

エージェントベースシミュレーションを用いた ムカデゲームにおける被験者の意思決定および学習に関する分析

菅生 雄矢*・林田 智弘**・西崎 一郎***

Mechanism of decision making and learning of human subjects
in centipede games by using agent-based simulation

Yuya SUGEO, Tomohiro HAYASHIDA and Ichiro NISHIZAKI

Abstract: This paper constructs an agent-based simulation model for behavioral analysis in a laboratory experiments of the centipede game using the human subjects. The centipede game is one of the games in extended forms which multiple players make decision in a predetermined order. In any game in extended forms, it is well known that the sub-game perfect equilibrium is a powerful solution concept to predict the behavior in the game. However, some experimental results are reported such that the equilibrium does not predict the behavior of human subjects, the centipede game is one of them. In the theoretical analysis approaches, it is assumed that players make decision are rational, though the decision making of human are not always rational but they make decision with the mechanism of trial and error. In this paper, a simulation model for behavioral analysis in the centipede game using artificial adaptive agents.

Key words: centipede game, agent-based simulation, decision making, neural network

1. はじめに

複数のプレイヤーが順番に意思決定を行う逐次手番ゲームを展開型ゲームといい、サブゲーム完全均衡とよばれる解概念を用いて、人間の行動を解析的に予測できることが知られている。しかし、ムカデゲーム³⁾や最終提案ゲーム⁴⁾など、いくつかの展開型ゲームの被験者実験においてサブゲーム完全均衡による予測と一致しないことが報告されている。

本論文では、ムカデゲームにおける被験者の意思決定や学習に関する分析を行い、サブゲーム完全均衡と被験者実験の結果が一致しない原因を明らかにすることを目的とする。ここで、均衡理論ではプレイヤーは利己的で合理的な意思決定をすると仮定されているが、これが現実の人間の意思決定構造を必ずしも厳密に表現できているとはいえないことが、サブゲーム完全均衡と被験者の行動が一致しない理由として考えられる。このような観点から、従来研究では静的確率選択モデルと動的確率選択モデルによる行動分析が行われた。静的確率選択モデル

は人間の意思決定におけるエラーを考慮したモデルであるが、被験者の時間的な行動の変化を説明できない。動的確率選択モデルは被験者が直前の行動分布に基づいて確率的に意思決定していると仮定し、プレイヤーの行動選択確率を更新していくモデルであり、被験者の時間的な行動の変化を説明できているが、ゲームの結果としてプレイヤーが得られる利得が考慮されていないなど、問題点もある。以上のことから、確率的な行動規則は、部分的には被験者の行動を説明できているが、必ずしも十分であるとはいえない。人間は試行錯誤的で相手の行動や自身の経験を参考にしながら意思決定を行っていると考えられる。

ニューラルネットワーク (NN) は脳の構成要素であるニューロン、およびニューロン間のシナプス結合を介した相互作用をモデル化したものであり、構造が単純であるにもかかわらず、人間の脳の情報処理機構に驚くほど類似した性質を多く持っており、エージェントが複数の情報を相互作用した意思決定を行うことが可能である。また、Holland and Miller¹⁾ は多くのシステムは複雑な

* 広島大学 大学院工学研究科 システムサイバネティクス専攻 博士課程前期 学生

** 広島大学 大学院工学研究院 電気電子システム数理部門 助教

*** 広島大学 大学院工学研究院 電気電子システム数理部門 教授

適応型システムであると考え、これらをモデル化するためには人工適応型エージェントを適用した分析が有効であることを指摘している。本論文では、人間は試行錯誤的で相手の行動や自身の経験を参考にしながら意思決定を行っていると考え、Nishizaki⁵⁾により提案された NN と遺伝的アルゴリズム (GA) に基づいた意思決定と学習機構を持つ人工適応型エージェントを用いたシミュレーション分析を行う。NN と GA を用いたシミュレーションシステムにより、試行錯誤的な人間の意思決定構造を実装できる。NN と GA を用いた適応型モデルについては、Axelrod²⁾ や林田ら⁹⁾ など多くの研究によりその妥当性が示されている。本論文では、NN と GA を用いた人工適応型エージェントを用いたシミュレーション実験を行うことで、ムカデゲームに関する被験者実験における被験者の意思決定と学習に関する分析を行う。

本論文の構成は以下の通りである。2章ではムカデゲームについて、3章ではムカデゲームに関する被験者実験について、4章では人工適応型エージェントを用いたシミュレーションモデルについて、5章ではシミュレーション実験の結果について説明を行い、6章で本論文の成果をまとめる。

2. ムカデゲーム

ムカデゲームとは、複数のプレイヤーが順番に意思決定する展開型ゲームの1つである。3人のプレイヤーでプレイされるムカデゲームのゲームの木を図??に示す。

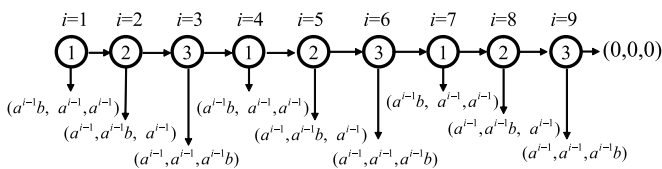


図 1: 3 人ムカデゲームのゲームの木

図??に示される i はノード番号、ノードの中の数字は意思決定を行うプレイヤーの番号を表す。意思決定ノードにおけるプレイヤーの戦略は *Down* もしくは *Right* であり、*Down* を選択した場合、各プレイヤーはそれぞれのノードに割り当てられた利得を受け取り、ゲームが終了する。*Right* を選択すると、次の順番のプレイヤーが意思決定する。ムカデゲームにおいて、プレイヤーが *Down* を選択した場合、そのノードに割り当てられた利得を受諾し、*Right* を選択した場合、その利得を拒否したと解釈できる。プレイヤーが *Down* を選択することで得られる利得が図??の各ノードの下に示されて

おり、左からプレイヤー 1, 2, 3 の利得である。多くのムカデゲームでは、*Down* を選択したプレイヤーが全てのプレイヤーのうち最も高い利得を得られるとされ、 $a > 1, b > 1$ とされる。

展開型ゲームにおいて、あるノードから先を独立した 1 つの展開型ゲームとみなしたゲームをサブゲームといい、すべてのサブゲームで各プレイヤーが自らの得る利得が最大となるような最適応答をしている場合、その戦略集合をサブゲーム完全均衡といい、後ろ向き帰納法により求められる。図??で表される 3 人ムカデゲームでは、プレイヤー 1 がノード 1 で *Down* を選択して、各プレイヤーが利得 $(a^0b, a^0, a^0) = (b, 1, 1)$ を得ることがサブゲーム完全均衡となる。一方、ノード 1 でゲームを終了することに比べて、ノード 7, 8, 9 など、ノード番号が大きいノードでゲームを終了するほうがすべてのプレイヤーの利得は大きくなる。サブゲーム完全均衡はプレイヤーが自己の利得を最大化するための合理的な意思決定の集合であることを考えると、サブゲーム完全均衡の実現は協調に失敗しており、大きい番号のノードでゲームを終了することは、協調に成功しているといえる。

3. 被験者実験

Rapoport *et al.*³⁾ は $a = 2, b = 10$ として 3 人ムカデゲームに関する被験者実験を行った。被験者は 15 人ごとの 4 つのグループに分けられ、各グループで 60 回のゲームが行われた。1 回のゲームごとに、15 人がランダムに 3 人ずつの 5 つの組に分けられ、それぞれランダムに 1, 2, 3 の番号をプレイヤー番号として割り当てられる。被験者実験の結果として、ノード 1, 2, 3 における *Down* が選択された割合の 10 期の移動平均を図??-??に示す。

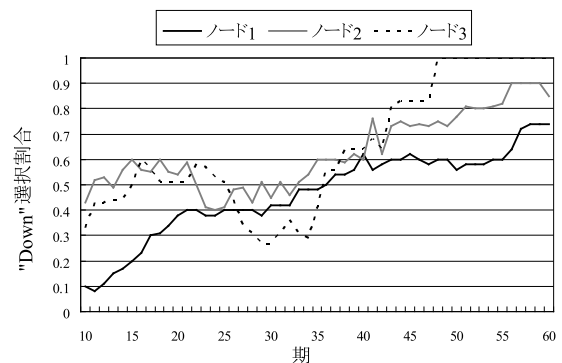


図 2: グループ 1 の *down* 選択確率の移動平均

横軸はゲームの期数、縦軸は各ノードにおける *Down* 選択割合を表している。実験の結果として、60 期間の

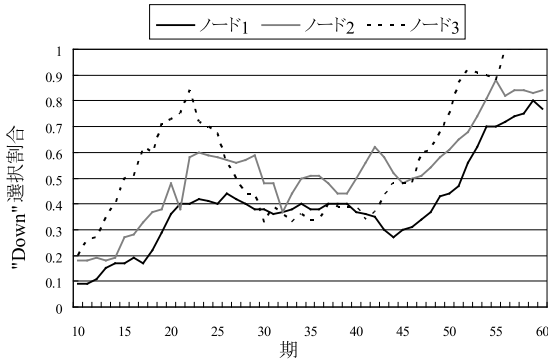


図 3: グループ 2 の down 選択確率の移動平均

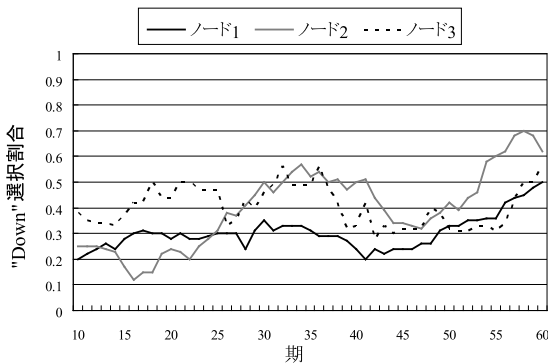


図 4: グループ 3 の down 選択確率の移動平均

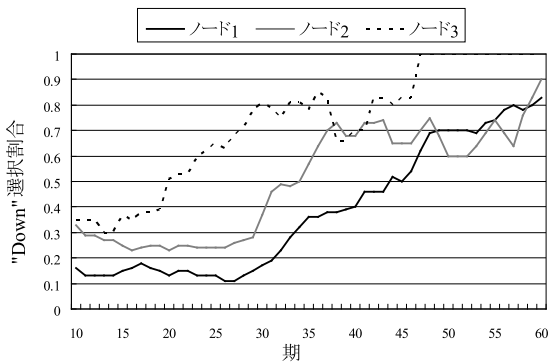


図 5: グループ 4 の down 選択確率の移動平均

実験における各グループのサブゲーム完全均衡が実現する割合、すなわちノード 1 でプレイヤーが *Down* を選択する割合はそれぞれ 46.8%, 39.3%, 30.3%, 40.7% となり、また、ゲームを繰り返し行くと、グループ 1, 2, 4 ではノード 1, 2, 3 で *Down* を選択する割合が高くなりサブゲーム完全均衡が実現する割合が高くなる。グループ 3 では均衡が実現する割合は他と比べてが低い結果となった。

Rapoport *et al.* は上述の被験者実験において、サブゲーム完全均衡の実現割合が低くなる理由を以下の確率的選択モデルにより、説明することを試みた。

3.1 確率的選択モデルによる被験者の行動分析

静的確率選択モデルと動的確率選択モデルに基づき、ムカデゲームにおける被験者実験³⁾の結果の説明が試みられている。

静的確率選択モデルとして、IPP モデル⁷⁾と AQRE⁸⁾が用いられ、これらのモデルでは、プレイヤーの意思決定におけるランダム性を合理的な意思決定におけるエラー行動と解釈し、その度合いが 1 つのパラメータによって表現されている。被験者実験の結果を用いた最尤推定法によるパラメータ値の推定によりモデルの妥当性が示されている。しかし、静的確率選択モデルでは、60 期間の実験における各ノードでの *Down* 選択の平均的な割合を説明することはできたが、被験者の行動の時間的な変化が説明されていない。一方、Rapoport *et al.*³⁾により提案された動的確率選択モデルは、プレイヤーの混合戦略を複数のパラメータに基づいて更新するモデルであり、これにより、静的確率選択モデルではできなかった被験者の行動の時間的な変化を再現することに成功した。しかし、動的確率選択モデルではゲームで得られる利得が考慮されておらず、また、混合戦略に関する確率分布の初期値は被験者実験のデータを用いていることから、これらの意思決定モデルで被験者実験の結果を十分に説明できているとはいえない。

4. シミュレーションモデル

本論文では、Rapoport *et al.*³⁾の被験者実験における被験者の行動分析のためのシミュレーションモデルを構築する。一般に、ゲーム論的状况において人間が意思決定するとき、自らの経験に基づき試行錯誤的に意思決定していると考えられるため、均衡理論で仮定されているような利得最大化のための合理的な意思決定モデルや確率的選択モデルのような意思決定におけるノイズやエラーを考慮したモデルだけでは、十分に説明できているとはいえない。本論文では、試行錯誤的な人間の意思

決定構造を模倣することのできる意思決定および学習機構を備えた人工適応型エージェントを用いたシミュレーションモデルを構築する。

適応的な学習システムとして、クラシファイアシステムや強化学習、NN などがある。クラシファイアシステムでは、エージェントは「もし～ならば～する」という条件部と出力部からなる行動ルールを複数保持し、もし条件が一致するルールがあれば行動部を実行する。そして、実行したことで良い結果が出ればそのルールの信頼度を上げ、逆に結果が悪ければ信頼度を下げ、信頼度が極端に低いルールは淘汰される。このようにして環境特性に適合したルールを残すシステムである。強化学習とは、試行錯誤を通じて環境に適應する学習制御の枠組みであり、環境からの報酬を手がかりに学習を行い、環境との相互作用を通じて最適または合理的な選択を学習するシステムである。NN は、その構造が単純であり、過去の自身の経験や相手の行動などの複数の基準を参照しながらエージェントが意思決定することができ、また、GA を用いて NN の構造を進化させることで、エージェントが試行錯誤的に学習を行うシステムを構築することができる。本論文では NN と GA に基づいたシステムを採用する。NN の入力値として現在のノード番号、自身のプレイヤー番号、前回のゲームにおける自分のプレイヤー番号と行動、利得、過去 t 期間にエージェントが獲得した利得と過去 t 期間のエージェントが選択した行動の重み付平均、現在のノードで *Down* を選択した場合と *Right* を選択した場合に得られるそれぞれの期待利得の 9 つの情報を与える。過去の情報の重み付平均は $\delta \in (0, 1)$ を割引率、 x_t は t 期前の利得や行動を表す数値とし、 $\sum_{t=1}^m x_t \delta^{t-1} / \sum_{t=1}^m \delta^{t-1}$ を NN の入力として与える。また、NN の出力値を R, D とし、エージェントは $D \geq R$ のとき *Down*、 $D < R$ のとき *Right* を意思決定として選択する。

4.1 シミュレーションモデルの概要

本研究では、3 節で述べた実験に対応したシミュレーション実験を行った。

- 手順 1: 初期集団 $N = 100 \times 3$ (プレイヤー数) を生成する。
- 手順 2: 集団からランダムに m 人ずつ選び、 N 個のグループを生成する。
- 手順 3: 各グループのエージェントにプレイヤー番号をランダムに割り当てる。
- 手順 4: 各グループにおいて、ムカデゲームを行う (N 回)。この時、各エージェントは毎回 NN に基づ

く意思決定を行う。

- 手順 5: 得られた利得をエージェントの適合度として、プレイヤー番号ごとに GA を用いてエージェントの学習を行う。
- 手順 6: $t = T$ ならば終了。そうでなければ、 $t := t + 1$ として手順 2 にもどる。

ただし、GA はルーレット選択、1 点交叉、突然変異、エリート保存則により行う。図 6 に、シミュレーションモデルのフローチャートを示す。

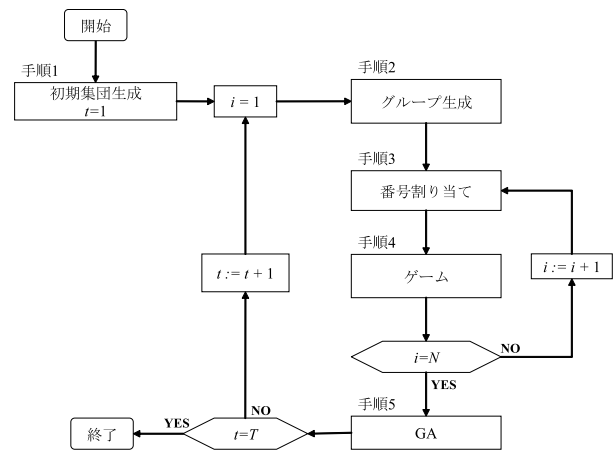


図 6: シミュレーションモデルの概要

本論文では、ムカデゲームにおけるプレイヤー数の影響を分析するために、ゲームを行うプレイヤー数を 2-5 と変動させた実験を行う。ただし、プレイヤーの得られる利得は最大利得が 3 人ムカデゲームの最大利得と同様に 2560 となるように、 a, b の値を調整する。

5. シミュレーション結果

本論文では、エージェント数 300、交叉確率 $P_c = 0.3$ 、突然変異確率 $P_m = 0.005$ 、世代間ギャップ $G = 0.7$ 、過去の情報に対する割引率 $\delta = 0.6$ としてシミュレーション実験を行った。

5.1 被験者実験に対応したシミュレーション実験

事前実験の結果、シミュレーション実験の 2 期を被験者実験の 1 期に対応させて、分析を行うことが適切であるとした。本論文では、同じ条件のシミュレーション実験をそれぞれ 100 回ずつ行った。

シミュレーション結果として、最初はノード 7-9 でゲームが終了しており、協調が成功しているが、ゲームを繰り返し行うことで、エージェントの学習が進み、サブゲーム完全均衡が実現した。均衡への収束について、

次の3つに分類することができる。1つ目は、ノード7-9でゲームが終了する協調関係が継続することなく、すぐにサブゲーム完全均衡に収束する結果である。2つ目は、ノード7-9でゲームが終了する協調関係が長期間継続し、緩やかにサブゲーム完全均衡に収束する結果である。3つ目は、緩やかに低いノード番号でゲームが終了することが多くなり、サブゲーム完全均衡に収束しかけるが、もう一度ノード7-9でゲームが終了する協調関係が構築された後に、サブゲーム完全均衡に収束する結果である。

シミュレーション結果の一部で、図??-??に示した被験者実験の4つのグループに近い結果を、実験結果より得ることができた。被験者実験の4つの行動パターンに対応するシミュレーション結果を図??と同じ形式で図??-??に示す。図??-??は、被験者実験との比較のため、シミュレーション期間の2期間ごとの平均値を横軸の1期に対応させている。すなわち、図??-??はシミュレーションの360-460期の結果を被験者実験に対応する形式で示している。このとき、シミュレーション実験の1-359期はエージェントの事前学習期間と解釈する。

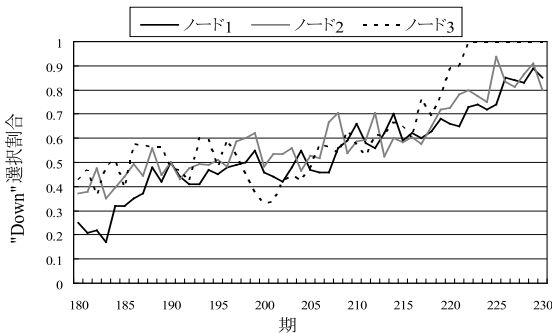


図 7: グループ 1 に対応した Down 選択確率の推移

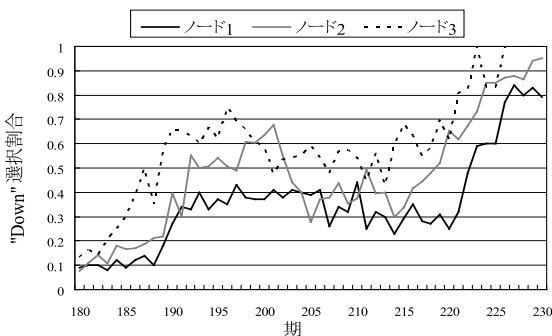


図 8: グループ 2 に対応した Down 選択確率の推移

図??-??より、グループ 1, 2, 4 ではノード 1, 2, 3 で Down 選択する割合がゲームを繰り返し行うと高く

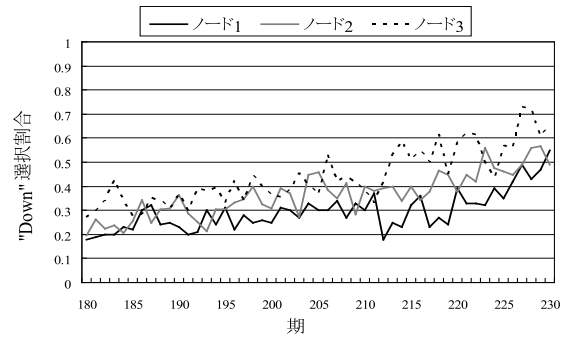


図 9: グループ 3 に対応した Down 選択確率の推移

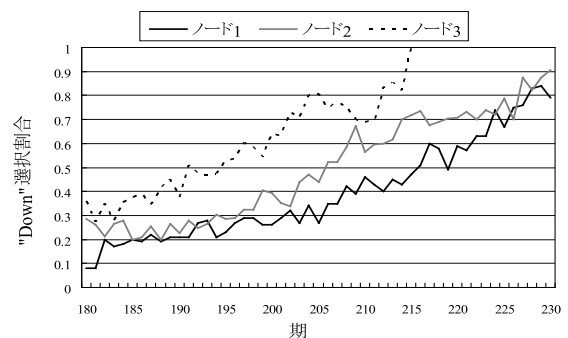


図 10: グループ 4 に対応した Down 選択確率の推移

なる傾向がある。また、グループ 1 はゲームを繰り返し行うと割合が単調増加しており、グループ 2 ではグラフが盆地型をしており割合の変動が大きく、グループ 4 ではノード間の割合の差があるといった各グループのグラフの特徴が観測できる。また、図??より、ゲームを繰り返し行っても確率が上昇しない行動パターンであるグループ 3 に対応する結果も得ることができた。

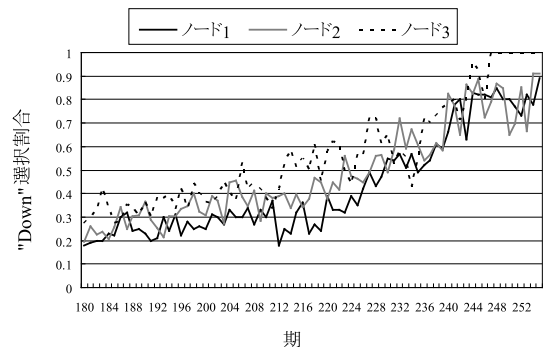


図 11: グループ 3 のシミュレーション結果

図??は、グループ 3 に対応したシミュレーションである図??に示される結果の 3000 期までの結果を示している。

図??より、グループ 3 に対応した結果も、ゲームを繰り返すことで他のグループと同様にノード 1, 2, 3 での *Down* 選択を選択する割合が上昇しており、協調が成功しなくなっている。次に、サブゲーム完全均衡が実現する割合の 100 試行の平均値の推移を図??に示す。グラフの縦軸はサブゲーム完全均衡の実現、すなわち、ノード 1 で *Down* を選択した割合を表しており、横軸はシミュレーションの期数を表している。

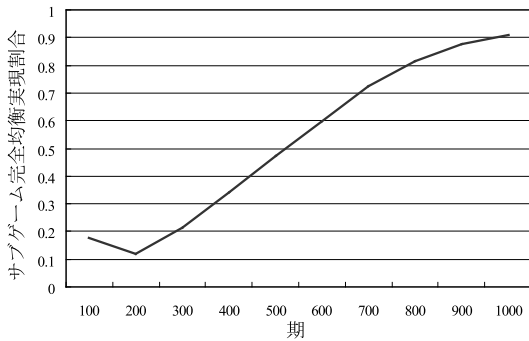


図 12: サブゲーム完全均衡の実現割合の推移

図??より、ゲームを繰り返し行うことでサブゲーム完全均衡が実現する割合は上昇し、901-1000 期の 100 期間では、サブゲーム完全均衡が実現する割合の平均が 90% 以上となっている。このことより、長期間ゲームを繰り返すことで、サブゲーム完全均衡が高い割合で実現していることがわかる。

図??-??に 1000 期間のシミュレーションにおける終了ノードの割合の推移をグラフに示す。

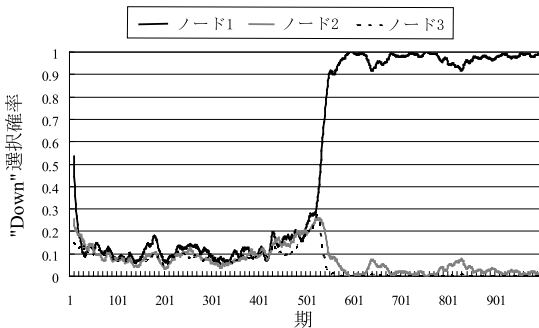


図 13: 終了ノードの割合 (ノード 1-3)

図??より、まずはじめにノード 9 まで到達するゲームがなくなり、ノード 8 に到達したすべてのエージェントが *Down* を選択している。さらに学習が進むと、ノード 8 に到達しなくなり、ノード 7 まででゲームが終了する。また、図??, ??から、他のノードにおいても同

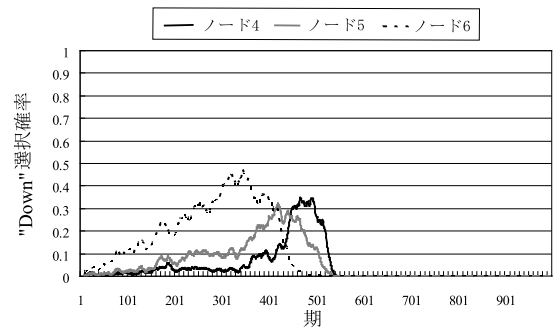


図 14: 終了ノードの割合 (ノード 4-6)

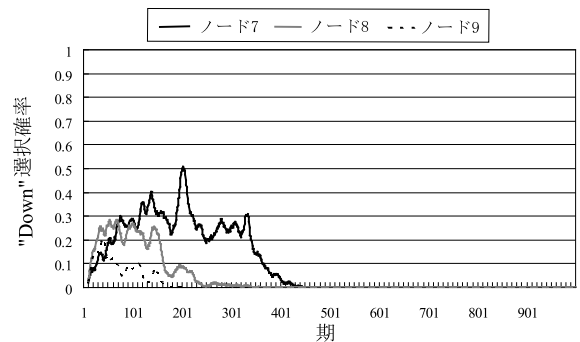


図 15: 終了ノードの割合 (ノード 7-9)

様の結果が示され、ゲームを繰り返し行うことでノード番号の高いノードから順番に到達することがなくなり、500 期付近からは、90% 以上の高い割合でサブゲーム完全均衡が実現している。以上の結果は、ムカデゲームの構造に大きく影響していると考えられる。実験の初期においては、エージェントの行動はランダム性が強いのでノード 7-9 に到達するエージェントが存在し、協調関係が成功するが、ノード 9 に到達する回数が増減してくると、プレイヤー番号 3 のプレイヤーは高い利得を得ようとノード 6 で *Down* を選択するようになる。すると、今度は、ノード 7, 8 に到達する回数も減少することになるので、次第に、プレイヤー全体がノード番号の小さいノードでゲームを終了する傾向となる。ゲームを繰り返し行い、以上の過程を経ることで、ノード 1 で *Down* を選択するサブゲーム完全均衡が実現すると考えられる。したがって、3 人ムカデゲームにおいてはゲームを繰り返すことでサブゲーム完全均衡の予測と同様に、ノード 1 で *Down* を選択してゲームを終了するようになるが、このことは、初めのランダムな状態において成立した協調が高い割合で成功している関係が、ノード番号の小さいノードで *Down* を選択するという裏切り行為によって崩れることにより、非協調な関係へと変

化するからである。しかし、協調的な関係が継続する期間がグループにより異なっていることから、この関係が長期間続いたグループが、被験者実験のグループ 3 に示されているノード 1-3 での *Down* を選択する確率が上昇しない結果であるといえる。

5.2 追加実験：プレイヤー数変動実験

プレイヤー数を 2, 3, 4, 5 と変動させ、各 100 回のシミュレーション実験を行った結果を図??と同じ形式で図??に示す。

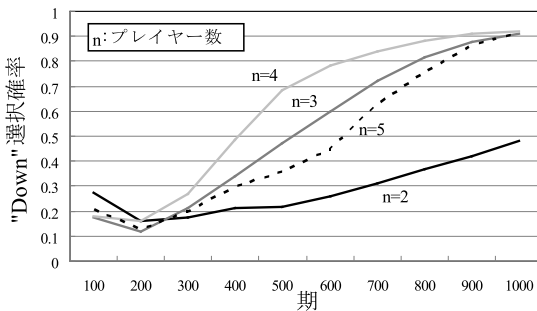


図 16: プレイヤー数変動実験

このとき、プレイヤー数が 4, 5 のムカデゲームではゲームを繰り返し行うと協調が成功しにくくなり、サブゲーム完全均衡が実現されやすくなる。しかし、プレイヤー数 2 のシミュレーション結果はゲームを繰り返し行ってもサブゲーム完全均衡が実現する割合がそれほど高くなることはなく、また、901 期-1000 期のノード 1 で *Down* を選択する割合は 42% であった。長期間ゲームを繰り返し行うことでエージェントの行動に変化が現れるのかを確認するため、3000 期のシミュレーション実験を行い、結果を図??に示す。

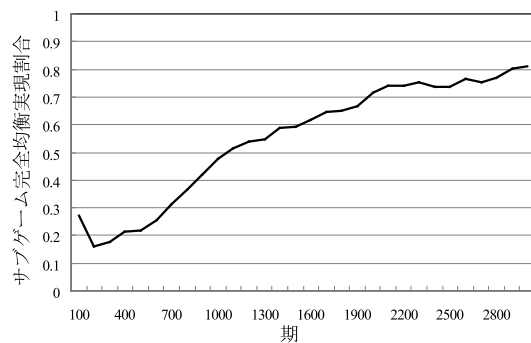


図 17: サブゲーム完全均衡の実現割合の推移 (プレイヤー数 2)

この結果、さらにゲームを長期間繰り返し行うことでサブゲーム完全均衡が実現する割合が 80% を超えた。したがって、プレイヤー数 2 のシミュレーション結果も長期間学習を行うことで、サブゲーム完全均衡の割合が上昇していき、サブゲーム完全均衡が実現されやすくなった。これらの結果から、プレイヤー数 3, 4, 5 では自分以外のプレイヤー数が 1 人から複数に増えることで、プレイヤー間の協調関係を長期間維持することが難しく、早い段階でサブゲーム完全均衡が実現すると考えられる。また、プレイヤー数を 5 としたシミュレーション実験で収束が遅いのはエージェント数が多くなり、他に比べて効率的な学習が行われなかったことが原因だと考えられる。

6. おわりに

本研究では、人工適応型エージェントを用いたシミュレーションシステムによるムカデゲームにおける被験者の行動分析を行った。エージェントのニューラルネットワークと遺伝的アルゴリズムに基づく意思決定機構により、試行錯誤的な人間の意思決定構造を模倣することで、ムカデゲームにおける被験者の意思決定および学習過程に関する分析を行った。シミュレーション分析の結果、提案モデルが 3 人ムカデゲームにおける被験者の試行錯誤的な意思決定構造を適切に表現できていることを示した。さらに、長期的な学習により、協調が必ずしも成功するとは限らず、繰り返しゲームを行うことで、サブゲーム完全均衡が実現されやすくなる可能性があることを示した。

参考文献

- [1] J. H. Holland and J. H. Miller. "Adaptive Intelligent Agents in Economic Theory", *American Economic Review* **81**, pp. 365-370, 1991.
- [2] R. Axelrod, "The Complexity of Cooperation," *Princeton University Press*, pp.14-29, 1997.
- [3] A. Rapoport, W. E. Stein, J. E. Parco and T.E. Nicholas, "Equilibrium Play and Adaptive Learning in a Three-person Centipede Game", *Games and Economic Behavior*, **43**, pp. 239-265, 2003.
- [4] E. Fehr and K. M. Pfaffrey, "A Theory of Fairness, Competition, and Cooperation", *The Quarterly Journal of Economics*, **114**, pp. 817-868, 1999.
- [5] I. Nishizaki. "A general Framework of Agent-Based Simulation for Analyzing Behavior of Play-

ers in Games,” *Journal of Telecommunications and Information Technology*, pp. 28–35, 2007,

- [6] J. A. Sundali, A. Rapoport and D. A. Seale. “Co-ordination in Market Entry Games with Symmetric Players”, *Organizational Behavior and Human Decision Processes* **64**, pp. 203–218, 1995.
- [7] K. G. Zauner, “ A Payoff Uncertainty Explanation of Results in Experimental Centipede Games”, *Games and Economic Behavior*, **26**, pp. 157–185, 1999.
- [8] R. D. McKelvey and T. H. Palfrey, “ Quantal Response Equilibria for Extensive Form Games”, *Experimental Economics*, **1**, pp. 9–41, 1998.
- [9] 林田 智弘, 西崎 一郎, 片桐 英樹, “社会的評判を考慮したネットワーク形成に関するエージェントベースシミュレーション分析,” 日本経営システム学会誌, **25**, pp.21–32 , 2009

平成 22 年 10 月 29 日