

## Java KK-MAS を用いた対戦型「囚人のジレンマ」ゲーム

鈴木一敏<sup>1</sup>

KK-MAS は、その名の通り、複数のエージェントの相互作用を再現するのに便利な環境である。このシミュレータを用いた研究の大部分は、プログラムされたエージェント同士の相互作用による秩序を生みだし、それを分析する事に主眼を置いている。たとえば、過去第1回から第5回の KK-MAS コンペティションで発表されたモデルは60ほどあるが、研究部門・アイデア部門を問わず、ほぼ全てがこうしたタイプのモデルである<sup>2</sup>。

もちろん、こうしたモデルは、複数のエージェントによって創り出されている秩序をより良く知る上で役立っている。しかし、教育分野で利用する場合など、既存のモデル(たとえば分居モデルや囚人のジレンマ)の秩序を知って貰うという目的の上では、上記以外にも有効な方法があるだろう。すなわち、プログラムされたエージェント同士の相互作用の中に、人間がエージェントとして参加する方法である。

一般にシミュレーションといわれるものには、フライトシミュレータのように人間が参加するものもある。参加する事による学習を目的としたゲームは、近年、「シリアスゲーム」と呼ばれ、にわかに注目を浴びている<sup>3</sup>。たとえば、ゲームを用いて消防士などに対する緊急時のトレーニングを行う hazmat や<sup>4</sup>、米国陸軍の訓練内容 PR を目的とした America's Army など<sup>5</sup>、近年実用化されている例もある。

社会秩序の研究・教育においても、「人間が参加する」というコンセプトは役立つ可能性があるし、そうした可能性をあらかじめ排除してしまう必要はないだろう。今後マルチエージェントシミュレーションが普及し教育分野での利用が増加すれば、人間がコンピュータの中の秩序を体感できるモデルの利用価値は高くなる可能性が高い。たとえば、本稿で紹介する「囚人のジレンマ」は、多くのゲーム理論の入門書で、例として活用されている。その名前の元となっている逸話(数学者のアルバート タッカーによる)によってかなり理解しやすくなっているが、繰り返し回数の変化や、通信の可否、賞罰の強弱など様々なバリエーションが可能であり、その奥が深さゆえ、説明は長く複雑になりかねない。しかし、もし、仮想刑務所に放り込んでその選択を自ら体験させたのならば、たとえ子供であって

---

本稿で紹介する対戦型囚人のジレンマモデルは、山影研究室(<http://citrus.c.u-tokyo.ac.jp/>)にて配布している。JAVA 版 KK-MAS については、2005年8月現在フィールドテストが行われている。詳しくは構造計画研究所(<http://www.kke.co.jp/it/mas/index.html>)参照。

<sup>1</sup> 東京大学大学院総合文化研究科

<sup>2</sup> 構造計画研究所の MAS コミュニティ(<http://www2.kke.co.jp/mas/>)参照。例外は2~3以下である。

<sup>3</sup> <http://www.seriousgames.org/>

<sup>4</sup> <http://www.etc.cmu.edu/projects/hazmat/>

<sup>5</sup> <http://www.americasarmy.com/>

も、ゲームの構造の細かな部分までよりよく理解するだろう。人間が参加するマルチエージェント・シミュレーションは、社会秩序の教育にも役立つ筈である。

また、教育目的以外でも、人間自体の行動の研究や、そこから発展してより人間に近い行動をするプログラムを作成するといった用途も考えられる。たとえば、モデル内の秩序に参加している人間がどのように決断するかを分析する事で、より現実的な避難モデルを作成できるかもしれない。

現在、JAVA 版 KK-MAS には複数の PC を用いて多数の人間が参加できる通信機能も開発されつつあるが、人間が参加するモデルの実例は未だほとんど無い。そこで本稿は、人間が参加できるゲームの一例として、「囚人のジレンマ」トーナメントのモデルを作成して紹介する。

### 「囚人のジレンマ」のシミュレーション

ロバート・アクセルロッドの「囚人のジレンマ」トーナメントは<sup>6</sup>、社会科学におけるコンピュータシミュレーションの分野では、最も基本的な古典の一つに挙げられる。この実験はもともと、様々な分野の専門家に応募してもらった 14 にランダムを加えた 15 の戦略による総当たりで、200 回の「繰り返し囚人のジレンマ」を行うものであった。その結果、最も単純なルールである「しっぺ返し戦略(Tit-for-Tat 以下 TFT、アナトル・ラパポートが考案)」が好成績を収めた。また、その結果を踏まえて行われた第二回トーナメントにおいても、TFT は最高点を出している。

実験の有名さゆえ、そこで用いられた戦略のなかには、マルチエージェントシミュレーションを行う者には良く知られるようになったものもある。たとえば、Freedman(一度でも裏切られると二度と協力しない)や Joss(初回協力し、相手が裏切ったら 100%、協力なら 10% の確率で裏切る)といった戦略である。しかし、このような戦略を知識としては知っていても、これらと実際に対戦した事がある者はどれほどいるだろうか。この実験を良く知っている研究者なら臨機応変に対応して TFT 以上の成績を残せるのだろうか。残せるとしたら彼はどの様な戦略を用いているのだろうか。一方、このゲームや戦略を全く知らない人はどの様に反応するだろうか。様々な戦略にどういう感情を持ち、どう反応するだろうか。興味は尽きない。そこで、コンピュータのエージェント達に混じって「囚人のジレンマ (Prisoner's Dilemma)」をするゲームを作成した。

モデルは、人間のプレイヤーが手を入力すると、プログラムされたエージェントが自らの戦略に基づき手を選択し、結果が計算されるように設計した。具体的には、コントロー

---

<sup>6</sup> Axelrod, Robert の “Effective Choice in the Prisoner’s Dilemma”, Journal of Conflict Resolution, vol.24. (1980)、ならびに、The Evolution of Cooperation, Basic Books.(松田裕之 訳、『つきあひ方の科学』、ミネルヴァ書房、1988 年)参照。

	協力	裏切り		協力	裏切り
協力	R, R	S, T	協力	3, 3	0, 5
裏切り	T, S	P, P	裏切り	5, 0	1, 1

図表 1 「囚人のジレンマ」と本モデルのペイオフマトリクス

ルパネル上のボタンもしくはキーボードの C および D キーからプレイヤーの入力が無い場合には、そのステップではルールが実行されず空回しされるようにした。ステップ間にごくわずかな実行ウェイトを設定すれば、ステップを空回ししてもソフトウェアの反応が鈍くなることは無い。

ペイオフはツリーで設定できるようになっている。デフォルトは上述のアクセルロッドの実験と同じであり、相互に協力した場合(Reward)に 3 点、自分だけが協力した場合(Sucker)に 0 点、相手だけが協力した場合(Temptation)に 5 点、両方が裏切った場合(Punishment)には 1 点が与えられる。このペイオフは、 $S < P < R < T$  となっており「囚人のジレンマ」の要件を満たしている。また  $2R > T + S$  を満たし、(裏切り, 協力) (協力, 裏切り) (裏切り, 協力)・・・というパターンを繰り返すことによって双方が得られる平均点数が互いに協力した場合より低く抑えられており、標準的囚人のジレンマと呼ばれる<sup>7</sup>。もちろん、実験や体験では、こうした条件にとらわれる必要はない。裏切りの誘因の強さを変化させるばかりでなく、「チキン」など全く別のゲームにすることもできるし、ペイオフを動的に変化させる事も簡単なルールの追加により可能である。単純なゲームばかりでなく、より現実的なモデルを構築して、それに参加できる点は、インタラクティブなマルチエージェント・シミュレーションの一つの利点であろう。

このモデルでは、同じ相手と 50 回の選択を繰り返し、その後、相手をランダムに組み替えて繰り返す。<sup>8</sup> コンピュータのエージェントは、以下の 10 の戦略の中から初回にランダムに選ばせることも、個別エージェントに対してコントロールパネルから指定する事もできる。

- "Random" : ランダムに協力または裏切りを選択する
- "allD" : 全て裏切る
- "allC" : 全て協力する
- "TFT" : 初回協力し、その後は前回の相手の手をまねる
- "TFTT" : 初回協力し、相手が二回連続で裏切ったら一度裏切りかえす
- "rTFT" : 初回裏切り、その後は前回の相手の手をまねる

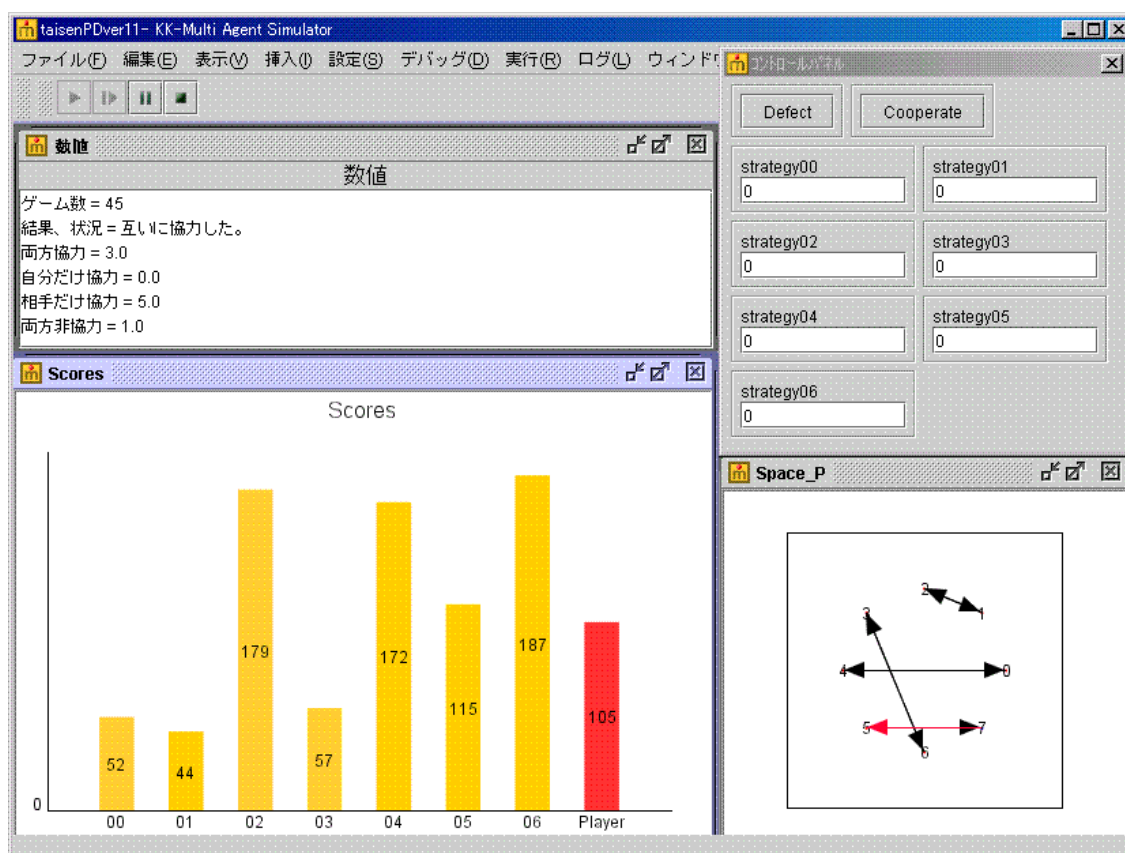
<sup>7</sup> 鈴木光男 著、『新ゲーム理論』、勁草書房、1994 年。

<sup>8</sup> アクセルロッドのシミュレーションでは 200 回繰り返していたが、被験者の忍耐を超えないように回数を減らした。

- "Friedman" : 初回協力し、一度でも裏切られたら、裏切り続ける
- "Davis" : 最初の 5 回協力し、その時相手に裏切られなければずっと協力的
- "Joss" : 初回協力し、相手が裏切ったら 100%、協力なら 10%の確率で裏切る
- "Tullock" : まず 5 回協力し、その時の相手の裏切り率より 10%高い確率で裏切る

図表 2 の右下のウィンドウにある矢印は組み合わせを、左下のグラフは合計点数を表している。得点は相手が組み替えられても持ち越されるようになっている。左上のウィンドウには、ペイオフや現在の繰り返し回数、そして、前回の状況などが表示される。これらの表示を制御する事で、実験や体験の条件(相手の戦略の開示の有無、ペイオフの開示の有無、一度対戦した相手を特定できるかどうか等)を変更する事ができる。

難易度はエージェントの戦略の組み合わせに大きく左右されるが、何度か実際にプレイしてみると、一位を維持するのは平均的に難しいことが分かるだろう。このゲームでエージェントよりも高得点を出すためには、All C や Random、TFTT といった戦略を相手に搾取するのが有効なので、対戦相手の戦略を推測することが重要になってくる。しかし、反応を見るために D を出すことが相手の協力する確率に影響を与える戦略(Davis や Tullock)のほ



図表 2 実行画面

か、一度でも D を出すと最後まで D を出し続ける Friedman が混じっている可能性もある。現状では個々のエージェントの戦略は固定されているうえ、50 回の対戦が終わるとそれまでの相手がどんな戦略だったのかを表示するようにしているので、何度か同じ相手と対戦すれば予想が付くようになっている。こうした表示をなくせばゲームはより難しくなるだろう。また、アクセルロッドの実験よりも戦略を少なくしてあるが(15 10)、より複雑な戦略の数が増加すれば、相手エージェントの戦略を見抜きにくくなり、難易度が上がるものと思われる。今後、ゲーム理論を理解している学生や大学院生を被験者として、単純な戦略に対して、どの程度のスコアを上げられるのかを様々な条件下で実験してみたい。

おわりに

ゲームを通してエージェントの秩序を体験する事に一定の学習効果があることは容易に想像が付く。そして、この学習効果は「囚人のジレンマ」モデルに限られたものではない。勢力均衡や国際交渉といったモデルにエージェントとして参加する事で、そのモデルの相互作用を観察者の視点から眺めるだけでなく、ミクロな視点から体験できるだろう。モデルの設計者にとってすら、意外な発見があるかもしれない。

また、商取引や避難のシミュレーションなどでは、参加した人間プレイヤーの判断をより深く観察する事で、より人間的な振る舞いをするエージェントを作る事が可能だろう。マルチエージェント・シミュレーションでは、エージェントルール小さな違いが、創り出される秩序に大きく影響する事も少なくない。したがって、既存のエージェントと、人間的なエージェントの創り出す秩序に違いがあるかどうかは重要な点だと言えよう。

こうした利用方法に加えて、現在開発されつつあるネットワーク機能を使えば、人間同士の対戦も実現できる。そうなれば、多人数囚人のジレンマなど、複数の人間を含む大規模な実験も容易になり、より利用範囲が広がるだろう。