

# 好意・感情を持つエージェント系の構築と集団形成の解析

成瀬 継太郎 [北海道大学大学院工学研究科/助手]

## 背景・目的

計算機やネットワーク技術の急速な発達、インターネットのようない新しいコミュニケーションメディアを提供し、地理的な距離には関係ない新しい種類のコミュニティが発生している。しかしこれらはメールなどの主として文字情報による通信であり、直接顔を会わせないため、感情のもつれなどによる集団の崩壊も時に発生する。問題は感情の相互作用があまりに複雑で未解明である点にある。感情の完全なモデル化は事実上不可能であるが、本研究開発では、構成論的な手法により感情の相互作用を一側面を明らかにすることを目的とする。

## 内容・方法

以下では、「感情」の要素として、感情と好意を導入する。

- 感情のモデルには様々なものが提案されているが、ここではプラチックにより提案されたモデル(基本感情として、喜び-悲しみ、怒り-恐れ)を採用する。
- 好意と感情のあいだの相互作用には、ニューカムのABXモデルを採用。このモデルは定性的なものであるため、いくつかの仮定を置きながら動力学系を構成する。
- 各個人を人工エージェントとして計算機上に実装し、そこで得られる集団の性質を議論する。エージェント数が少ない場合は数学的に安定性の解析を行い、多数のエージェントの場合は計算機によるシミュレーションにより実験的に明らかにする。
- マルチロボットの制御へ応用する。具体的には、問題の達成度などに応じてロボットの好意値、感情値が変化するようにする。シミュレーション実験を行い、ロボットが構成する集団の安定性の解析を行う。

## 結果・成果

エージェント $i$ は次の3項組で定義される。

$$Agent_i = (A_{ij}, E_i^k, \mathbf{P}_i). (1)$$

ここで $A_{ij}$ は $i$ から $j$ への好意値である。正の値は好意を、負の値は嫌悪を表す。値の大きさは、好意または嫌悪の大きさを表す。値の範囲は、 $(-1, 1)$ である。 $E_i^k$ は $i$ の $k$ 軸の感情値を表す。 $k=1$ のときを便宜的に喜び-悲しみ軸、 $k=2$ のとき怒り-恐れ軸とする。値の大きさは感情の大きさを表す。範囲は $(-1, 1)$ である。 $\mathbf{P}_i$ は $i$ の仮想空間での位置とする。以下では2次元空間を対象とする。

エージェントは仮想空間上のすべてのエージェントと相互作用するものとする。しかし、近い距離ほど大きく相互作用するため、ある一定距離以上離れたエージェントとの相互作用は事実上ないに等しい。エージェントは、相互作用により好意値と感情値を変化させ、好意値の値に従って移動方向と距離を決定

する。以下では、時間は離散化して考え、エージェント系の動力学は差分方程式として表現される。

$$E_i^k(t+1) = E_i^k(t) + \Delta E_i^k(t), (2)$$

$$\Delta E_i^k(t) = a_1(1 + E_i^k(t))(1 - E_i^k(t))\{A_{ij}(t)A_{ji}(t)\exp(-lr_{ij})\} - a_2\tau E_i^k(t). (3)$$

ここで $a_1$ は感情の大きさの変化率、 $r_{ij}$ はエージェント $i$ から $j$ への距離、 $l$ は距離に関する減衰定数、 $a_2$ は自然減衰率、 $\tau$ は自然減衰の時定数である。2つの感情の軸に対して、同じ式で更新されるが、その定数は軸によって違う値が設定可能とする。

一方、好意値は次の式で表現される。

$$A_{ij}(t+1) = A_{ij}(t) + \Delta A_{ij}(t) (4),$$

$$\Delta A_{ij}(t) = b(1 + A_{ij}(t))(1 - A_{ij}(t))(c_1 E_i^1(t) E_j^1(t) + c_2 E_i^2(t) E_j^2(t) + c_3 A_{ji}(t)). (5)$$

ここで $b, c_1, c_2, c_3$ は各項に関する係数である。

最後にエージェントの移動に関しては、エージェント $i$ は $j$ への好意値が正のときにその方向に引力が働き、負のときに斥力が働くものとする。さらに、各エージェントは近づきすぎないように斥力が働くものとする。したがって、次のように表現される。

$$\mathbf{P}_i(t+1) = \mathbf{P}_i(t) + \sum_j (d_1 A_{ij} \exp(-lr_{ij}) - d_2 \frac{1}{r_{ij}^2}) \mathbf{e}_{ij}. (6)$$

ここで $\mathbf{e}_{ij}$ はエージェント $i$ から $j$ 方向への長さ1のベクトルを表す。

本研究では、以上のように設定されたマルチエージェント系を定式化し、その安定性や形成される社会の性質を、理論的および計算機実験により明らかにしている。

- 感情に関するパラメータの設定により、多彩な社会が形成されることを計算機実験により明らかにした。具体的には、相互に好意を持ち合う安定した社会や、ちょっとした外乱で社会が散逸する不安定な社会、集団が構成される個体が点在する社会などが実現されることを確認した。
- 提案した系の限定的な場合の安定性について、理論的考察を行った。その結果、互いに好意を持っているエージェント間の関係は安定、互いに嫌悪しているエージェント間の関係は発散、その他の場合は不安定であることが明らかとなった。
- 応用例として、提案した感情エージェント系を、頻繁にデッドロックが発生しうる箱だし問題に適用した。提案手法の人工社会形成能力が、デッドロック解消に有用であることを示した。

## 今後の展望

マルチロボットの制御に関しては、従来の「全体の目的を個々に分解する」というアプローチから、感情の相互作用により「集団の中の数個のロボットの制御するだけで全体が制御可能」という新しい方法論が可能になる。これは、昨今のロボットの小型化と台数の増加を鑑みると、将来的に必要な技術である。またこれは物理的なロボットだけに限らず、ネットワーク上のエージェントへも応用可能である。今後は、本研究で得られた基礎的な系の性質を利用して、ロボット制御、エージェント制御などの実問題に応用していきたい。