

## 選好の自己強化を伴うエージェント集団の多様性

岩永 佐織<sup>\*1</sup>, 生天目 章<sup>\*2</sup>

1: 海上保安大学校, \*2: 防衛大学校

多数の均質なエージェントが自らの選好(利得構造)を自己強化しながら二者択一の意味決定を繰り返す状況を扱う。集団全体の意思決定プロセスを安定性、効率性、そして、公平性の観点から評価する。また、最初は均質なエージェント集団がどのような多様なエージェント集団に進展するかを分析する。エージェントが移動せず同じ近傍のエージェントと繰り返し相互作用する場合には、同じ選好をもつ者どうして多数の小規模なニッチを形成する。一方で、利得が低く不満なエージェントが移動する場合には、集団全体で二つの大規模なニッチを形成する。そのことにより、一人ひとりが高いレベルの利得を獲得する安定した状態に収束することを示す。

## Diversity of Agents with Self-enforcing their interests

Saori IWANAGA<sup>\*1</sup>, Akira NAMATAME<sup>\*2</sup>

1: Japan Coast Guard Academy, \*2: National Defense Academy

We handle the situation that a lot of homogeneous agents repeat a decisionmaking of a choice between two alternatives while doing self-enforcing of the interest (the payoff structure). We estimate decision making process of the population from stability, the efficiency and the angle of the equity. We analyze for what kind of various population a population consists of a homogeneous agents develops. In the case when an agent doesn't move and acts on it mutually repeatedly with an agent in the same neighborhood, person with the same interest forms a lot of small-scale niches. On the other hand when a profit is low and a dissatisfied agent moves, two large-scale niches are formed out in the population. An then, we show that population settles in the stable state by which an individual gets high payoff.

### 1. はじめに

人間は社会の中で多くの人々と様々な人間関係を築く。人間の行動は、その人の性格などによってのみ決まっているわけではなく、関係する相手と取る行動に依存することが多い<sup>1)</sup>。そして、相互作用を扱う際にゲーム理論を用いる研究が着目されている。また、複雑な社会現象の解明を目的とした社会システムの分析に社会シミュレーションの1つのアプローチが用いられる。その中で要素と要素間の相互作用をモデリングするボトムアップ的手法を用いたマルチエージェントが社会現象の理解などの手段として着目されている<sup>2)</sup>。

### 2. シミュレーションの設定

#### 2.1 相互依存関係

人間の相互作用のあり方は協業と分業に分けられ、エージェントの研究でも協調的な関係<sup>3)</sup>と相互補完的な関係(以下、相補的な関係という。)に関する研究が多くある。多くの人が協力することで相乗効果が生まれる関係を「協調的な関係」という。一方で、各人が分業しあうことで相乗効果が生まれる関係を「相補的な関係」という。このうち、本論文では協調的な関係を扱っていく。

エージェントは2つの選択肢( $S_1$ と $S_2$ )のうち、相手と取る行動を知らないままどちらかを1つを選択する。このとき、エージェントはどちらを選択するであろうか。エージェントどうしが協調的な関係にあるとき、エージェント $A_i$ の利得をTable1の利得行列で与える。Table1は、エージェントの得る利得が自分の選択と相手の選択に依存し、相手と同じ選択をするならば利得を得られることを意味している。

ここで、簡単な2人のエージェントについて考えてみ

る。各エージェントはTable1の利得行列を持ち、相手と同じ選択をすることにより自分も相手も利得を得られる。ただし、エージェント毎に利得パラメタ $\theta_i$ は異なる。 $\theta_i$ がエージェント $A_i$ の異質性を表す。そして、利得パラメタが $\theta_i < 0.5$ の条件を満たすエージェントは、 $S_1$ を選択した場合に得る利得が $S_2$ を選択した場合に得る利得より大きいので、 $S_1$ の方が有利であると考える(タイプ1)。一方で、 $\theta_i > 0.5$ であるエージェントは $S_2$ の方が有利であると考える(タイプ2)。このように、利得パラメタ $\theta_i$ は各エージェントの選好を表しており、選好の違いによってエージェントを2つのタイプ(タイプ1:  $S_1$ を選好, タイプ2:  $S_2$ を選好)に分類できる。

Table1 Payoff matrix of  $A_i$  ( $0 \leq \theta_i \leq 1$ )

|                |       |                |            |
|----------------|-------|----------------|------------|
|                |       | Other agent    |            |
|                |       | $S_1$          | $S_2$      |
| Agent<br>$A_i$ | $S_1$ | $1 - \theta_i$ | 0          |
|                | $S_2$ | 0              | $\theta_i$ |

#### 2.2 相互作用モデル

近傍の者だけと相互作用するものとして扱うのが局所的な相互作用である。様々なエージェントで構成されている集団において局所的に相互作用が行われる場合には、エージェントがどのような相手と相互作用するかが重要であり、ランダムな相互作用と選択的な相互作用に分けて考えることができる<sup>4)</sup>。エージェントが様々なタイプの相手と相互作用するのがランダムな相互作用であり、特定のタイプの相手と相互作用するのが選択的な相互作用である。さらに、選択的な相互作用は同じタイプどうしの相互作用と異なるタイプどうしの相互作用に分類できる。エージェントが協調的な関係にある場合には、同じタイプどうして相互作用をするときに最も良い協調関係が成立する。

### 2.2.1 集団の多様性

利得行列を構成している利得パラメタ  $\theta_i$  によりエージェントの異質性を定義する。そして、多数のエージェントで構成されている集団の多様性を利得パラメタ密度関数により定義する。  $N$  人の集団  $G$  で利得パラメタ  $\theta$  をもつエージェントの数を  $n(\theta)$  で表し、  $n(\theta)$  を  $N$  で割った値を利得パラメタの密度関数  $f(\theta)$  とする。

$$f(\theta) = n(\theta)/N \quad (1)$$

### 2.2.2 局所的相互作用

「格子モデル」は、生物の生態学的特性などを明らかにし、時間による変化を明らかにするために用いられることが多い。そして、Axelrod<sup>5)</sup> は社会に協力関係が発生するために空間構造の重要性を指摘している。本論文でも格子モデルを採用している。  $50 \times 50$  格子のトラス状の2次元平面社会を設定し各格子にエージェントを1人ずつ配置し、各エージェントが隣接した4人と相互作用するものとする。

### 2.3 適応ルール

本論文では、各エージェントが「短視眼的な」仮定に基づき、離散時間において決定論的に意思決定するものとする。各エージェントは、相手の行動が前回から変化しないと予想して合理的な方法で行動を決定する。このとき、エージェントは自己の利得を計算し利益の高い方の選択肢を選ぶ。これを最適反応学習と定義する。

ある時点  $t$  において、エージェント  $A_i$  の近傍4人の中で  $S_1$  を選択するエージェントの割合を  $p_i(t)$  ( $0 \leq p_i(t) \leq 1$ ) で表す。協調的な関係にあるエージェント  $A_i \in G$  は  $S_1$  または  $S_2$  を選択して利得を得る。その際、エージェントが1相互作用あたりに得る期待利得  $U_i(S_1)$  または  $U_i(S_2)$  はそれぞれ次式で与えられる。

$U_i(S_1) = (1-\theta_i)p_i(t)$ ,  $U_i(S_2) = \theta_i(1-p_i(t))$ (2)  
 エージェント  $A_i$  は、これらと比較して次の時点  $t+1$  で次式のように選択をする (最適反応学習)。

$$p_i(t) > \theta_i \rightarrow S_1, \quad p_i(t) < \theta_i \rightarrow S_2 \quad (3)$$

ただし、  $p_i = \theta_i$  の場合には、前回と同じ選択をする。

### 2.4 移動ルール

エージェントが高い利得を得ようとするならば、2次元平面上の他の格子へ移動して相互作用の相手を変更することは有効な行動であり、主体的に相互作用の相手を選ぶことは重要である<sup>3)</sup>。エージェントは予め設定されたしきい値  $\mu$  をもち、獲得した利得が  $\mu$  を下回るときに他の格子へ移動する。エージェント  $A_i \in G$  が選択肢  $S_k$  を選択することにより得た利得を  $U_i(S_k)$  ( $k = 1, 2$ ) で表すと、移動ルールは次式で表される。

$$U_i(S_k) \geq \mu \rightarrow \text{stay}, \quad U_i(S_k) < \mu \rightarrow \text{move}(4)$$

利得が  $\mu$  以上の者はそのまま今の相手と相互作用するが、  $\mu$  より低く不満足な者は移動し相手を変更する。

ところで、利得がしきい値を下回り移動しようとしているエージェント (以下、移動対象エージェントという) が集団の中で一斉に移動をする。したがって、多数のエー

ジェントがそれぞれどこに移動するかが問題になってくる。ここでは、移動対象エージェントどうしがランダムに2人ずつのペアを作りペアどうしが格子を入れ替わるという方法をとる。すなわち、各移動対象エージェントはペアを作った相手の格子へ移動することにより、移動対象エージェントが全員同時に移動する。

### 2.5 選好関係の自己強化

本論文では、エージェントの異質性や集団の多様性がどのようにして生まれるかを明らかにしていく。各エージェントはTable1の利得行列をもち、式(3)の適応ルールに従って合理的に選択肢を選ぶ。その際、各エージェントは固定の利得パラメタ値をもつだけでなく、相互作用によって獲得した利得に基づき自らの利得パラメタの値を更新し、これを「選好関係の自己強化」と定義する。

すなわち、エージェント  $A_i$  がある一定の時間 ( $T$ ) 同じ選択肢  $S_1$  をとりつづけたならば、選択肢  $S_1$  に対する利得パラメタ値  $\theta_i$  を  $\Delta\theta$  だけ増加させ、その分  $S_2$  に対する利得パラメタ値を  $\Delta\theta$  だけ減少させる。エージェントが同一の選択肢をとりつづけるのは、その方が利得が高いためであり、利得パラメタの増加は当該選択肢をより選択する傾向が強くなることを示す。反対に、選択肢  $S_2$  を  $T$  回連続して選択したときは、選択肢  $S_2$  に対する  $\theta_i$  を  $\Delta\theta$  だけ増加させる。これをTable2に示す。

Table2 Payoff matrix of agent  $A_i$  ( $0 \leq \theta_i \leq 1$ )

|             |       |                               |                           |
|-------------|-------|-------------------------------|---------------------------|
|             |       | Other agent                   |                           |
|             |       | $S_1$                         | $S_2$                     |
| Agent $A_i$ | $S_1$ | $(1-\theta_i) + \Delta\theta$ | 0                         |
|             | $S_2$ | 0                             | $\theta_i - \Delta\theta$ |

### 2.6 評価基準

集団の中で多数の異質なエージェントが相互作用することによって最終的に生まれる集合行為を評価する。評価基準として、安定性、効率性、平等性の3つを用いる。集合行為は個々のエージェントの行動を集約することによって表されることが多く、ここでも集団の中で選択肢  $S_1$  を選択しているエージェントの割合を集合行為と定義する。集合行為の安定性は、システムの安定性を計る尺度であり集合行為を扱う上で重要である。シミュレーションでは、最初に各エージェントがとる選択肢を任意に与え、最終的に自己組織化される集合行為を求める。そして、選択肢  $S_1$  を選択するエージェントの初期割合  $p(0)$  (以下、集合行為の初期値という) と最終的に  $S_1$  を選択するエージェントの割合  $p^*$  (以下、集合行為の最終値という) との関係を集集合行為の安定性と定義する。すなわち、集合行為の最終値  $p^*$  が集合行為の初期値  $p(0)$  に依存しにくいことを集合行為が安定であるという。

次に、1相互作用あたりの平均利得を効率性と定義する。最終的な相互作用において、1相互作用あたりに各エージェントが得る利得の分布関数を  $g(u)$  で表すならば、効率性  $\bar{U}$  (平均利得) は次式で与えられる。

$$\bar{U} = \int_0^1 ug(u)du \quad (5)$$



集合行為の評価をする上で、全体の効率性だけを尺度とした研究が多い。しかし、人間は感情的に公平さを重んじ、公平さから生じる振る舞いが長い目で見て人間あるいは社会全体に利益をもたらしてきた。そこで、公平性の尺度からも集合行為を評価するため、エージェント間の利得格差を平等性と定義する。

ここで、経済学で用いられているローレンツ関数  $L(x)$  を導入する。ローレンツ関数  $L(x)$  は、利得の最も低いエージェントから並べ、その利得が全体に占める割合を横軸に示し累積利得を縦軸に表した曲線であり、曲線以下の部分の面積が平等性の尺度となる。この面積が大きいほど平等性の高い社会を示す。利得の分布関数  $g(u)$  を用いると、ローレンツ関数  $L(x)$  は次式で与えられる。

$$L(x) = \int_0^x \tau g(\tau) d\tau / \int_0^1 \tau g(\tau) d\tau \quad (6)$$

ただし、 $x = \int_0^w g(\tau) d\tau$ ,  $w$  は獲得した利得である。そして、平等性  $E$  は次式で定義される。

$$E = 2 \int_0^1 L(x) dx \quad (7)$$

### 3. シミュレーション結果

#### 3.1 エージェントが固定されている場合

これまでの研究成果から選好関係の同質な集団よりも多様な集団において望ましい集合行為が自己組織化されることが分かっている<sup>6)</sup>。本論文では、均質な集団に選好の多様性が自己組織化されるのかを調べる。最初は全員が同じ利得パラメタ  $\theta_i = 0.5$  をもつ均質な集団とし、Table2 において  $T = 10$ 、利得パラメタの増加分を  $\Delta\theta$  をそれぞれ、0.5, 0.1, 0.01 と設定してシミュレーションを行った。また、 $\Delta\theta$  が 0.5, 0.1, 0.01 の場合の適応回数をそれぞれ 1000, 100, 100 とした。

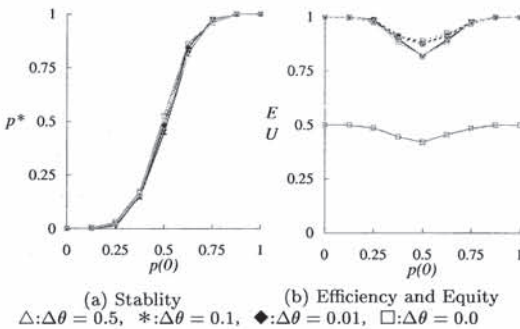


Fig.1 Evaluation of Collective Behavior: Cooperative

自己組織化された集合行為の評価を Fig.1 に示す。Fig.1 には自己強化しない集団、すなわち、 $\Delta\theta = 0.0$  である場合の集合行為の評価も併せて示している。Fig.1(a) は集合行為の安定性を表しており、横軸は初期状態で選択肢  $S_1$  を選択するエージェントの割合  $p(0)$ 、縦軸は最終的に  $S_1$  を選択するエージェントの割合  $p^*$  を表してい

る。集合行為は利得パラメタの増加分  $\Delta\theta$  によらず初期値に依存している。そして、初期値  $p(0) = 0.5$  を境にして集合行為  $p^*$  は約 0 (ほぼ全員が  $S_2$  を選択する状態) または約 1 (ほぼ全員が  $S_1$  を選択する状態) に収束する。選好関係を自己組織化しない場合と同様である。

最終時点で得られる集合行為の効率性と平等性を Fig.1(b) に示す。横軸は集合行為の初期値、縦軸には最終時点での効率性を実線で平等性を点線で示している。選好関係を自己強化する場合には、自己強化しない場合よりも効率性と平等性がともに向上していることが分かる。また、効率性と平等性も  $\Delta\theta$  によらず初期値に依存し、初期値  $p(0)$  に偏りがある (小さいまたは大きい) 場合には効率性と平等性は約 1 となるが、0.5 付近では効率性と平等性がともに低い。

自己強化された選好関係の多様性も初期値  $p(0)$  に依存している。初期値に偏りがある場合には、均質なまま利得パラメタが自己強化しやすい。すなわち、初期値  $p(0)$  が小さい場合にはほぼ全員の利得パラメタ  $\theta_i$  は変化し次第に 1 になる。一方、初期値  $p(0)$  が大きい場合には、ほぼ全員の利得パラメタ  $\theta_i$  は 0 になる。ほとんどの者の利得パラメタが同じ (1 または 0) であるため、ほぼ全員が同じ選択をし効率性と平等性がともに高い。

ところが、初期値が 0.5 付近である場合には、 $\theta_i = 0$  をもつタイプ 1 と  $\theta_i = 1$  をもつタイプ 2 が半数ずつ占める集団になった。エージェントの配置を調べてみると同じタイプどうしの相互作用をしている<sup>4)</sup>。

#### 3.2 エージェントが移動する場合

エージェントが集団の中を移動する場合について考える。ここでも、利得パラメタの増加分を  $\Delta\theta$  をそれぞれ、0.5, 0.1, 0.01 と設定してシミュレーションを行った。移動ルールにおけるしきい値は  $\mu = 0.5$  とした。

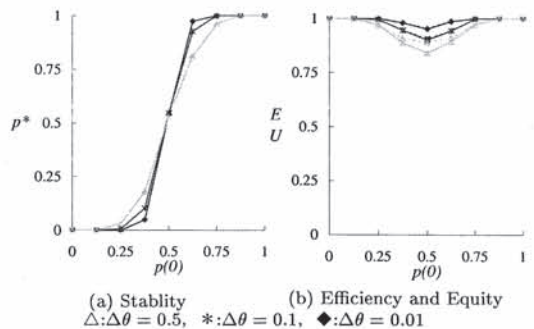


Fig.2 Evaluation of Collective Behavior: Cooperative, move

集合行為の評価を Fig.2 に示す。ここでも最終的な集合行為は集合行為の初期値に依存している。しかし、今回は利得パラメタの増加分  $\Delta\theta$  に依存して集合行為が異なる結果となっている。 $\Delta\theta$  が 0.5 のときには、エージェントが移動しないとときと類似した結果であるが、 $\Delta\theta$  が小さくなるに従って、集合行為は初期値に鋭敏になって

おり安定性は低下している。ところで、エージェントが移動しない場合には、例えば、 $p(0)$  が 0 付近であっても、 $\theta_i$  が 1 である集団の中にごく僅かに  $\theta_i$  が 0 である者が残存し多様性が維持されたが、移動する場合には少数派は淘汰され全員の利得パラメタが同じになった。

集合行為の効率性と平等性も利得パラメタの増加分  $\Delta\theta$  に依存している。 $\Delta\theta = 0.5$  の場合には、エージェントが移動しない場合と類似しているが、 $\Delta\theta$  が小さくなるに従って効率性と平等性がともに向上している。

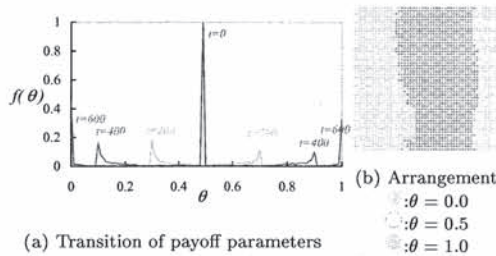


Fig.3 Self-reinforcement of preference: $p(0) = 0.5$ , move

集団の多様性はエージェントが移動しない場合と同様に初期値  $p(0)$  に依存している。例えば、 $p(0) = 0.5$  である場合に最終的に生まれる利得パラメタの密度関数 (Fig.3(a)) はエージェントが移動しない場合とほとんど同じである。しかし、エージェントの配置を調べてみると、同じタイプとの選択的な相互作用が促進されていることが分かる (Fig.3(b))。エージェントが移動することにより、各エージェントが適切な場所で相互作用ができるようになり、より望ましい相互作用が形成されている様子が分かる。そのため、効率性と平等性も移動しない場合に比べて向上している (Fig.2(b))。以上から、エージェントが移動する場合には、利得パラメタの増加分  $\Delta\theta$  を小さく設定する必要があることも分かる。

#### 4. 考察

Axelrod<sup>7)</sup> らは、多様性の創出には山岳帯などによる群生が必要であると示唆した。これは、本論文でエージェントが移動できない場合に各エージェントの周囲で選択肢が固定化し全体として内生的な内部属性である選好が多様化することに相当している。このように、エージェントの移動性と選好関係の多様性にはトレードオフがある。さらに、Sandholm<sup>8)</sup> は、多様性の創出するために速度の異なる二段階の学習プロセス及び集落の有効性を示した。本論文でも、適応ルールと選好関係の自己強化という二段階の学習を用い、 $\Delta\theta$  の値を小さくして自己強化を遅くすることによって学習速度に差を設けることによって、選好関係の多様性と選択的な相互作用が生まれやすくなることを示し、多様性創出に対する二段階の学習の重要性について示している。

一方、Schelling<sup>3)</sup> は、異なる言語を話す人々が言語に基づいて自発的にすみ分けすることを示した。言語と行

動を同一視しており観察可能な行動によるすみ分けといえる。本論文では、観察不可能な内部属性である選好に基づいてエージェントが自発的に適切な配置を実現できることを示している。以上から、自発的に多様性を創出するためには、まず、観察可能な行動についてのすみ分けを作りだし、さらに、自己強化によって選好 (観察不可能な内部属性) についてのすみ分けを生じさせることが有効である。そのために、エージェントを固定化して選択についてのすみ分けを作ることが必要であるが、エージェントが移動している場合には、自己強化の速度を遅くする必要がある。このように、選好関係の多様性と選択的な相互作用には相乗的な関係があり、エージェント個人と集団全体にとって効果の高い集合行為が自己組織化されることが分かる。

#### 5. おわりに

多数の均質なエージェントが自らの選好 (利得構造) を自己強化しながら二者択一の意思決定を繰り返す状況を扱った。集団全体の意思決定プロセスを安定性、効率性、そして、公平性の観点から評価する。また、最初は均質なエージェント集団がどのような多様なエージェント集団に進展するかを分析し、エージェントが移動せず同じ近傍のエージェントと繰り返し相互作用する場合には、同じ選好をもつ者どうしで多数の小規模なニッチを形成することを確認した。一方で、利得が低く不満なエージェントが移動する場合には、集団全体で二つの大規模なニッチを形成することを確認した。そして、一人ひとりが高いレベルの利得を獲得する安定した状態に収束することを示すことができた。

#### 6. 参考文献

- 1) Subrahmanian, V.S., Bonatti, P., Dix, J., Eiter, T., Kraus, S., Ozcan, F. and Ross, R.: *Heterogeneous Agent Systems*, MIT Press (2000).
- 2) 石田亨, 寺野隆雄, 鳥居大祐, 村上陽平: 社会シミュレーションと参加型デザイン, 人工知能学会誌, Vol.48, No.3, pp.271-277 (2007).
- 3) Schelling, T.: *Micromotives and Macrobehavior*, Norton (1978).
- 4) 岩永佐織, 生天目章: 局所的で異質な意思決定の集合現象, 情報処理学会論文誌, Vol.43, No.5, pp.1528-1537 (2002).
- 5) アクセルロッド R.: つきあい方の科学—バクテリアから国際関係まで, ミネルヴァ書房 (1998).
- 6) 岩永佐織, 生天目章: 選択的な相互作用と異質なエージェントの集合行為の評価, 電子情報通信学会論文誌, Vol.86-D1, No.8, pp.505-513 (2003).
- 7) アクセルロッド R., コーエン M.D.: 複雑系組織論—多様性・相互作用・淘汰のメカニズム, ダイアモンド社 (2003).
- 8) Sandholm, W.H.: Preference Evolution Two-Speed Dynamics and Rapid Social Change, *Review of Economic Dynamics*, Vol.4, pp.637-679 (2001).