

# 重ねあわせを用いたエージェントの評価関数について

星 謙作      加藤 貴司      白鳥 則郎

東北大学 電気通信研究所

{kenstar,p-katoh,norio}@shiratori.riec.tohoku.ac.jp

**概要** 本稿では、状況が動的に変化し、またエージェントの得られる情報やエージェント同士のコミュニケーションが不完全である環境において、エージェントが協調行動を行うための手法として、ポテンシャルを導入する。すなわち、環境に含まれる各要素の影響力をポテンシャルとして表現し、それらを重ねあわせることで、環境全体の評価関数として用いる事を提案し、この評価関数に基づいたエージェントの行動決定法について議論する。さらに、本手法を RoboCup Soccer Simulation League に適用し、その有効性を示す。

**キーワード** マルチエージェントシステム、評価関数、エージェント協調、ロボカップ

## Superposition of Evaluation Function for Agents

Hoshi Kensaku      Katoh Takashi      Shiratori Norio

Research Institute of Electrical Communication, Tohoku University, Japan

{kenstar,p-katoh,norio}@shiratori.riec.tohoku.ac.jp

**Abstract** In a dynamically changing environment, the agents cannot acquire complete information about the environment and so they are unable to achieve complete communication between them. In such an environment, it is difficult for the agents to accomplish their job, by cooperating with each other. In this paper, we introduce the concept of potential that each element of the environment is carrying and a method with which these agents cooperate between them. The potential expresses the expanse of influence of each element in the environment. The superposed potential of the each element in the environment is used to evaluate the potential of the whole environment. We used this evaluation function to decide agent's action. Furthermore, we apply this method to RoboCup Soccer Simulation League and show the effectiveness of our method.

**Keywords** Multi-agent System, Evaluation Function, Cooperation of agents, RoboCup

## 1 はじめに

災害救助の分野などの人間が活動できない危険な環境において、人間の代わりに活動するロボットの登場が望まれている。このような状況では、複数台のロボットが協調する事により、より効果的にタスクを実行することができる。その際、ロボットが協調作業を行なうために、互いがコミュニケーションをとり、状況に合わせて行動を選択する必要がある。しかし、このような環境では、ロボットの得られ

る情報が不完全である事や、周囲の状況が動的に変化してゆく事が考えられる。また、他のロボットから得られる環境に関する情報も不完全であることが考えられる。

すなわち、ロボット同士のコミュニケーションが保証されないため、動的に変化し、得られる情報が不完全である環境においてもロボットが協調作業を行うためには、他ロボットの行動を考慮した上で、適切な行動を選択する必要がある。

そこで本稿では、環境の情報が不完全で動的に変化するという特徴を備える環境に対しての評価関数としてポテンシャルを提案し、環境の一例としてのRoboCup Soccer Simulation League[1] に対して適用し、本手法の有効性を示す。

## 2 評価関数としてのポテンシャルの提案

### 2.1 ポテンシャルの導入

#### 2.1.1 ポテンシャルによる環境の表現

エージェントが協調動作をする場合、複数のエージェントが互いについての情報を持つ必要がある。しかし、エージェントが認識する情報が不完全な環境では、他エージェントから送られてくるそのエージェントについての情報についても正確さを期待できない。このような場合、エージェント同士が協調行動を行うためには、他エージェントの動作を考慮して行動しなければならない。そのためには、他エージェントの動作を評価する必要があり、他エージェントの状況も含めた環境全体の状態の把握が重要となる。

このように情報に多くの誤差を含み不完全である環境において、状況を把握する際には、誤差の影響を小さくすることが重要である。そのためには、一見詳細な情報が得られたとしても、一定の幅を持たせて情報を評価することが有効である。

そこで本稿では、情報が不完全なマルチエージェント環境においての状況把握の手法として、各要素の影響力をポテンシャルとして表現することを提案する。環境全体のポテンシャルを用いて適切な行動を求めることができれば、情報が不完全な環境においても、ポテンシャルを利用することによって、適切な行動を選択することが可能となる。

環境へのポテンシャルの適用は以下のように行う。まず、物体のおよぼす影響力をポテンシャルとして計算する。次に、環境の各要素のポテンシャルを重ねあわせることで、環境全体のポテンシャルを表現する。

仮に物体  $i$  のおよぼす影響力が、絶対座標に依存している場合、

$$f_i(x, y)$$

と表現することができる。このとき環境のポテンシ

アルは

$$\sum_i f_i(x, y)$$

となる。

ポテンシャルとは、物体の影響力を数値化したものである。つまり、ポテンシャルの絶対値が大きいとエージェントに対しての影響力が大きく、絶対値が小さいと影響力が小さいという事である。また、ポテンシャルが低いほど望ましいとした場合、ポテンシャルが高く絶対値の大きい場所は望ましくない影響力が大きいという事になる。

#### 2.1.2 ポテンシャルを用いた行動決定法

上で導入したポテンシャルは、ある物体がエージェントに与える影響力を表している。従って、影響力の分布を見ることにより、環境を大局的に見ることができ、自分にとって適切な行動を選択する際の材料として利用することができる。

具体的な行動決定法は適用する問題領域に依存するが、ポテンシャルの大きさや傾きを利用する方法が考えられる。

火災現場を例にとると、ポテンシャルの低い部分には要救助者がいることに対応する。またポテンシャルが高いとは、危険な物質が存在することなどに対応する。従って、要救助者のポテンシャルの絶対値をあらかじめ大きくしておく事によって、両者が存在する際に適切な動作を選択することができるようになる。

また、スポーツの場合では、フィールドにおける状態の有利不利を判断するものとしても利用できる。すなわち、環境を大局的に認識することにより、これから起こそうとする行動のリスクや方向性を判断できる。例えば、ボールをパスする際、どちらの方向が成功率が高いか、などが考えられる。また影響力の小さい位置を知る事により、これから影響力を広げるべき方向が分かる。

このように、環境の評価にポテンシャルを導入することによって、状況に応じた動作決定を行うことができる。

## 2.2 RoboCup Soccer Simulation League について

ここで本手法の適用例として、RoboCup Soccer Simulation League(以下 RoboCup)における協調行動の一つであるパスについて考える。

RoboCup とは、Soccer Server と呼ばれる共通の環境を利用し、マルチエージェントにおける分散協調や強化学習、チームモデリングなどの問題設定として利用されるように考えられたものである。

Soccer Server とクライアントは、UDP/IP プロトコルを用いて通信する。

クライアント プレイヤーの脳に相当する。各クライアントは、サーバから伝えられた情報を元に次の行動を決定し、サーバに送信する。各クライアントは一つのプロセスであり、一チームは11個のクライアントから構成される。また、クライアント同士の通信は行わない。

サーバ (Soccer Server) クライアントからの指示を元に、サーバ上のプレイヤーを移動させる。また、ボールなどの移動も制御する。そして、クライアントが得られる情報を作成し、その情報をクライアントに送信する。また、審判の役割も果たす。

また、クライアントが利用できる外部の情報は以下の二つである [2]。

**視覚情報** プレイヤーの向いている方向により、見える範囲が変化する(自分の背面方向は見る事ができない)。また、距離に応じて、他プレイヤーまでの距離などの値や背番号、敵味方の区別などの、得られる情報の正確さが変化する(図 1)。

**聴覚情報** プレイヤーは発声することができる。しかし、聞き取ることができるのは、単位時間に一人からの声のみである。また、届く範囲も一定距離に限られている。

各プレイヤーは単位時間(300ms)ごとに1つの行動をとることができ、2チーム22人が、一斉に行動する。

このように、RoboCup は、動的に状況が変化し、情報が不完全な環境の一つとなっている。以下では、本手法を RoboCup に適用した場合について議論する。

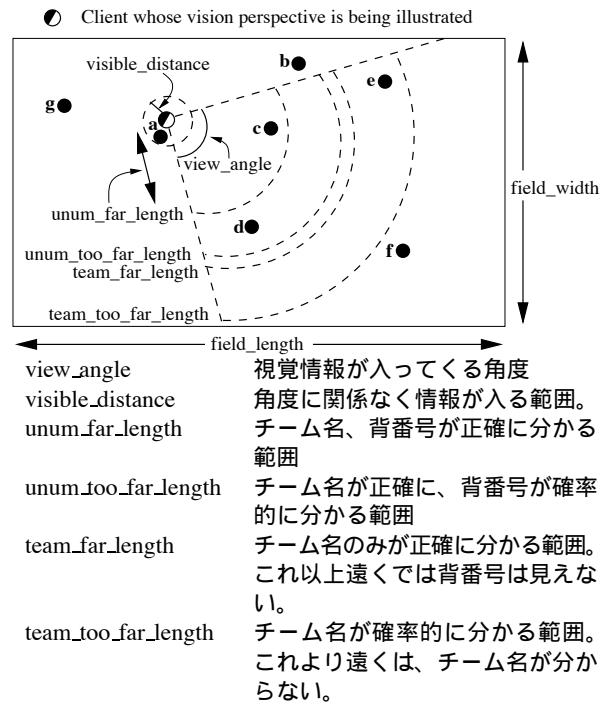


図 1: RoboCup における視覚情報 ([2] より引用)

## 2.3 ポテンシャルの RoboCup への適用

### 2.3.1 プレイヤーへポテンシャルの導入

RoboCup における環境の要素は、プレイヤーとボール、フィールドからなる。本稿では、これらの要素の中で最も環境に与える影響の大きい、プレイヤーに対してポテンシャルを適用する。

RoboCup におけるプレイヤーの影響力は絶対位置に依存し、プレイヤーの周囲に広がっており、プレイヤーの影響を与える範囲については、以下の特徴を持つと考えられる。

- 進行方向を軸として、左右対称  
フィールド上の物体には慣性があり、物体はすぐに曲がる事ができない。また、右に曲がる場合と左に曲がる場合では移動時間に差がないため、プレイヤーの影響力は左右対称であると考えられる。
- 進行方向前方が、後方よりも広い  
慣性があるため、いま進んでいる逆方向へ進むのには時間がかかってしまう。そのため、前方により広く影響力を持っていると考えられる。

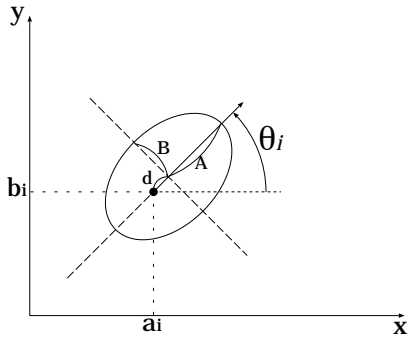


図 2: プレイヤーの影響力

以上の要件を満たす関数として、

$$f_i = \exp \left( C - \frac{(x \sin \theta_i + y \cos \theta_i - a_i - d)^2}{A^2} - \frac{(-x \cos \theta_i + y \sin \theta_i - b_i)^2}{B^2} \right) \quad (1)$$

が考えられる (図 2)。

ただし、エージェントが  $(a_i, b_i)$  に存在し、進行方向が  $\theta_i$  であるとする。 $(A, B, C, d)$  はエージェントの影響を与える範囲を定める定数である。これらは、環境に応じて決定される。 $d$  は速度により変化すると考えられるが、本稿では簡単のため定数とする。

### 2.3.2 行動決定法

次いで、RoboCup における基本的な協調行動としてパスを取り上げ、本手法を用いた適切なパスの実現方法について述べる。

味方のプレイヤー、敵のプレイヤーそれぞれは、ポテンシャルの形は同じとし、味方のポテンシャルは負となるように取る。また、ゲームの性質上、味方へのパスは望ましいのに対し、敵へのパスは望ましくない。従って、味方と敵のポテンシャルは正負が逆であるとする。

すなわち、エージェント  $j$  からみた味方チーム、敵チームのポテンシャルをそれぞれ

$$T_j(x, y) = - \sum_{\{i|i \in \text{味方}, i \neq j\}} f_i$$

$$O_j(x, y) = \sum_{\{i|i \in \text{敵}\}} f_i$$

とし、自分から見たフィールドのポテンシャルを

$$F_j(x, y) = T_j(x, y) + O_j(x, y)$$

とする。すると、

$$\vec{d} = - \begin{pmatrix} \frac{\partial F_j}{\partial x} \\ \frac{\partial F_j}{\partial y} \end{pmatrix}$$

で与えられる方向が、図 3 で見られるように、敵よりも味方の影響力が強い方向である。

しかし、図 3 から分かるように、傾きのみで蹴る方向を決めてしまうと、結果的に望ましくない方向へ蹴ってしまう可能性がある。そのため、 $\vec{d}$  にもっとも近いプレイヤーの方向を選択するなどの補正を行うことが望ましい。

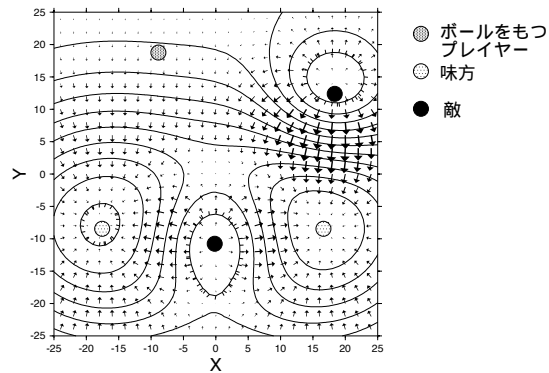


図 3: パス選択の際のポテンシャル

このようにポテンシャルを用いて環境を大局的に評価することにより、今まで煩雑であったパス方向の決定を容易に行うことができる。

## 3 実装と評価

### 3.1 目的

本実験では、動的に状況が変化し、エージェントが得る情報が不完全という特徴を持つ環境の一つである RoboCup に対し本手法を適用し、エージェントがポテンシャルを用いた環境の評価を行うことにより、適切な行動を選択することができることを示す。

なおこの実験は、AT 互換機の Linux 上でサーバ、クライアントを起動し行った。サーバは Soccer Server Ver7.09 を使用し、クライアントはプレイヤーに必要とされる基本的な能力を備えた CMUnited99 のソース [3] を元に C++ で作成した。

### 3.2 実験方法

RoboCupにおける協調行動の一つであるパスは、他エージェントの状況を考慮してボールを蹴り出す必要がある。また、実際の試合においてパスは、ある範囲内のプレイヤー敵味方3,4人ずつが関わるものである。そこで、フィールド上に一定範囲を定め、その中のパスの成功率を測定することにする。

プレイヤーのポテンシャルの式は、式(1)を用いる。但し、味方プレイヤー*i*のポテンシャルは $-f_i$ とし、敵プレイヤー*j*のポテンシャルは、 $f_j$ を用いる。

まず始めに、パスプレーを行う範囲を定める(図4の実線区域)。そして、パスを受けるプレイヤーを配置する範囲を定める(図4の破線区域)。

次に、プレイヤーを配置する。パスを受けるプレイヤーを、先にきめた範囲内(上記破線区域内)にランダムに配置し、パスをするプレイヤーをその範囲の外側に配置する。

次に、ボールをパスをするプレイヤーのそばにおき、すぐにパスを出せる状況にしておく。これにより、パスの能力のみを見ることが可能になる。

一人目のプレイヤー(パスをするプレイヤー)はポテンシャルの傾きの方向に最も近い方向にいるプレイヤーに向けてパスをする。その後ボールが他のプレイヤーに触れた時点、もしくはボールが範囲外に出た時点で終了とする。

ボールに触れた一人目と二人目のプレイヤーが同じチームであれば、パスは成功したと見なし、異なるチームであれば敵にパスをカットされ失敗したと見なす。また、二人目に触れることなくボールがパスプレーを行う範囲(図4の実線区域)からはずれてしまった場合はプレーが成立しなかったとする。

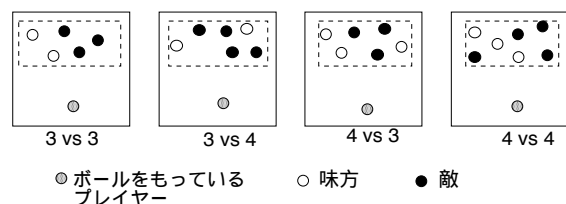


図4: 実験 初期位置の例

以上のような設定において、パスをするプレイヤーを含めて、味方対敵の数がそれぞれ3対3、3対4、4対3、4対4の条件で実験を行った。

### 3.3 実験結果

それぞれの条件で約50回ずつ実験を行った。その結果を表1に示す。

(括弧内は割合(%))

|     | 3vs3    | 3vs4    | 4vs3    | 4vs4    |
|-----|---------|---------|---------|---------|
| 成功  | 38 (76) | 34 (67) | 36 (71) | 37 (73) |
| 失敗  | 12 (24) | 16 (31) | 15 (29) | 14 (27) |
| 不成立 | 0       | 1 (2)   | 0       | 0       |

表1: ポテンシャルを用いた実験結果

また比較実験として、ポテンシャルを用いず、ランダムにパス相手を選択する手法を取った際の結果を表2に示す。

(括弧内は割合(%))

|     | 3vs3    | 3vs4    | 4vs3    | 4vs4    |
|-----|---------|---------|---------|---------|
| 成功  | 33 (65) | 21 (42) | 33 (65) | 35 (67) |
| 失敗  | 18 (35) | 29 (58) | 18 (35) | 16 (31) |
| 不成立 | 0       | 0       | 0       | 1 (2)   |

表2: ポテンシャルを用いなかった実験結果

### 3.4 考察

表1,2より、ポテンシャルを用いず、ランダムにパスの相手を選択した場合には、パスの成功率は40%から70%の範囲であるのに対し、ポテンシャルを用いた場合には、65%から75%ほどの成功率になり、成功率が上昇していることが分かる。それにともない、ポテンシャルを用いることにより失敗する確率を下げることができた。

また、不成立の場合がほとんど見られないのは、実験の設定上、味方や敵が多く存在している方向に蹴ることになるため、領域から出ることがほとんどないからである。

以上から分かるように、エージェントがポテンシャルを用いてパスの方向を決定することによりパスの成功率が向上した。これは、エージェントがポテンシャルを用いてフィールドの大局的な状態を把握する事により、適切な方向を選択することができていることによると考えられる。従って、ポテンシャルを

用いて環境を大局的にとらえるという手法は、エージェントの行動選択の際に有効であると言える。

今実験において、パスに失敗している原因は、主に2つ考えられる。

一つ目は、キックの強さとプレイヤーの位置関係によるものである。蹴る方向そのものは正しいものの、蹴る強さを調節していないために、ボールのスピードが速すぎる場合がある。この場合は、プレイヤーが受け取る前にボールが通りすぎてしまい、敵にとられてしまう。また遅すぎる場合は、ボールを受け取る前に敵に横からとられてしまうことがある。この問題は、ポテンシャルの傾きの大きさを利用して蹴る強さを調整するなどして解決される。

二つ目は、初期配置による問題である。味方のプレイヤー全てが敵のプレイヤーの影に隠れてしまい、有効なパスを行えない場合がある。今回、そのような状態でもパスを試みているが、ポテンシャルの傾きにより、パスをするべきかを判断できると考えられる。

このようにポテンシャルの傾きの大きさも利用することにより、さらにパスの成功率が高くなると期待できるが、これらの検証は今後の課題である。

## 4 おわりに

本稿では、状況が動的に変化し、情報が不完全な環境において、エージェントが協調するための評価関数としてポテンシャルを用いた手法を提案した。この手法により、エージェント同士の明示的なコミュニケーションがなくても協調行動をとることができるようになる。また、RoboCupにおけるパスに適用し、その有効性を示した。

今後は、プレイヤーの位置どりなど参加人数の多い協調行動に本手法を適用した場合の分析を行う予定である。

## 参考文献

- [1] M. Asada, H. Kitano, I. Noda, and M. Veloso. Robocup: Today and tomorrow - what we have learned, 1999.
- [2] Fredrik Heintz Mao Chen, Ehsan Foroughi. *Users Manual RoboCup Soccer Server for Soc-*

*cer Server Version 7.07 and later*, June 2001. <http://prdownloads.sourceforge.net/sserver/>.

- [3] Manuela Veloso Peter Stone, Patrick Riley. CMUnited99 source code. <http://www.cs.cmu.edu/pstone/RoboCup/CMUnited99-sim.html>, 1999.