

利害の衝突回避のための交渉コミュニケーションの学習と個性の発現

柴田 克成、伊藤 宏司 東京工業大学 Email : shibata@ito.dis.titech.ac.jp

Learning of Communication for Negotiation to Avoid Some Conflicts of Interests and Emergence of Individuality

Katsunari SHIBATA and Koji ITO Tokyo Institute of Technology

Abstract - We believe that communication in multi-agent system has two major meanings. One of them is to transmit one agent's observed information to the other. The other meaning is to transmit what an agent is intending. Here we focus the latter, and aim to the emergence of the autonomous and decentralized arbitration through communication among some agents. The communication contents and representation of them are not prescribed and are acquired by learning using a reinforcement signal, which is given to the agent after its action. The reinforcement signal is not shared with the other agents. Since the agent often has to make a decision from the past communication signals, the system architecture using recurrent type (Elman) neural network is proposed. The ability of this architecture was examined by two and four agents negotiation problems. A variety of negotiation strategies, that can be called individuality, emerged among the agents through the learning to avoid the conflicts after their decisions.

1. はじめに

コミュニケーションはマルチエージェントシステムにおける協調や調停に有効である。しかし、何をどのようにコミュニケーションすれば良いか予めわからなかったり、それを予め設計してしまうことで、環境の変化に対応できず、柔軟性を失ってしまうことも考えられる。さらに、我々生物のコミュニケーションは予め外部から与えられるものではなく、個々の学習によって獲得される。したがって、マルチエージェントシステムにおいて創発的手法によるコミュニケーションの獲得は大変重要であり、かつ可能であると考えられる。

筆者らは、コミュニケーションには大きく2つの意味があると考えている。一つは、エージェントの観測結果を他のエージェントに伝えることである。これは、受け手の観測が不十分である場合に有効である。有名な G. M. Werner らの研究 [1]における雄と雌の出会いの問題や中野らが行った共通言語の創発に関する先駆的な研究[2]は、いずれもこの範疇に属すると考えることができる。

コミュニケーションのもう一つの意味は、自分の意志を伝えることである。これは、複数のエージェント間の協調行動や利害の衝突を避ける際に有効である。このようなコミュニケーションに関する研究は、既に分散人工知能の分野で行われている[3][4]。しかし、いずれの場合も、コミュニケーションの内容やその表現方法および戦略が予め設計され、その設計の善し悪しが議論されている。

そこで、本研究では、「コミュニケーションの学習」と「自分の意志の伝達」に焦点を当てる。まず、予めコミュニケーションの内容や戦略を与えなくても、適切なコミュニケーション信号と行動を生成するためのエージェントの構成と学習方法を提案する。そして、エージェントが常に報酬を得るためには、相手のエージェントに応じて行動を変化させなければならないという問題を取り上げて、その能力を検証する。さらに、利害の衝突を回避するために、エージェントによる行動の違い（個性）が発現することを観察する。また、各エージェントが、コミュニケーション後の行動によって与えら

れた報酬や罰（強化信号）を他のエージェントと分配することなく、すべてのエージェントが満足する解が得られるかどうか一つのポイントである。

2. エージェントの構成と学習方法

図1に、提案するエージェントの構成を示す。エージェントは、他のエージェントおよび自分の1単位時間前のコミュニケーション信号を受け取り、ニューラルネットに入力する。ニューラルネットは、Elman 型のリカレントネット[5]を用いた。つまり、1単位時間前の中間層の値も入力値として与えられる。これにより、過去の履歴を利用した交渉や行動の決定が可能になる。ニューラルネットの各ニューロンの出力関数は、-0.5 から 0.5 の値域を持つシングモイド関数とした。出力は、コミュニケーション用と行動用の2つ用意した。コミュニケーション信号と行動は 1 と-1 のどちらかとし、1を出す確率はニューラルネットの対応する出力に 0.5 を足した値とした。エージェントは、コミュニケーション信号を同期して交換し合い、最後に行動を決定する。交渉の最初は中間層からのフィードバック信号も含めて入力をすべて0とした。これによって、最初のコミュニケーション信号が 1 である確率は、他のエージェントによらず常に一定となる。

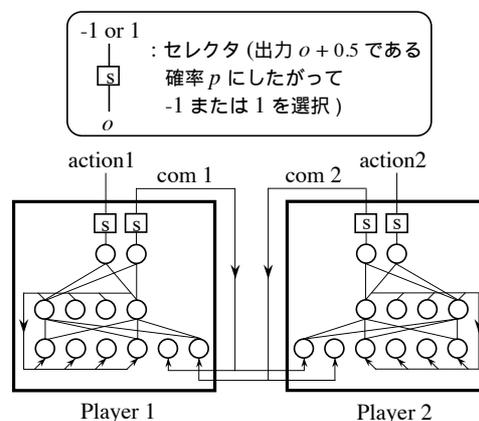


図1 Elman 型リカレントネットを用いたコミュニケーションを行うエージェントの構成

エージェントは、行動後に得られる強化信号を元にコミュニケーションと行動を学習する。エージェントが報酬を得た時には、その際にとっていたコミュニケーションと行動の生起確率を大きくし、罰の場合には小さくする。ニューラルネットワークは BP(Back Propagation) [6] または BPTT(Back Propagation Through Time) [6] という教師あり学習で各出力に対してそれぞれ次のような教師信号を内部生成してネット内の重み値とバイアスを共に学習する。

$$x_{ideal} = x + \eta r x'$$

x_{ideal} : 教師信号

$x=f(u)=1/(1+\exp(-u))-0.5$: ニューラルネットの出力

u : 出力ニューロンの内部状態

η : 学習係数 (ここでは0.1)

r : 強化信号 (ここでは報酬の場合 1、罰の場合 -5)

$x'=dx/du=(0.5-x)(0.5+x)$

o : 実際のコミュニケーション信号または行動

x' は、確率 p が 0 か 1 に近い場合に、安定になるように加えた。入力層から中間層への結合の初期値は微小な乱数とし、中間層から出力層へはすべて 0.0 とした。よって、学習前の出力の値はすべて 0.5 となる。BPTT では、誤差信号は入力層からフィードバック結合を通して中間層に戻り、交渉の始めまで時間をさかのぼって学習に用いる。一方、通常の BP では、誤差信号はフィードバック結合を通して時間をさかのぼらず、教師信号を与えた時間での状態のみに基づいて重み値が更新される。

3. シミュレーション

3.1 2 エージェント交渉問題

3.1.1 設定

始めに、2 エージェント交渉問題について述べる。図 2 にシミュレーションの環境を示す。まず、16 エージェントの中から 2 つのエージェントをランダムに選ぶ。エージェントは、相手がどのエージェントかを知らずに、3 回のコミュニケーションの機会が与えられ、各機会毎に、前節のように、1 か -1 のコミュニケーション信号を出力する。最初の機会には、相手のコミュニケーション信号を知ることなく出力を決定する。2 回目は、自分と相手の 1 回目のコミュニケーション信号を入力として出力を決定する。3 回目には、Elman ネットが 1 回目のコミュニケーション信号に関する情報を保持していれば、1 回目と 2 回目の両方の信号を考慮して出力を決定できる。そして最後に、エージェントは自分が通るルートを \circ (行動出力=1)か ∇ (行動出力=-1)のどちらかに決定する。もし、両者が同じルートを選んだ場合は罰($r=-5$)が、違うルートを選んだ場合は報酬($r=1$)が与えられる。この強化信号を元に、エージェントは前節の要領で 3 回のコミュニケーション信号と行動を学習する。この問題では、コミュニケーションの機会が 3 回なので、そのパターンから最大 $2^3=8$ のエージェントの区別が可能である。さらに、コミュニケーションパターンが同じ 2 つのエージェントが違う行動をとることもで

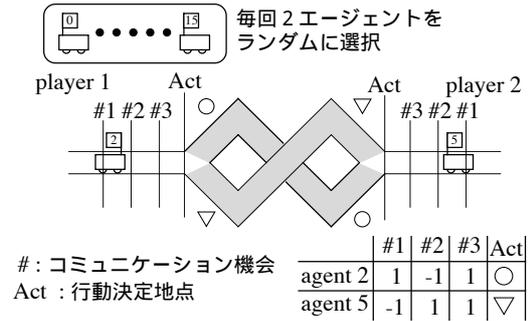


図 2 2 エージェント交渉問題のシミュレーション環境

きる。理想的には $2^4=16$ のエージェントの中から任意の 2 つのエージェントを選んで常にも違うルートをとることが可能である。

3.1.2 結果

表 1 の右側に、学習が成功した時の交渉の例を示す。学習が成功したほとんどの場合、各コミュニケーション信号や行動の選択確率は 0 か 1 に近い値となった。最初の 3 つの例について詳しく説明する。

例 1 : エージェント 0 とエージェント 1 が選ばれた場合。両者は 3 回ともコミュニケーション信号 1 を出力し、最終的にエージェント 0 はルート \circ を、エージェント 1 はルート ∇ を選択した。

例 2 : エージェント 0 とエージェント 2 が選ばれた場合。エージェント 0 は例 1 と同一のコミュニケーション信号と行動をとった。一方、エージェント 2 は、コミュニケーション信号 1 を 2 回出力した後、3 回目は -1 を出力。そして、ルート ∇ を選択した。

例 3 : エージェント 1 とエージェント 2 が選ばれた場合。エージェント 1 は、コミュニケーション信号 1 を 3 回出力した後、ルート \circ を選択。一方、エージェント 2 は例 2 と同様のコミュニケーション信号と行動をとった。

コミュニケーション信号の意味は学習前にエージェントに与えなかったが、学習後に獲得されたコミュニケーション信号に筆者らが意味づけをした。上述の例におけるエージェント 0 のように、相手のエージェントによらず常に同じルートをとったエージェントは、そのコミュニケーション信号のシーケンスも、相手によらず一定であった。このようなエージェントを、そのルートを通ることを主張し続けたと定義する。したがって、例 1 のエージェント 1 は、コミュニケーションの 3 回の機会でもルート \circ を通ることを主張し続けたが、最終的に、エージェント 0 に譲歩し、ルート ∇ を選択したと解釈できる。ところが、例 3 では、相手のエージェント 2 が先に譲歩したため、主張通りルート \circ をとったと解釈できる。

表 1 に、すべてのエージェントペア間での交渉結果を示した。表に出てくるエージェントの順番は、ルート \circ をとる確率が多いものから並べた。● は、上述の例におけるエージェント 0 のように、ルート \circ をとることを主張し続け、最終的に主張通りルート \circ をとったことを示す。上付き添え字のある \circ は、例 7 におけるエージェント 1 2 のように、添え字の

表1 学習後の選択されたエージェントペアに対するコミュニケーション信号と行動.

		エージェント No.															交渉の例						
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	agent	#1	#2	#3	Act	
エージェントNo.	0	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	(1)	①	1	1	1	○
	1	▽ ³	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	(1)	①	1	1	1	▽
	2	▽ ²	▽ ²	○ ₃	●	●	●	●	●	●	●	●	●	●	●	●	●	(2)	①	1	1	1	○
	3	▽ ²	▽ ²	▽ ²	●	●	●	●	●	●	●	●	●	●	●	●	●	(2)	②	1	1	-1	▽
	4	▽ ¹	▽ ¹	▽ ¹	▽ ¹	○ ₂	○ ₂	○ ₂	●	●	●	●	●	●	●	●	●	(3)	①	1	1	1	○
	5	▽ ¹	▽ ¹	▽ ¹	▽ ¹	▽ ₁₃	○ ₂	○ ₂	●	●	●	●	●	●	●	●	●	(3)	②	1	1	-1	▽
	6	▽ ¹	○ ₂₃	●	●	●	●	●	●	●	●	●	●	(4)	①	1	1	1	○				
	7	▽ ¹	▽ ¹	●	●	●	●	●	●	●	●	●	●	(4)	⑧	-1	-1	-1	▽				
	8	▼	▼	▼	▼	▼	▼	▼	▼	○ ¹	○ ¹	○ ¹	○ ¹	○ ¹	○ ¹	○ ¹	○ ¹	(5)	④	1	-1	1	○
	9	▼	▼	▼	▼	▼	▼	▼	▽ ₂₃	○ ¹	○ ¹	○ ¹	○ ¹	○ ¹	○ ¹	○ ¹	○ ¹	(5)	⑤	1	-1	-1	▽
	10	▼	▼	▼	▼	▼	▼	▼	▽ ₂	▽ ₂	○ ₁₃	○ ¹	(6)	⑥	1	-1	-1	○					
	11	▼	▼	▼	▼	▼	▼	▼	▽ ₂	▽ ₂	▽ ₂	○ ¹	(6)	⑦	1	-1	-1	▽					
	12	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○ ²	○ ²	○ ²	○ ²	(7)	⑫	-1	-1	1	○
	13	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▽ ₃	○ ²	○ ²	○ ²	(7)	⑬	-1	-1	1	▽
	14	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	○ ³	○ ³						
	15	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼						

番号のコミュニケーションの機会までルート▽を主張したが、その次の機会に主張を変更して最終的にルート○をとったことを示す。下付き添え字のある○は、例5のエージェント4や例6のエージェント6のように、最初はルート○を主張したものの添え字の番号のコミュニケーション機会でもルート▽を主張し、再びルート○を選択したことを示す。▼と▽は、最終的にとったルートが▽の場合であるが、見方はそれぞれ●と○と同様である。このように、各エージェントは、自律分散的に集団の中で順序づけられ、衝突を避けるようになっていることがわかる。また、上記の場合、ルート○の主張は、(1, 1, 1)であったが、シミュレーションによっては、(-1, -1, -1), (1, -1, 1), (-1, 1, -1) のこともあった。これは、各エージェントのニューラルネットの初期値の値、および、タスクにおけるエージェント選択の確率的要因のためである。しかし、(1, 1, -1), (1, -1, -1), (-1, 1, 1), (-1, -1, 1) にはならなかった。これは、自分の前回のコミュニケーション信号も入力となるため、前の信号を常に繰り返すか、前の信号の常に反対の信号を出すことを学習することがニューラルネットにとって容易であるからと考えられる。また、エージェントの数を8に減らすと、解はより簡単に得られた。この場合、2回以上主張を変えるエージェント、つまり、表1で下付き添え字のあるエージェントは出現しなかった。

この問題は、各エージェントの中間層ニューロン数2個だけで解くことができた。ただし、解に至る確率は非常に小さかった。この時の、表1のエージェント2のニューラルネットを図3に示す。そして、相手のエージェントを変えた4つの場合の中間層の値の変化を図4に示す。これらから、このエージェントの基本的な戦略は、

1. 中間層から両出力への結合の重み値はほぼ等しいので、両出力はほぼ同じ値となる。
 2. 中間層ニューロン2のパイアスが正で、そこからコミュニケーション出力への結合の重み値が正なので、交渉の最初のコミュニケーション出力は1となる。
 3. 相手のエージェントのコミュニケーション信号が - 1 ならば、入力層ニューロン1 中間層ニューロン2の負の結合と中間層ニューロン2から出力層への正の結合によって、両出力は1になる。これは、図4の(b)(c)(d)でも観察できる。
 4. 入力層ニューロン2 中間層ニューロン1、中間層ニューロン1 出力ニューロン1の2つの負の結合と、入力層ニューロン4 中間層ニューロン2への正のフィードバック結合により、前の出力を保持しようとする。
 5. もし、相手のエージェントのコミュニケーション信号が1であると、入力層ニューロン1 中間層ニューロン2への負の結合により中間層ニューロン2の出力が減少する。そして、自分と相手のコミュニケーション信号が共に1の場合が2回続くと、中間層ニューロン2の出力がさらに下がって0より小さくなり、コミュニケーション出力が - 1となる。これは、図4の(a)(b)で観察される。
- この項目5では、コミュニケーションの2回目と3回目の機会では入力が同じであるにもかかわらず、出力が変化していることになる。これは、中間層から入力層へのフィードバック結合により、文脈情報が保持されたことによって実現された。リカレントでない通常の階層型ニューラルネットで学習させると、中間層ニューロン数を増やしても、このような機能を持つエージェント、つまり、表1におけるエージェント

2, 3, 12, 13は出現しなかった。

ニューラルネットの中間層ニューロン数を4個とし、結合の初期値を変化させて100回のシミュレーションを行った。BPTTで学習させた場合、41回で解が得られ、通常のBPの場合、52回で解が得られた。Elman型ニューラルネットでは、中間層からフィードバックされた入力と中間層との結合は、文脈情報の保持と出力への反映の両方に使われる。BPTTでは文脈の保持まで考慮した学習ができるが、通常のBPでは文脈情報の出力への反映だけが学習される。しかしこの場合は、文脈情報を出力へ反映させる学習が、文脈の保持にも有効であったと考えられる。

3.2 4 エージェント交渉問題

3.2.1 設定

次に、4 エージェント交渉問題について述べる。前述の2 エージェント交渉問題では、全エージェントを順序付けし、選ばれた2 エージェント間でコミュニケーションを通して相手との順序関係を知り、その相手との関係だけで衝突の回避ができた。しかし、4 エージェントの場合、他の一つのエージェントとの間で衝突が回避できても、それによって逆に、それ以外のエージェントとの間で衝突が起きる可能性もある。さらに、どの4つのエージェントが選ばれるかだけでなく、他の3つのエージェントが相対的にどのような位置関係で配置されるかも考慮する必要がある。従って、我々が各エージェントのコミュニケーション手順を設計するのは困難である。

図5に、4 エージェント交渉問題の設定を示す。全部で8個のエージェントの中から4つのエージェントが毎回ランダムに選ばれ、4つの入り口のどれかにランダムに置かれる。エージェントは、-1か1のコミュニケーション信号を前の問題と同様に3回送ることができる。各エージェントは、一度に4個のコミュニケーション信号を受け取る。1つは自分自身の信号、2つ目は向かいのエージェントから、3つ目は、ルート○側の隣から、4つ目は、ルート▽側の隣からの信号である。ただし、エージェントは、どのエージェントが選ばれているか、選ばれたエージェントがどこに配置されているか、どの信号がどこから来た信号かといった情報は知らされない。そして、最後に、○か▽のどちらのルートをとるかを決める。そして、隣のエージェントと衝突すると罰 ($r=-5$) が、衝突しないで通り抜けると報酬 ($r=1$) が与えられる。得られた強化信号は、他のエージェントと分配しない。エージェントの構成は、図1とほぼ同じであるが、コミュニケーション信号の入力の数が増える。この問題は、選ばれたすべてのエージェントが、図5の菱形のルートを時計回りか反時計回りに回る場合が解となる。そのためには、向かい側のエージェントとは同じで、隣とは違うルートをとれば良いことがわかる。

3.2.2 結果

この問題で解ける最大のエージェント数は8より大きく、解には様々なバリエーションがあった。ある解では、特定のエージェントが2 エージェント問題での表1のエージェント

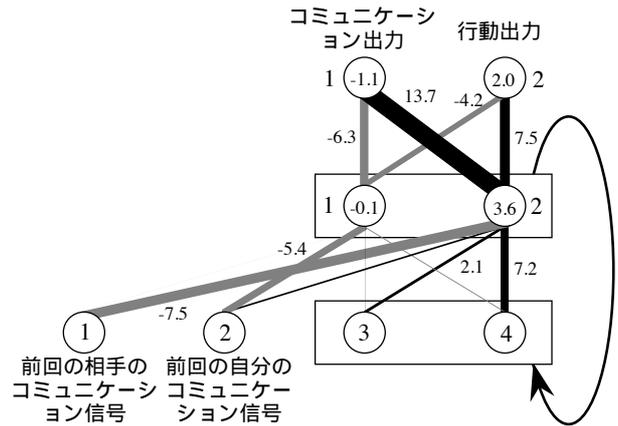


図3 学習後のエージェント2のニューラルネットの結合の重み値とニューロンのバイアス

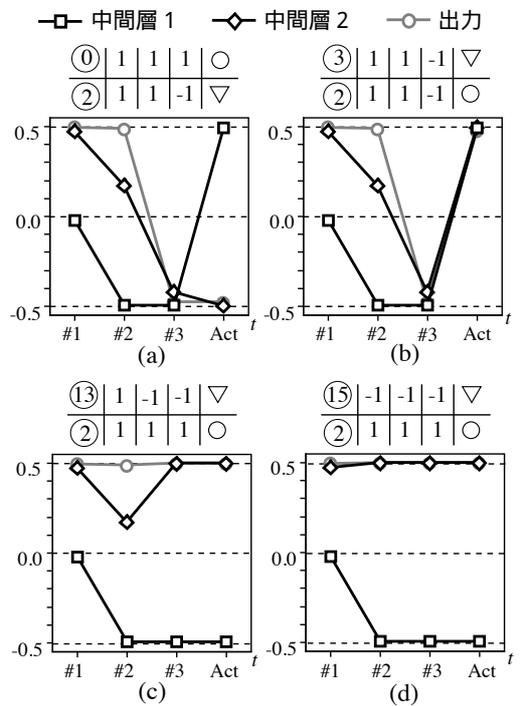


図4 相手のエージェントによるエージェント2の中間層の値の時間的変化の比較

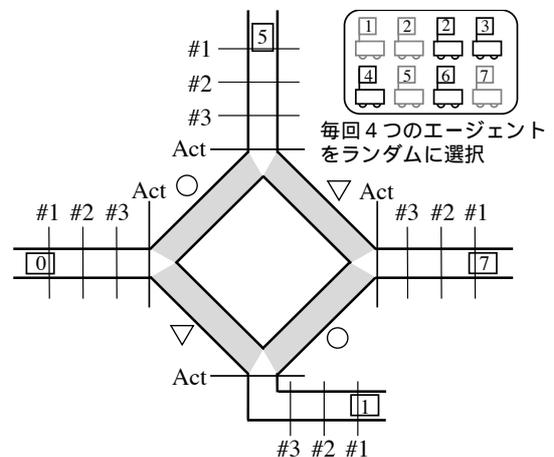


図5 4 エージェント交渉問題のシミュレーション環境

0のように、相手や場所によらず常に同じコミュニケーション出力と行動をとるという場合があった。しかし、ほとんどの場合、すべてのエージェントが状況に応じてルートを変化させていた。

この問題も、解ける確率は少ないものの、各エージェントの中間層ニューロン数を2個としても解を得ることができた。この時の解の詳細を述べる。表2に、典型的な交渉の例を示す。表2(a)では、エージェント0, 2, 3, 7が選ばれている。最初と2回目のコミュニケーション信号は4つのエージェントとも-1で同じであるが、エージェント7だけが、3回目に信号を1に変化させている。そして、最終的に、エージェント0と3が行動1(ルート○)をとり、エージェント2と7が行動-1(ルート▽)をとった。表3に、学習後にすべての組み合わせで交渉を行った場合の各コミュニケーション信号と行動の確率とコミュニケーション・行動のパターンの数を示した。表1の場合のようにエージェントのソートは行っていない。この表の中の確率の値は、1を出力する確率を p として、 $2p-1$ とした。したがって、符号は、ルート○かルート▽かの好みを表し、絶対値がその好みに対する固執度を表している。これより、エージェントが状況に応じてコミュニケーション・行動パターンを変化させていることがわかる。最初のコミュニケーション信号は、常に一定であるため、とりうるすべてのパターンは8個であるが、エージェント0は、この8パターンのうち、7個のパターンをとっており、非常に柔軟で適応的である。また、すべてのエージェントにおいて、3回のコミュニケーション信号の好みの符号は等しく、行動の好みの符号は反対である。このことから、この解では、コミュニケーション信号-1が行動1(ルート○)の主張であると解釈できる。エージェント間のすべての組み合わせでの交渉結果を見てみると、次のようないくつかのルールを見つけることができた。

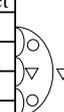
- (ケース1) 選ばれた4つのエージェントの好みのルートがすべて○の場合、つまり、エージェント0, 2, 3, 7が選ばれた場合(表2(a))、エージェントの配置によらず、コミュニケーションの機会#3でエージェント7がコミュニケーション信号を変化させ、最終的に向かいのエージェントと共にルート▽をとった。
- (ケース2) 表2(b)のように、3つのエージェントの好みのルートが○で、他のエージェントの好みのルートが▽である場合、▽を好むエージェントおよび隣接する2つのエージェントは意志を変えず、向かいにいるエージェントが意志を変えてルート▽を選択した。
- (ケース3) 4つのエージェントの好み半々の場合、さらに2つの場合に分けられる。1つは、同じ好みのエージェントが向かい合っている場合で、この場合は、全エージェントが最初の意志を貫いた(表2(c))。
- (ケース4) 好み半々に分かれた場合で、同じ好みのエージェントが隣同士の場合は、エージェントのコミュニケーションと行動は複雑で、簡単なルールを見つけていくことができなかった(表2(d))。

表2 4エージェント問題における学習後の典型的なコミュニケーションパターン

	#1	#2	#3	Act
agent 0	-1	-1	-1	1
agent 2	-1	-1	-1	-1
agent 3	-1	-1	-1	1
agent 7	-1	-1	1	-1



	#1	#2	#3	Act
agent 2	-1	-1	-1	1
agent 7	-1	1	1	-1
agent 0	-1	-1	-1	1
agent 5	1	1	1	-1



	#1	#2	#3	Act
agent 2	-1	-1	-1	1
agent 5	1	1	1	-1
agent 7	-1	-1	1	-1
agent 4	1	1	1	-1



	#1	#2	#3	Act
agent 2	-1	-1	-1	1
agent 4	1	1	1	-1
agent 5	1	1	1	1
agent 7	-1	-1	1	-1

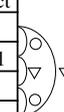


表3 学習後の各エージェントのコミュニケーション信号と行動の確率およびコミュニケーションパターン数 (確率の数値は p を実際の確率とした時の $2p-1$ で表記)

	com #1	com #2	com #3	Action	パターン数
agent 0	-1.0	-0.429	-0.457	0.352	7
agent 1	1.0	0.029	0.286	-0.257	4
agent 2	-1.0	-1.000	-1.000	0.581	2
agent 3	-1.0	-0.086	-0.343	0.181	5
agent 4	1.0	1.000	1.000	-0.571	2
agent 5	1.0	0.771	0.533	-0.486	5
agent 6	1.0	0.429	0.324	-0.286	5
agent 7	-1.0	-0.771	-0.552	0.486	4

ケース1とケース2のそれぞれと対称な場合もあるが、同様な戦略であった。このシミュレーションと同様に、ニューラルネットの初期値を変えた他のシミュレーションでも、8個のエージェントの好みは半々になるという対称性を予測していた。しかし実際には、最も非対称な場合で、2つのエージェントがルート○を好み、残りの6個がルート▽を好むという結果になった場合もあった。

図6にエージェント0, 2, 7のニューラルネットの結合の重み値を示す。各エージェントの特徴は次のように記述することができる。

[全エージェント]すべてのエージェントで、入力1および入力4から中間層への結合の符号はほぼ同じで、入力2および入力3からの結合の符号はその逆になっている。これによって、向かいのエージェントとは同じ行動を、隣のエージェントとは違う行動を得る力が働く。

[エージェント0]全般的に、他のエージェントのコミュニケーション信号の入力である入力1, 2, 3から中間層への結合の重み値の絶対値が大きく、逆に、文脈入力および自分のコミュニケーション信号入力から中間層への結合の重み値の絶対値は小さい。つまり、エージェント0は自分の意志よりも、他のエージェントの意向を優先してコミュニケーション、行動を決定している。したがって、相手によって様々なパターンが生成されると考えられる。

[エージェント 2] 中間層からコミュニケーション出力への結合の重み値が 0 に近いため、エージェント 0 とは逆に、他のエージェントの意向は参考にせず、コミュニケーション信号を生成していることがわかる。このことは、表 3 からわかる。

[エージェント 7] 表 2 の(a)(d)のように、最初の 2 回のコミュニケーション信号が同じでも、2 回目と 3 回目に違う信号を生成することができる。これも、図 6 (c) の結合から大まかに説明できる。中間層ニューロン 2 はバイアスが小さいため、# 1 では出力値が非常に小さい。したがって、# 2 では、入力 6 からの負の結合により中間層ニューロン 1 の値が大きくなり、コミュニケーション出力は小さくなる。しかしこの時、入力 3 からの負の結合によって中間層ニューロン 2 の値は大きくなり、# 3 では、中間層ニューロン 1 の値が小さくなってコミュニケーション出力が 1 に近くなる。

中間層を 4 個にして、ニューラルネットの結合の初期値を変えて 100 回シミュレーションを行った。BPTT 適用時には 42 回で解が求まり、通常の BP では、26 回で解が求まった。このことから、多少 BPTT の方が優位であるとは言え、あまり大きな変化がないことがわかった。

4. 結論

マルチエージェント間でのコミュニケーションを、観察した情報を伝えることと自分の意志を伝えることの 2 つに大きく分類することを提案した。そして、後者のためのコミュニケーションを各エージェントが強化信号から自律的に学習するアーキテクチャを提案した。例題のコミュニケーションによる衝突回避交渉の問題では、エージェントにコミュニケーションの内容やその表現法および戦略を予め与えなくても、学習によってコミュニケーションを通して衝突を回避できるようになった。この時、リカレントニューラルネットが必要に応じて過去の情報を保持しながら、相手のエージェントに対して適応的に交渉を進めていることがわかった。また、エージェントによって自分の主張を固辞する度合いに差ができた。これは個性と呼べるものと考えられる。また、1 対 1 の交渉だけではすまない 4 エージェント間の交渉の問題において、強化信号を他のエージェントと分け合わなくても、全エージェントが満足する解を得ることができた。この過程は、自律分散的な変化によって系のポテンシャルを落ちていくという意味で、ホップフィールドネット[7]を使った最適化の過程に類似している。また、本学習は、意思伝達のコミュニケーションに特に有効であるが、コミュニケーション一般の学習に利用できると考える。

謝辞

本研究の一部は、日本学術振興会未来開拓学術研究推進プロジェクト“生物型適応システム” (JSPS-RFTF96I00105) および文部省科学研究費基盤研究(B) (No. 10450165) の補助の下で行われた。

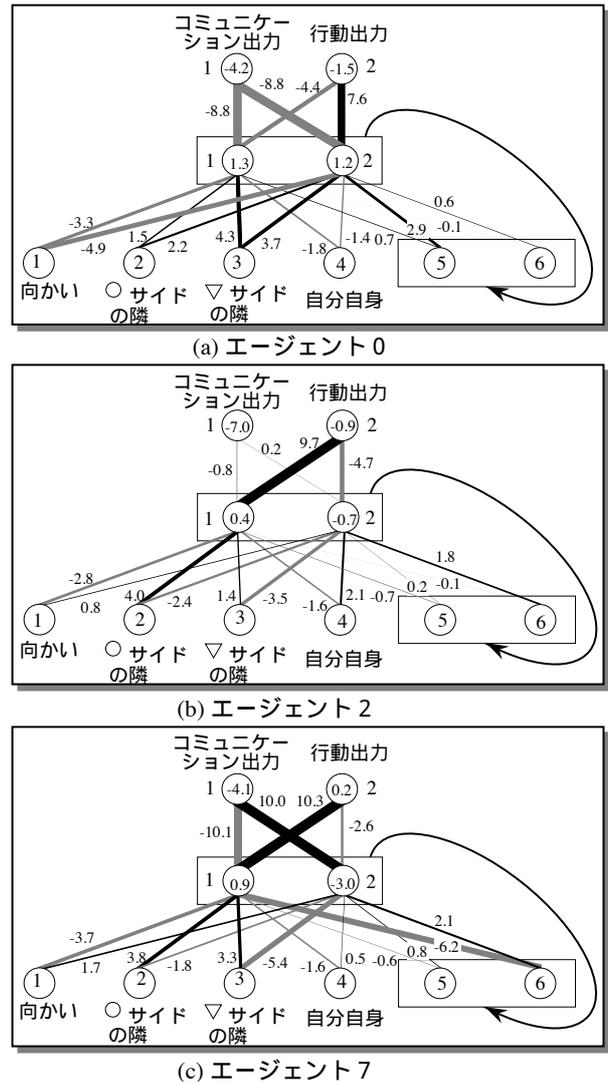


図 6 4 エージェント問題学習後のニューラルネットの結合の重み値とニューロンのバイアス値

参考文献

- [1] Werner, G. M. & Dyer, M. G., "Evolution of Communication in Artificial Organisms", Proc. of Artificial Life II, pp.1-47 (1991)
- [2] Nakano, N., Sakaguchi, Y., Isotani, R. & Ohmori, T., "Self-Organizing System Obtaining Communication Ability", Biological Cybernetics, 58, pp.417-425 (1988)
- [3] Davis, R. & Smith, R. G., "Negotiation as a Metaphor for Distributed Problem Solving", Artificial Intelligence, Vol. 20, Np. 1, pp. 63-109 (1983)
- [4] Kreifelts, T. & von Martial, F., "A Negotiation Framework for Autonomous Agents, Demazeau, Y. & Muller, J.-P. eds., Decentralized A. I. 2, pp.71-88 (1991)
- [5] Elman J. L., "Finding Structure in Time", Technical Report CRL 8801, Center for Research in Language, Univ. of California, San Diego (1988)
- [6] Rumelhart, D. E., Hinton, G. E. & Williams, R. J., "Learning internal representations by error propagating", Parallel Distributed Processing, Vol. 1, MIT Press, pp. 318-362 (1986)
- [7] Hopfield, J. J., "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", Proc. of Natl. Acad. Sci. USA, Vol. 79, pp.2554-2559 (1982)