

北海道大学工学部 ○石田崇 鈴木恵二 嘉数侑昇

要 旨

本研究では、進化ゲームを行うマルチエージェント環境において、各エージェントが自らの行動判断基準となる利得表、すなわち価値観をある程度自由に持つモデルについて考察する。このような環境を構築するにあたり、ここでは、非協力二人ゲームにおいて最も複雑性の高いといわれる囚人のジレンマを取り上げ、これを足がかりとして、異なる価値観を持つエージェント間の相互作用についての実験を行う。

1. はじめに

囚人のジレンマに関する研究はこれまで進化ゲームの立場から、適応的戦略による協調^[4]、戦略的複雑性の発現^[3]、または人工社会の形成といった観点から、数多くなされてきた。これらの研究においてゲームルールの題材の一つとしてしか扱われなかった囚人のジレンマを、マルチエージェント環境におけるエージェントの行動価値観とみなすことで、その新たな可能性を見出す。

実世界において、各人の価値観が統一されていないように、ここでもエージェントが自らの価値観を変動させることができるメカニズムを設計する。そして、進化ゲームを行う過程において、エージェントがその価値観をどのように変化させるかに着目した計算機実験を行う。このような研究を行う意義としては、行動規則の統一されていないエージェント間の協調行動や、機械学習において報酬規則すらも獲得可能なメカニズム構築などの将来展望の模索が挙げられる。

2. マルチエージェントモデルにおける囚人のジレンマ

例としてFig.1に示すように、荷物運搬作業を行う複数のエージェントモデルを考えてみる。エージェントは1step 毎に一つの荷物を運搬するかしないかの判断に迫られる。荷物を運搬する際にはエネルギーEを消費し、すべてのエージェントは運搬した荷物の数iに応じた報酬R*iを環境から平等に受け取る。このとき、エネルギーを消費したエージェントも荷物を運ばずに休んでいたエージェントも同じ報酬を受け取ることになる。

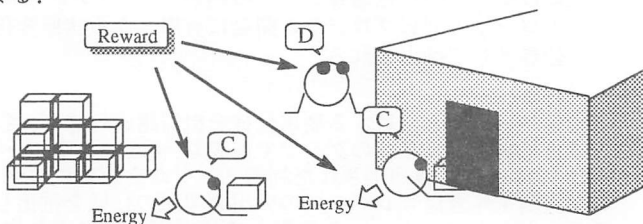


Fig.1 : Prisoner's Dilemma with Multi-Agent Model.

この事から、利己的なエージェントは次のように考える。「自分が働かずに他のエージェントたちが仕事をしてくれれば、自分のエネルギーを減らさずに報酬を受け取ることができる。しかし、他のエージェントも同じ

事を考えるので、一向に仕事ははかどらず、報酬ももらえない。皆で働いて平等な賃金をもらった方がはるかに良い事は分かるが、自分だけが働いてその報酬を分配されてしまうことだけは避けたい。」これが、「囚人のジレンマ」の価値観である。

ここで、荷物を運搬するという行動をC(協調: Cooperation)、運搬しないという行動をD(裏切り: Defection)と表現することにより、付加条件をなるべく取り除いた形で上記のようなマルチエージェントモデルを囚人のジレンマの定式に当てはめる事ができる。

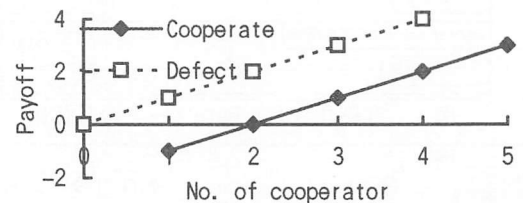


Fig.2 : Payoff function (E=2, R=1).

Fig.2はエージェントが選択した行動(C or D)に対して得られる利得を表したグラフである。ここで、横軸は荷物を運搬することを選択したエージェントの数であり、報酬はエネルギーで支払われることとする。例えば、3人のエージェントが働いた場合、全員に3のエネルギーが配分されるが、働いた3人は荷物の運搬のために2のエネルギーを消費しているため、差し引き1の利得となる。これに対し、働かなかった2人はエネルギー省費がないので3の利得となる。ここでDへの魅力が生ずるが、全員が働かなかったときの0の利得に対し、全員が働けば全員が3の利得を得ることができることから、ジレンマをひきおこしている。

繰り返し囚人のジレンマゲーム^[1](Iterated Prisoner's Dilemma : IPD)は、このような相互作用を繰り返し行い、その総得点を競うものである。

3. 価値観モデルの導入

前節で述べたように、囚人のジレンマはエージェントの価値観の一つである。各エージェントがある程度自由に価値観を持つということは、Fig.2に示す利得グラフを各人に独立に設定することであり、ここでは、傾きと切片のある制約条件の下に、可変となるよう設計した。その概念としては次のように考えられる。

エージェントの能力に個人差があり、荷物を運ぶのに消費するエネルギー量が異なる場合。また、受け取る単位報酬の価値が各エージェントにとってそれぞれ異なる場合。前者は利得グラフにおいて二本の直線の切片の差を表し、差が大きいほど大きなエネルギーが必要であることを示す。後者は二本のグラフの傾きを表し、傾きが大きいほど、そのエージェントにとって報酬の価値が大きいことを示す。ただし価値観が変動することによって、ジレンマには陥らずにどちらかの選択が常に優位となる場合もある。さらに、個人の価値観は自らの行動を反映して変動できるメカニズムが必要となる。それにより、相互作用の中で各エージェントが自分の価値観を獲得していく挙動を観察することができる。

このようなメカニズムを実現するために、ここでは、Karma と名づけられた変数を定義する。Karma は各エージェントが独立に持ち、協調するたびに 1 加え、裏切るたびに 1 を減ずるものとする。よって、Karma はその行為を積み重ねたものであり、そのエージェントの性格をあらわす数値ともいえる。全プレイヤーの Karma の合計と自分の Karma の比を算出し、そのプレイヤーが相対的に協調を多く取ってきたか否かで、二本の利得グラフの傾きと幅を決定する。ここでは挙動を単純化するため、協調が多いとき、グラフの傾きは減少し、幅を広げるものとし、協調が少ない(裏切りが多い)時には傾きは増大し、幅は狭くなるものとする。

4. 戦略の進化的適応

変動する価値観により、評価すべき利得が変化する環境においても、戦略を適応させることのみによって、エージェントが環境を正しく認識可能であることを確認するために、ここでは進化ゲームにおいて一般的に用いられる以下のような手法^[2]を用いる。

各エージェントは前回の相互作用において協調した人数に応じた次の行動を戦略としてバイナリコード化したストリングを複数個持つものとする。繰り返しゲームを行った後、その得点を適応度としてエージェントの遺伝子プール内で単純 GA 操作を行い、これを 1 世代とする。これにより、エージェントは価値観だけでなく、戦略も自立的に振る舞うことができ、価値観という利得の変動と、他のエージェントの戦略の変化にも適応できる能力を持つと考えられる。

5. 計算機実験

実験は、5 人のエージェントが 20 回で 1 世代の相互作用(IPD)を行うものとし、GA のパラメータは交差率 0.5、突然変異率 0.01 とし、エージェントの有する戦略 string は一律 20 個とした。

Fig.3,4 は、上記の設定で 1000 世代の相互作用を観察したものである。Fig.3 に示すように、Karma の値は 100 世代頃、戦略が適応した結果協調が発現するために上昇をはじめ。それに同期して、Fig.4 に示すようにジレンマの価値観(Fig.1 の利得グラフ)からの各エージェントの価値観のずれが一斉に 0 に収束する。また、Karma の値が減少をはじめると、価値観も次第にバラバラになって行くことが観察される。

つまり、協調することで価値観は統一され、裏切ることでその統一は崩れるといえる。この事は、協調行動を獲得するためには、エージェントの価値観を「囚人のジレンマ」という価値観に統一する必要があるととらえることもできる。

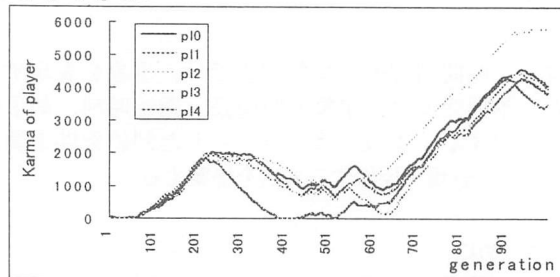


Fig.3 : Karma value of each agent.

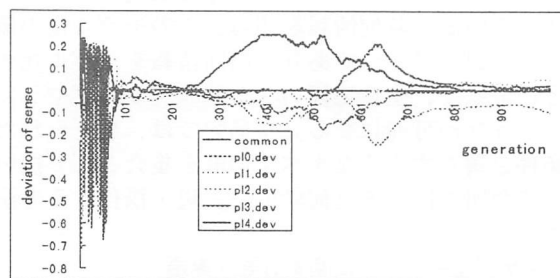


Fig.4 : Deviation from Prisoner's Dilemma.

また、エージェントの価値観とは別に、環境の常識(common sense)が変化する(実際に支払われる報酬や、運ぶべき荷物の重さが変化する)ような環境において、common sense を把握しているエージェントと、自らの価値観にしたがって行動するエージェントとの相互作用により、常識を知らされないエージェントが common sense の変化を、自らの価値観をずらすことで感知するといった結果も得られた。この事は、環境から直接報酬を受け取らなくても、Karma という変数を通じることで、環境が評価する行動を相互作用のみから感知することの可能性を示している。

6. まとめ

ゲーム理論の題材の一つである囚人のジレンマをマルチエージェント環境における価値観のモデルとして説明を試みた。また、ある制約条件の下にエージェントが価値観を自由に持つモデルを設計し、相互作用の中で、価値観変動の挙動に関する計算機実験を行った。

参考文献

- [1] Axelrod, Robert. "The evolution of cooperation" Basic Books, 1984.
- [2] Bankes, Steve. : "Exploring the Foundations of Artificial Societies: Experiments in Evolving Solutions to Iterated N-player Prisoner's Dilemma" Artificial LIFE IV 337-342. 1995
- [3] Ikegami, Takashi. : "From genetic evolution to emergence of game strategies" Physica D 75 310-327. 1994
- [4] Kristian Lindgren : "Evolutionary Phenomena in Simple Dynamics" Artificial LIFE II 295-312. 1991