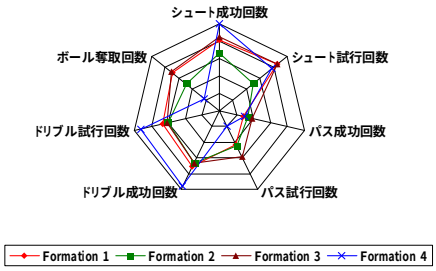

図 15. 各フォーメーションを適用した際の各アクション数の比較（対戦相手: フォーメーション 2，アルゴリズム A ）

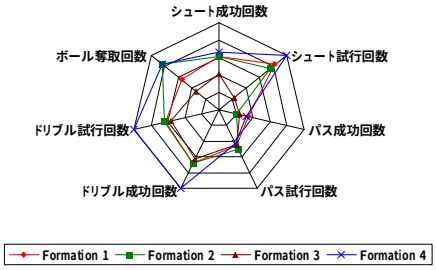

図 16. 各フォーメーションを適用した際の各アクション数の比較（対戦相手: フォーメーション 1，アルゴリズム B）

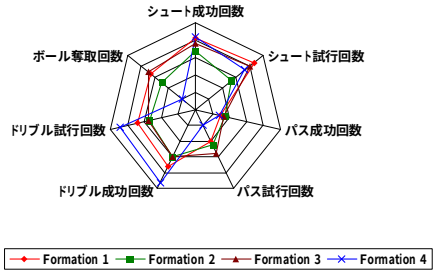

図 17. 各フォーメーションを適用した際の各アクション数の比較（対戦相手: フォーメーション 2，アルゴリズム B）

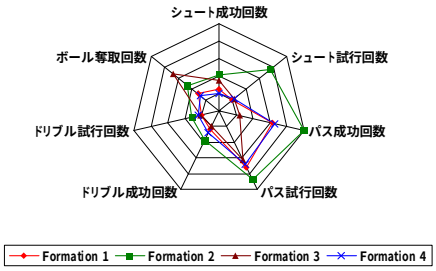

図 18. 各フォーメーションを適用した際の各アクション数の比較（対戦相手: フォーメーション 1，アルゴリズム C）

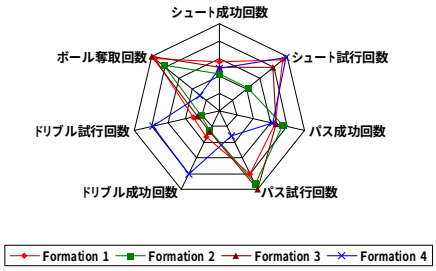

図 19. 各フォーメーションを適用した際の各アクション数の比較（対戦相手: フォーメーション 2，アルゴリズム C）

5. 考察

5.1. 学習効率への影響

(1) 自チームのフォーメーションを変更することによる学習効率への影響

図 8，図 9 から，自チームに設定したフォーメーションによって，同じ対戦相手でも勝率の立ち上がり，収束先に差が生じていることが分かる．特に図 8 の結果から，(1)自チームにフォーメーション 2 を適用した場合には勝率の立ち上がりが早く，収束先も高いこと，(2)自チームにフォーメーション 3 を適用した場合には勝率の立ち上がりが遅く，収束先も遅い，という 2 点は，自チームに適用したフォーメーションに依存した学習成果の差を顕著に表していると言える．次に，図 8 と図 9 を比較して考察する．図 8 では自チームにフォーメーション 2 を，図 9 では自チームにフォーメーション 4 を設定した際に勝率が高く保たれている．また，図 8 ではフォーメーション 3，図 9 ではフォーメーション 1 を自チームに適用した場合に勝率の立ち上がりが遅く，また低い値に収束することを示している．このことから，自チームのフォーメーションだけでなく，対戦相手のフォーメーションによっても自チームの学習効率が左右されているということが分かる．以上から，自チームに適用されるフォーメーション，また対戦相手に適用されるフォーメーションの組み合わせによって，自チームの学習効率が上がるケースと下がるケースが存在すると考えられる．まず，自チームに適用されるフォーメーションに依る学習効率の変化について考察する．エージェントの学習は，エージェントが環境から報酬値を受け取ることによってなされる．つまり自チームのエージェントの学習効率は，各プレイヤーの行動成功率に依存すると考えられる．つまり，学習効率が悪くなる理由は（勝率の立ち上がりが遅い，ないし低い値に収束先する），エージェントが学習開始当初に，環境から報酬よりむしろ罰を受けている回数が多くなっていることが原因であると考えられる．エージェントが学習開始当初に得ることの出来る報酬の大小は，エージェントにあらかじめ組み込まれているアルゴリズム部（非学習部）の内容に依る影響が大きいと言える．この点から，アルゴリズム部と，設定されたフォーメーションの相性が学習効率に大きな影響を与えていると考えられる．言い換えると，あらかじめ組み込まれたアルゴリズムによって起こされる戦術を行うに当たって，設定されたフォーメーションが適しているかどうかが問題になると言える．具体的な例を挙げると，攻撃的なアルゴリズム部を有するエージェントに守備的なフォーメーションを適用した場合，学習の効率が悪くなる可能性があると考えられる．

(2) 対戦相手のフォーメーションを変更することによる自チームの学習効率への影響

次に，対戦相手のフォーメーションが変化することに依る自チームの学習効率の変化について考察する．図 8 と図 9 の両図におけるフォーメーション 4 の実験結果を比較すると，図 8 では立ち上がりが遅く，収束先はフォーメーション 2 を自チームに適用した際に得ら

れる実験結果とほぼ同じ値である一方，図 9 では非常に立ち上がりが早く，収束先も他のフォーメーションを適用した場合に比べて高い値を示している．また，図 8 中では自チームにフォーメーション 2 を適用した場合に高い学習効率を挙げることが出来ることが示されている一方，図 9 中でのフォーメーション 2 の実験結果は立ち上がり，収束先，共にフォーメーション 4 よりも低い値を示している．以上の点は，対戦相手のフォーメーションによって，学習に適する自チームのフォーメーションが変化していることを表していると言える．つまり，フォーメーション同士には相性が存在していると考えられる．自チーム，相手チーム双方のプレイヤーの配置によって，「MF のプレイヤーから FW のプレイヤーに向けてのパスが通りやすい」「FW のプレイヤーがドリブルしやすい」等の状況が発生しているため，エージェントの学習にも影響が現れていると考えられる．

(3) 対戦相手のアルゴリズムを変更することによる自チームの学習効率への影響

続いて，対戦相手のアルゴリズムの違いがエージェントの学習に及ぼす影響について考察する．最初に，図 8，図 10，図 12 に示されている，対戦相手のフォーメーション 1 を適用した際に得られた実験結果について比較し，考察を行う．3 図より，対戦相手にアルゴリズム A，もしくはアルゴリズム C を適用した場合，自チームにフォーメーション 2 を適用すると他のフォーメーションを適用した場合に比べて学習効率が高くなることが分かる．一方，対戦相手にアルゴリズム B を適用した場合は，自チームにフォーメーション 4 を適用すると学習効率が高くなることが分かる．また，対戦相手のアルゴリズムをアルゴリズム C とした場合，自チームにフォーメーション 4 を適用すると自チームの学習効率が非常に低くなることが図 12 に示されている．これらの要素から，対戦相手のアルゴリズムと自チームに適用されるフォーメーションが関係し合って，自チームの学習効率に影響を及ぼしていることが分かる．つまり，対戦相手の戦術と自チームのフォーメーションの組み合わせによって，エージェントの学習が機能しなくなる可能性があることを示していると考えられる．

以上 3 点の考察より，エージェントの学習効率と自チームないし対戦相手チームのフォーメーションは，大きく関係しあっていることが示された．以上の考察から，サッカーゲームにおけるエージェントの学習効率向上のためには，チームのフォーメーションについて考慮することが一つの有効な手段であると考えられる．

5.2. チーム戦術への影響

（1）自チームのフォーメーションを変更することによるチーム戦術への影響

図 14 より，自チームにフォーメーション 2 を適用した場合に，他のフォーメーションを自チームに適用した場合に比べてシュート試行回数が少なくなっていることが分かる．一

方，本研究で採用した 4 つのフォーメーションのうち，FW の数が最も多い設定（3 人）となっているのがフォーメーション 2 である．つまり，チームのシュート試行回数は，チームの FW の人数に依存していないということが分かる．また同図より，フォーメーション 4 を適用したチームは，他のフォーメーションを適用したチームに比べてドリブル成功回数，ドリブル試行回数が多いことが分かる．つまり，フォーメーション 4 を適用したチームは他の 3 種のフォーメーションを適用したチームに比べ，ドリブルを中心とした戦術を獲得していると考えられる．同様にフォーメーション 3 を適用したチームについては，ボール奪取を中心とした守備的な戦術を獲得していると考えられる．また，フォーメーション 3 を適用したチームはパス試行回数が高くなっている一方，パス成功回数が最も低くなっていることが分かる．この結果から，フォーメーション 3 を適用したチームはパス成功率が低いと言える．フォーメーション 3 は他の 3 種のフォーメーションに比べて MF の人数が多く設定されている（5 人）．つまり，チームのパス活用度は MF の数に依存していないことが分かる．以上より，チーム戦術としての攻撃性は，FW のプレイヤーの数では決まらないことが分かる．同様に，MF のプレイヤーの数の増加に伴って，パスを中心としたチーム戦術が表れるという傾向も表れていないことが分かる．

（2）対戦相手のアルゴリズムを変更することによるチーム戦術への影響

次に，図 16 に着目する．図 16 より，フォーメーション 4 を適用したチームがドリブル成功回数，ドリブル試行回数について他のフォーメーションを適用したチームよりも高い値を示していることが分かる．フォーメーション 4 を適用したチームがドリブルを有効に活用していると考えられる要素は，図 14 において示された結果と同様の傾向が表れていることが分かる．続いて図 18 に着目する．先に示した図 14，図 16 に表れている傾向と異なり，フォーメーション 4 を適用したチームのドリブル成功回数，試行回数が非常に小さい値になっていることが分かる．一方，図 18 ではフォーメーション 2 を適用したチームがパス成功回数，パス試行回数について非常に大きな値を示している．勝率に関する実験結果（図 8，図 10，図 12）は，フォーメーション 4 を適用したチームは対戦相手のアルゴリズムがアルゴリズム A，もしくはアルゴリズム B のときに高い勝率を上げることが出来ることを示している．勝率に関する実験結果とチーム内行動の内訳を合わせて考察すると，フォーメーション 4 を適用したチームはドリブル成功回数，ドリブル試行回数が高い場合に同時に勝率が高くなる傾向があると考えられる．この考察を裏付けるため，対戦相手にフォーメーション 2 を適用した際に得られた実験結果の傾向に着目する．フォーメーション 4 を適用したチームのドリブル成功回数，ドリブル試行回数は図 15，図 17，図 18 に表れているように，対戦相手のフォーメーションにフォーメーション 2 を適用した場合にはいずれも高い値を示している．前節において，対戦相手のフォーメーションにフォーメーション 2 を適用した場合，自チームにフォーメーション 4 を適用すると勝率が高くなることが示されている．つまり，いずれの実験結果においても，「自チームにフォーメーション

4 を適用し，かつドリブル成功回数，ドリブル試行回数が高い場合，自チームの勝率が高くなる」という傾向が表れていることが分かる．この傾向は，「フォーメーション 4 はドリブルを中心とした戦術を実行しやすいフォーメーションである」ことを意味していると考えられる．上記と同様の傾向が，フォーメーション 2 を適用したチームにも表れている．図 18 に示されている実験結果では，フォーメーション 2 はパス成功回数，パス試行回数が非常に高くなっている．図 12 に示されている勝率に関する実験結果と合わせて考察すると，「フォーメーション 2 は，パスを中心とした戦術を組み立てやすいフォーメーションである」と言える．

上記の考察は，対戦相手の戦術ないしフォーメーションの特徴を露見させることが出来る可能性を示していると言える．具体的な例を以下に示す．図 14，図 16 より，フォーメーション 4 を適用したチームはドリブルを有効に活用する戦術を学習する可能性があることを既に示した．一方，図 18 に示されている実験結果では，フォーメーション 4 を適用したチームはドリブルを有効に活用出来ていないと言える．さらに，図 18 におけるフォーメーション 4 に関する実験結果では，図 14，図 16 と比較してパス成功回数，パス試行回数が比較的高くなっていることが分かる．これらの要素から，図 18 における対戦相手はドリブルによる攻撃は有効でなく，一方でパスによる攻撃が有効である，といった傾向があることが考えられる．また，図 18 におけるフォーメーション 2 に関する実験結果がこの傾向を裏付けていると言える．

（3）対戦相手のフォーメーションを変更することによるチーム戦術への影響

図 18 と図 19 を比較して考察する．図 18 における対戦相手には，パスを中心とした戦術を用いることが有効である可能性があることを既に示した．一方，図 13 ではフォーメーション 4 を適用したチームが高い勝率を収めている．また図 19 より，フォーメーション 4 を適用したチームはドリブルを有効に活用する戦術を行っていること分かる．同時に，フォーメーション 2 を適用したチームはパスを有効に活用する戦術を行っている一方，勝率は低い値に収束している．つまり，対戦相手のアルゴリズムに変化がない場合でも，フォーメーションの変化によって有効な戦術が変化していることが分かる．また同時に，フォーメーションの変化によってチームの特徴が変化していると言える．具体的には，フォーメーション 1 を適用した場合にはドリブルを中心とした攻めに強く，パスを中心とした攻めに弱いという特徴を持つ．フォーメーション 2 を適用した場合には，ドリブルを中心とした攻めに弱く，パスを中心とした攻めに強いという特徴が表れていると言える．

上記 3 点の考察より，自チームないし対戦相手のフォーメーション，対戦相手のアルゴリズムと学習によって獲得されるチーム戦術は，大きく関係しあっていることが示された．以上の考察から，(1)フォーメーションによって実行しやすい戦術が存在すること，(2)アルゴリズムとフォーメーションの組み合わせによって，有効な戦術が変化すること，の 2 点

を示した．以上の考察より，エージェントの学習効率向上のためにはチームのフォーメーションについて考慮することが有効な手段であることを示すと同時に，本研究で用いられているエージェントの強化学習の手法は，作成されたアルゴリズムの強い部分，ないし欠点を探し出すことに有効である可能性を示した．

6. まとめ

本稿では，サッカーゲームにおけるチームフォーメーションが，エージェントの学習に及ぼす影響について調査・検討した結果を報告した．また，各プレイヤーの意思決定について分析し，チームの一連の行動となって現れる戦術にフォーメーションが与える影響に着目して考察した．結果として，フォーメーションの設定はエージェントの学習に大きな影響を及ぼすことを示し，合わせて自チームと対戦相手フォーメーションの組み合わせによって学習効率が変化することを示した．また設定されたフォーメーションによって，エージェントが獲得する意思決定の内容が変化することを示した．以上より，サッカーエージェントの学習効率の向上には，フォーメーションの設定を考慮することが有効である可能性を示した．

参考文献

[1] Holland, J. H. : Escaping brittleness, *Machine Learning, an artificial intelligence approach*, Vol. II, R. S. Michalski, J. G. Carbonell and T. M. Mitchell (eds.), pp. 593-623, Morgan Kaufmann (1986).

[2] Belew. R. K and Gherrity. M, “Back Propagation for the Classifier System” In *Proceedings of the Third International Conference on Genetic Algorithms*. Morgan Kaufmann Publishers, CA, 1989, pp. 275-281

[3] J. H. Holland, “Escaping brittleness: The possibilities of general-purpose learning algorithms applied to parallel rule-based systems”, In Michalski, R. S. et al. (eds.): *Machine Learning II*, Morgan Kaufmann Publishers, CA, 1986, pp. 593-623

[4] R. L. Riolo, “Bucket brigade performance: I. Long sequences of classifiers, genetic algorithms and their application”, In *Proceedings of the Second International Conference on Genetic Algorithms*. Lawrence Erlbaum Associates, Publishers, 1987, pp. 184-195.

[5] R. L. Riolo, “Bucket brigade performance: II. Default hierarchies”, In *Proceedings of the Second International Conference on Genetic Algorithms*. Lawrence Erlbaum Associates, Publishers, 1987, pp. 196-201.

[6] R. L. Riolo, “The emergence of coupled sequences of classifiers”, In *Proceedings of the Third International Conference on Genetic Algorithms*. Morgan Kaufumann Publishers, CA, 1989, pp. 256-263

[7] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, MA, 1989.

[8] Holland. J. H, “Adaptation in Natural and Artificial System”. The MIT Press, Cambridge, MA, 1992. The University of Michigan Press, Ann Arbor, 1975.

[9] J. H. Holmes, P. L. Lanzi, W. Stolzmann, S. W. Wilson, ”Learning Classifier Systems: New Models, Successful Applications”, *Information Processing Letters*, Vol. 82, 2002, pp. 23-30.

[10] C-H. Huang, and C-T. Sun, “Parameter Adaptation within Co-adaptive Learning Classifier Systems”, In *Proceedings of the Sixth Annual Genetic and Evolutionary Computation Conference*. Vol. 2, LNCS3103, Springer-Verlag, Berlin, Heidelberg, 2004, pp. 774-784.

[11] T. Kovacs, “What Should a Classifier System Learn and How Should We Measure It?” *Journal of Soft Computing*, Vol. 6, No. 3-4, 2002, pp. 171-182.

[12] Y. Sato and R. Kanno, “Event-driven Hybrid Learning Classifier Systems for Online Soccer Games”, In *Proceedings of the 2005 IEEE Congress on Evolutionary Computation*, IEEE Press, pp. 2091-2098 (2005).

[13] Y. Sato, Y. Akatsuka, and T. Nishizono, “Reward Allotment in an Event-driven Hybrid Learning Classifier System for Online Soccer Games”, In *Proceedings of the 2006 Genetic and Evolutionary Computation Conference*, ACM Press, pp. 1753-1760 (2006).

[14] Y. Akatsuka, and Y. Sato, “Reward Allotment Considered Roles for Learning Classifier System for Soccer Video Games”, In *Proceedings of the 2007 IEEE Symposium on Computational Intelligence and Game (2007)*.

[15] R. S. Sutton and A. G. Barto, *Reinforcement Lerning: An Introduction.* The MIT Press, Cambridge, MA, 1998.

[16] Wahl, Alfred, ”LA BALLE AU PIED HISTOIRE DU FOOTBALL”, 創元社, 2002.

付録 1. 遺伝的アルゴリズムを用いた準最適フォーメーションの獲得

以下では，本稿 4.1.2.で示した遺伝的アルゴリズムを用いたフォーメーションの獲得手法について記載する．

1. フォーメーション決定に際する問題点

本稿で既にサッカーエージェントの学習効率向上にあたって，チームフォーメーションを考慮することが有効である可能性を示した．一方，最も学習に適したフォーメーションの組み合わせは学習以前に知ることは出来ない．またサッカーのフォーメーションは選手の配置に関して自由度が高く，本稿で取り上げた 4 つのフォーメーション以外にも非常に多くのフォーメーションを作成することが出来る．そのため，膨大なフォーメーションの中から学習に適しているものを選び出すのは非常に困難であると言える．

2. 実装方法

本稿では上記の問題を解決する一つの方法として，遺伝的アルゴリズムを用いた準最適フォーメーションの自動的な探索手法を適用した．以下に実装方法を記す．

3.1. 染色体定義

エージェントが配置されるフィールドを横 9 マス×縦 9 マスで表現する．横の 9 マスを x 座標，縦の 9 マスを y 座標とし，x，y の値を用いてフィールド内の場所を特定出来るものとする．各エージェントには配置位置情報として x 座標，y 座標を持たせる．エージェントの配置位置は x，y に 0 から 8 の値を設定することによって定義される．GK を除く 10 エージェントの配置情報の列を染色体として定義する．つまり，1 つの染色体は 20 個の数値で構成される．図 20 に，既に本稿で取り上げたフォーメーション 1 を具体例として染色体の定義例を示す．

3.2. 遺伝的操作

3.2.1. 交叉

交叉手法として二点交叉を適用する．二点交叉とは，交叉点をランダムで二つ選び二つの交叉点に挟まれている部分を入れ換える方式を言う．本研究における具体的な交叉例を図 21，図 22 に示す．両図ではエージェント 4，５，６，７について座標の入れ替えが発生している．

3.2.2. 突然変異

突然変異は個体の遺伝子の一部を変化させる操作である．突然変異を用いることで親個体に存在しない遺伝情報を生成する効果がある．本研究で用いる突然変異の手順を以下に

エージェントNo	x座標	y座標
1	1	3
2	1	5
3	1	1
4	1	7
5	4	3
6	4	5
7	4	1
8	4	7
9	7	3
10	7	5

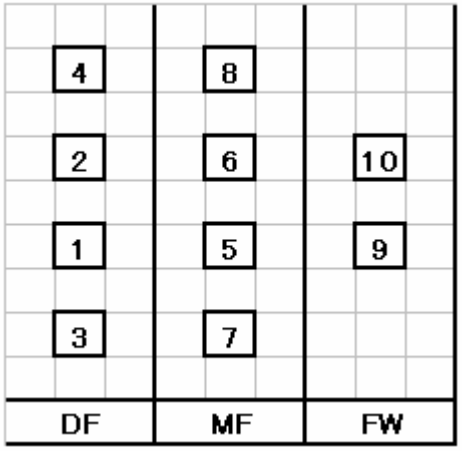

図 20. フォーメーション 1 の染色体定義

図 21. 交叉による座標の入れ替え（上: 交叉前 下: 交叉後）

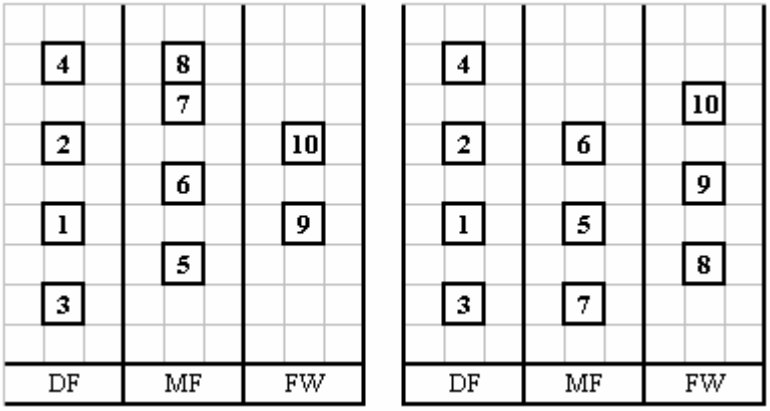

図 22. 交叉後に生成される 2 つのフォーメーション

示す.

1. あらかじめ設定された確率に従って全ての x 座標，y 座標に対して突然変異を行うか否かの判定を行う.
2. 突然変異を行う座標に対し，±2 の範囲でランダムに変化を加える.
   2.1. 変化を加えた結果座標が 9 以上になる場合, ランダムで 0 から 8 の値をセットする.
3. 全ての座標に対して突然変異が終わった後, 座標の重複の有無についてチェックを行う.
   3.1. 複数エージェント間で座標の重複が発見された場合はランダムで 1 エージェント選択する．選択されたエージェントについて，2. に示した操作から適用し直す.
4. 重複がなければ，突然変異を終了する

3.2.3. 評価方法

本稿 4 章で示した実験方法と同様に，対戦相手には学習を行わないチームを採用する．生成されたフォーメーションを自チームに適用し，対戦を 60 回繰り返す．対戦は 10 セット繰り返される．勝率の収束値を観測し，10 セットの平均値を算出する．算出した値を，自チームに適用したフォーメーションの評価点として定義する.

3.2.4. 選択

次世代に残す個体(子個体)を生成するための親となる個体(親個体)を決定する手法としてルーレット選択を適用する．ルーレット選択とは確率に基づいた選択手法であり，各個体の適合度に比例した割合で個体を選択する手法である.

3. GA のパラメータ設定

以下に GA を適用するに当たって用いられたパラメータ設定を示す.

個体数	10 **個体**
突然変異率	8%
交叉率	80%
対戦相手のアルゴリズム	**アルゴリズム** A
対戦相手のフォーメーション	**フォーメーション** 2（4-3-3）