

博士論文

空間構造上の協力行動に関する
エージェントベースモデル

48-47404 小野真裕

指導教員 石塚満 教授

東京大学大学院 情報理工学系研究科 電子情報学専攻

概要

科学の歴史において、デカルトの「方法序説」に始まる要素還元主義の影響は非常に大きい。要素還元主義とは、全体を細かく分割しその構成要素を観察しつくせば全体も理解することができる、という方法論である。この方法論に従った物理学や化学などはここ数世紀において大きく発展してきた。

しかし、全体が全ての要素を単純に足し合わせたものであるという前提は必ずしも正しくない。その例としてよく挙げられるのは、生命や知能である。一つ一つの原子、一つ一つの細胞を観察しても、生命や知能の源となるものは発見できないが、全体としては一個の生命や知能ができあがっている。

それに対し、複雑系科学と総称される研究領域が近年登場した。機械論的、決定論的でありながら、要素還元主義では説明が難しい現象を解明しようという試みである。この研究領域で議論の中心の一つとなっているものは、創発と呼ばれる現象である。創発とは、多数の要素が含まれる系において、それら要素間の相互作用が、要素それぞれからでは説明できない全体としての振る舞いを生み出す現象を指す。比喩的に言えば、要素間が線形的な関係であれば要素還元主義に沿い予測が可能であるのに対し、非線形的な関係では要素間の相互作用の考慮が不可欠となり創発現象は予測困難であることが特徴である。この考え方は計算機シミュレーションを強力な武器とし、計算機の発展と共に近年盛んになってきている。

世の中の一般的なシステムに共通する性質を知るとは、理学的な意味においても工学的な意味においても非常に重要である。本研究の根底には、創発現象の根本の法則を見出したい、という目的がある。しかしながら、現在のところ一足飛びにこの目的を達成することは困難であるため、一つの具体的な領域に限定し、その限定された現象から知見を得、最終的には一般システムへの適用を目指したい。

本研究では、空間構造上における意思決定主体間の協力行動、という事例を取り上げ、集団としての協力行動という創発現象に対する知見を得ることを目的とする。本研究ではこの目的に対して、構成論的アプローチをとる。構成論的アプローチとは、多数の要素からなる一つの系をつくり、ミクロなルールに従う多数の要素を実際に動作させ、それらを観察することによってマクロな創発現象に対する理解を深める、というものである。

具体的には、系の要素としては意思決定を行うエージェントを想定し、進化的なマルチエージェント・シミュレーションを行う。エージェントは、意思決定を数理的に扱うゲーム理論におけるゲームの一つである、囚人のジレンマ・ゲームを他のエージェントとプレイする。本研究の特徴として挙げられるのは、エージェント同士の接触を限定する空間構造という概念を導入している点である。ゲームをプレイする対戦相手のエージェントは、空間構造により決定される。空間構造は、近年研究の盛んなネットワーク・モデルを利用しており、そのノードとしてエージェントは存在する。

エージェントの意思決定は相互に影響を及ぼし合い、集団としての特性もそれらエージェントの意思決定が積み重なって生じる。そこで、エージェントの振る舞い、エージェント集団の性質といったものを観察し、創発現象に関する知見を得ることを試みる。また、本研究では新たなエージェントモデルを提案することにより、空間構造がエージェントの学習に及ぼす影響や、相手の意図の推定能力の有用性を考察することが可能となる。

以上のような背景・目的を踏まえ、本論文は以下の構成をとっている。

第一章では、本研究の背景および目的を述べる。これまで伝統的であった要素還元的な手法ではなく構成論的な手法の必要性を論じ、本研究では構成論的アプローチに従い、論を進めることを宣言している。

第二章では、ゲームとは何か、また関連する領域の中でどのような位置づけとなるかを説明している。本研究に関連する内容に関しては、これまで異なる分野においてそれぞれ研究がなされてきている。例えば、社会的にどのような場合にどのようなゲームが発生するか、慣習や道徳が歴史的にどのような経緯を経ているか、それらを整理することによって本研究の目的がより明確になる。関連分野の説明とそれらを統合したパースペクティブの提案、およびその中で本研究の位置づけを述べる。

第三章では、空間構造上において囚人のジレンマ・ゲームをプレイするエージェント集団の基礎的な振る舞いを調べる。プレイヤーは、過去一回のゲームの記憶を持ち、刺激 反射モデルに従う一次のメタ戦略をとる。計算機シミュレーションを行い集団を進化させて、エージェントの振る舞いや集団の特性を観察する。これまで生態学の分野で確認されていた、各生物種のランクと個体数の関係はべき乗に従うという現象を確認した。解析的、数値的に既に示されている単純なマトリクスによる進化実験だけではなく、ゲームを行うエージェントにおいても同様の現象が発生することがわかった。

第四章では、ゲームと連動して空間構造が変化する場合について調べる。エージェント間の接続をエージェントが操作可能であるとき、集団の規模を一定とした場合と集団が成長する場合について、空間構造が変化する様子をシミュレーション実験によって確かめた。どちらの場合においても、実世界において観察されるようなネットワークのように特定のエージェントへの集中が見られる。

第五章では、ゲームをプレイする新たなエージェント・モデルを提案し、空間構造が学習にあたえる影響を調べた。エージェントは強化学習を行い、ゲームの過程で自分の戦略を変化させることができる。そのため進化の過程において学習パラメータを観察することで、空間構造がエージェントの学習に与える影響を調査する。実験により、エージェント間の関係性が強くなれば将来の報酬を期待し協力しやすくなることがわかった。これは社会学における信頼と集団の流動性の関係に関する知見を支持するような結果である。また、学習は協力の弊害となることが明らかとなった。

第六章では、集団の制御可能性について検討する。工学的、社会工学的な応用を考える場合には、集団の特性を何らかの方法で望ましい状態に導くような技術が必要である。しかしながら、そもそも結果の予測が困難な複雑な系においては、そのような要件を満たす決定的な手法

は存在しない．ここではゲームととらえて定性的な議論を行い，考えうる方法を整理する．次に，空間構造を利用した制御方法を提案し，実験によりその効果を調べる．その結果，集団の多様性とのトレードオフがあるものの，ある程度空間構造に偏りがある場合には効果があることが確認された．

第七章では，他エージェントの意図を推測しゲームを行うエージェントを提案し，その効果を調べる．人類は他者の心を推測し，それも考慮に入れた上で自らの行動を決定する．その機構は社会的なゲームにおいては非常に重要であると考えられる．そこでその効果の重要性を定量的に評価した．対戦相手のエージェントの戦略を推定し，後方帰納法に基づき最も利得の得られる行動を選択するモデルを提案し，実験によりその効果を確かめた．その結果，反射的なモデル，学習するモデルと比較して著しく協力する可能性が高いことがわかった．

最後に第八章では，本論文をまとめ今後の課題について言及する．本研究では先述の目的に従い，空間構造上における協力行動を観察することにより，創発的な現象に関する知見を得ようと試みた．多主体意思決定環境においては，ある意思決定主体は他の意思決定主体から影響を受けるが，特に空間構造によってその影響の具合が異なること，集団の相互協力の度合いが異なることがわかった．また，複数の主体の意思決定モデルを提案し，それぞれについてその機能の効果を確認した．その結果，学習は協力の弊害になりうること，他のエージェントの戦略の推測は非常に有益であること，がわかった．

目次

第 1 章	序論	1
1.1	背景	1
1.2	本研究の目的	2
1.3	本研究のアプローチ	2
1.3.1	方法論	2
1.3.2	構成論的アプローチ	3
1.3.3	本研究の構成要素	4
1.4	本論文の構成	5
第 2 章	個体・集団の意思決定	7
2.1	はじめに	7
2.2	個体の振る舞い	8
2.2.1	脳科学と人工知能	8
2.2.2	経済学からのアプローチ	9
2.2.3	個体の振る舞いのまとめ	10
2.3	集団の振る舞い	10
2.3.1	数理生物学	10
2.3.2	自己駆動粒子	11
2.3.3	思想, 社会学	11
2.3.4	集団の振る舞いのまとめ	12
2.4	ゲーム	12
2.4.1	ゲームの種類	13
2.4.2	実社会におけるゲーム	18
2.4.3	情報	18
2.4.4	合理性	19
2.5	意思決定モデル	20
2.5.1	評価軸	20
2.5.2	不確実性	21
2.5.3	意思決定の基準	22

2.6	集団の意思決定	23
2.7	まとめ	23
第3章	空間構造上における囚人のジレンマゲーム	25
3.1	はじめに	25
3.2	空間構造	26
3.2.1	スモールワールドネットワーク	26
3.2.2	スケールフリーネットワーク	28
3.2.3	ネットワーク特徴量	28
3.3	エージェントモデル	29
3.3.1	遺伝子モデル	29
3.3.2	戦略の表現可能性	30
3.4	ネットワーク構造の影響	31
3.4.1	シミュレーション内容	31
3.4.2	シミュレーション結果	31
3.5	まとめ	35
第4章	ネットワーク形成	36
4.1	はじめに	36
4.2	自由生成条件におけるネットワーク形成	37
4.3	固定生成条件におけるネットワーク形成	39
4.3.1	ネットワーク形成モデル	39
4.3.2	シミュレーション実験	41
4.4	まとめ	44
第5章	空間構造の学習への影響	45
5.1	はじめに	45
5.2	学習エージェントモデル	46
5.2.1	強化学習	46
5.2.2	学習エージェント	46
5.3	シミュレーション実験	48
5.3.1	シミュレーション内容	48
5.3.2	シミュレーション結果	50
5.4	まとめ	58
第6章	集団の制御可能性	60
6.1	はじめに	60
6.2	提案制御方法	61

6.3	シミュレーション実験	62
6.4	空間構造と制御可能性	67
6.5	まとめ	68
第7章	他者意図の推定	69
7.1	はじめに	69
7.2	他者意図を推定するエージェントモデル	70
7.3	シミュレーション実験	73
7.3.1	異なるモデルとの対戦	75
7.3.2	推定精度	76
7.3.3	進化的安定性	77
7.4	まとめ	78
第8章	結論	79
8.1	本研究の結論	79
8.2	今後の課題	80
	謝辞	82
	参考文献	83
	関連発表文献	88

表目次

2.1	脳理解レベル	8
2.2	囚人のジレンマの利得行列	15
2.3	財の分類	18
3.1	メタ戦略例	30
3.2	対戦例	31
3.3	シミュレーション諸元	32
4.1	シミュレーション緒元	41
5.1	メタ戦略例(再掲)	48
5.2	シミュレーション緒元	49
6.1	シミュレーション緒元	62
6.2	空間構造と集団の性質	68
7.1	推定テーブル	72
7.2	シミュレーション緒元	73

目次

1.1	マイクロ マクロ・リンク・ダイアグラム	4
1.2	本研究において取り上げる構成要素	5
2.1	ゲーム理論の分類	13
3.1	スモールワールドネットワーク	27
3.2	エージェントモデル	29
3.3	各世代における手順	32
3.4	平均利得推移 ($R=0/1$)	33
3.5	世代交代例 ($R=0.2$)	33
3.6	時空間占有率と順位の関係	33
3.7	クラスタ係数に対する支配種の時空間占有率, 平均利得の変化	34
4.1	各世代における手順	37
4.2	時空間占有率と順位の関係	38
4.3	リンク数とエージェント頻度 (エージェント数=3000)	38
4.4	各世代における手順	42
4.5	平均利得の推移	43
4.6	進化ダイナミクス例 ($\text{noise}=0, \text{mutation}=0$)	43
4.7	定常状態におけるネットワーク	43
5.1	エージェントモデル	47
5.2	平均利得	51
5.3	Action Change Rate 定義	52
5.4	強化学習パラメータ	53
5.5	強化学習パラメータの推移例	54
5.6	エントロピー	55
5.7	平均利得	56
5.8	強化学習パラメータ	57
5.9	エントロピー	58

6.1	提案制御方法	61
6.2	平均利得	63
6.3	強化学習パラメータ	65
6.4	エントロピー	66
7.1	他者の意図を推定するエージェントモデル	71
7.2	推定の更新	72
7.3	バックワードインダクションによる手番決定	73
7.4	対戦例	74
7.5	協力率	75
7.6	協力に至るケースの割合	76
7.7	想定シグナル誤り率による差異	77
7.8	進化の過程例	78

第 1 章

序論

1.1 背景

我々の世界はどのように成り立ち動いているのか。これまで人類はこの問いに答えようと努力を続けてきた。

デカルトの「方法序説」に始まる要素還元主義は、その問いに答えるためのこれまでの科学の成功に大きく寄与している。これは全体を細かく分割しその構成要素を観察しつくせば全体も理解することができるという方法論である。この流儀に従った物理学や化学などはここ数世紀において大きく発展してきた。

しかし、全体が全ての要素を足し合わせたものであるという前提は必ずしも正しくない。その例としてよく挙げられるのは、生命や知能である。一つ一つの原子、一つ一つの細胞を観察しても生命の源となるものは発見できないが、全体としては一個の生命、知能ができあがっている。他にもアリの集団行動や人間組織なども、個体それぞれからでは説明できない集団の行動が発生することがある。

以上のような要素還元主義におさまらない事実を説明する試みは 20 世紀後半には始まる。その嚆矢として挙げられるのは、1973 年のベルタランフィの「一般システム理論」[Bertalanffy 73] である。これは、生物、社会など様々なものを一種のシステムとしてとらえ、システム共通の性質を探る試みである。ベルタランフィに端を発したシステム論の流れは、20 世紀後半に盛んとなった。第二世代と呼ばれる自己組織化の理論、第三世代のオートポイエーシス [Maturana 80] とその後進んだ。また、生物の説明に試みたオートポイエーシスの概念を社会学に取り入れたルーマンの社会システム理論も出現し、具体的な研究領域への応用も試みられてきた。

その後、複雑系科学と総称される研究領域が登場した。機械論的、決定論的でありながら要素還元主義では説明しがたい現象を説明するものである。この研究領域では、多数の要素が含まれる系において、それら要素間の相互作用が要素それぞれからでは説明できない全体としての振る舞いを生み出す、創発という現象が議論の中心の一つとなっている。創発現象の有無は比喩的に言えば、要素間が線形的な関係であれば要素還元主義に沿うが、非線形的な関係では

要素間の相互作用の考慮は不可欠、との差異となる。この考え方は計算機シミュレーションを強力な武器とし、計算機の発展と共に近年盛んになってきている。

1.2 本研究の目的

世の中の一般的なシステムに共通する性質を知ることは、理学的な意味においても工学的な意味においても非常に重要である。本研究の根底には、創発現象の根本の法則を見出したい、という目的がある。しかしながら、現在のところ一足飛びにこの目的を達成することは困難であるため、一つの具体的な領域に限定し、その限定された現象から知見を得、最終的には一般システムへの適用を目指したい。

本研究では、空間構造上における意思決定主体間の協力行動、という事例を取り上げ、そこから創発現象に対する知見を得ることを目的とする。

1.3 本研究のアプローチ

1.3.1 方法論

マクロな創発現象を説明する際には、どのような態度をとるか、を前もって考えておく必要がある。特に、社会科学は自由度が大きすぎるため、ある「主義」を持つことがそれに対応する。ここでいう主義とは研究の方法である。本論文は必ずしも社会科学の範疇に含まれるものではないが、ここでそれを整理しておくことは有益であると考えられる。

社会科学において1950~60年代を通じて、方法論的個人主義と方法論的集合主義の論争が行われた。方法論は社会科学をどのような研究方法をとるかを意味し、どのような研究方法をとるかという態度に対して定式化が行われた。金子によれば、方法論的個人主義は次の三つに分類できる [金子 03]。

- 還元論的個人主義：すべての社会現象の原因は個人の特性に帰着する
- 存在論的個人主義：社会の諸々の行為をとり行う基本単位は個々の人間である
- 同一性確定的個人主義：構造がはじめから不変である主体から理論がなりたっている

存在論的個人主義は、個人の特性に関する仮定はなく、基本単位の存在論的意味だけしか主張しておらず強い主張ではない。個人の特性まで規定した還元論的個人主義は、問題を人間個人に還元するため最も強い。社会現象の分析は「人間個人」の特性を研究すればよいという方法論的主張である。同一性確定的個人主義は、それら二つから独立である。例えば、企業などの集合体を同一性をもつ個人と扱うために考えられた。そのため、同一性をもつ個人が人間個人である場合には、還元論的個人主義と等しくなる。

個人主義に対応し集合主義も存在し、こちらは個人ではなく集合が説明となる。

- 存在論的集合主義：社会における基本的な意思決定単位は全体である

- 個 - 全体形成的集合主義：各個人が社会の中で形成され、それと同時に全体も構成される

存在論的集合主義とは「集合体の実在性に優先権を与え、集合体が呈する性質は、個人に帰属する性質から演繹できないとする立場」であり、伝統的な方法論的集合主義である。

個 - 全体形成的集合主義は、研究の方法論でなく研究対象の特性についての立場である。「存在論的個人主義」的な数学モデルによって、「個 - 全体形成的集合主義」的現象を捉える。

以上の個人主義と集団主義は双方とも個人と集団の両極端に振れすぎていると考える。社会は必ずしもどちらか一方の要因によってのみ説明することは不可能であり、両者の影響が含まれるに違いないからである。そこで、別の方法論としてジンメルの方法論的關係主義をとりあげる。彼の主張では、個人間の相互作用が重要であり、それが集合的な社会現象を説明するものである。

- 方法論的關係主義：社会における意思決定は個人の関係性が影響を与える

個人の関係性自体は社会とも言い換えることが可能であり、この主張がもっとも社会を考えるのに適当であると考えられる。また、この主張は、自己の意思決定が他者の意思決定に依存するゲーム的な状況をうまく説明できるものである。他者との関係をゲーム理論的にとらえた場合、他者のとる戦略に全く依存しない支配戦略が存在しないことは往々にしてありうる。支配戦略が存在する場合は完全に個人主義で考えることが可能であるが、そうでない場合は関係主義をとる方が良い、ということである。

以上から、本研究では、創発現象は個人・集団のどちらか一方で説明できるものではなく、個人の関係性が重要であるとの立場をとる。

1.3.2 構成論的アプローチ

本研究は構成論的アプローチをとる。構成論的アプローチとは、ミクロなルールに従う多数の要素を実際に動作させ、それらを観察することによってマクロな創発現象に対する理解を深めるというものである。このアプローチは人工生命 [有田 02]、人工社会 [Epstein 96][和泉 03] といった研究分野においてよく用いられる。また、このアプローチを人工知能にも適用しようという提案もある [中島 06]。

先述したように、本研究は、空間構造上における意思決定主体間の協力行動、に関する。エージェントと呼ばれる意思決定主体が複数存在する環境において、それらエージェント間の相互作用に基づき各エージェントの意思決定の振る舞いは変化する。それらエージェントの振る舞い、エージェント集団の性質といったものを観察し、創発現象に関する知見を得ることを試みる。特に本研究では、エージェント同士の接触を限定する空間構造という概念を導入している。エージェントが空間構造から受ける影響、およびエージェントが及ぼす空間構造への影響もまた研究対象である。

このようなミクロな現象からマクロな現象の説明に関しては、社会学において行われてき

た．ここでは，それを用いて本研究のアプローチを整理しておきたい．

本研究では，マクロな創発現象を説明しようとするとき，ミクロな行為の集積としてそれを説明する立場をとる．基本的に同一の立場をとると考える，社会学者のコールマンの合理的選択論をここでは説明する．まず，コールマンによるマクロ水準命題とミクロ水準命題の相関図 [Coleman 98] を図 1.1 に示す．

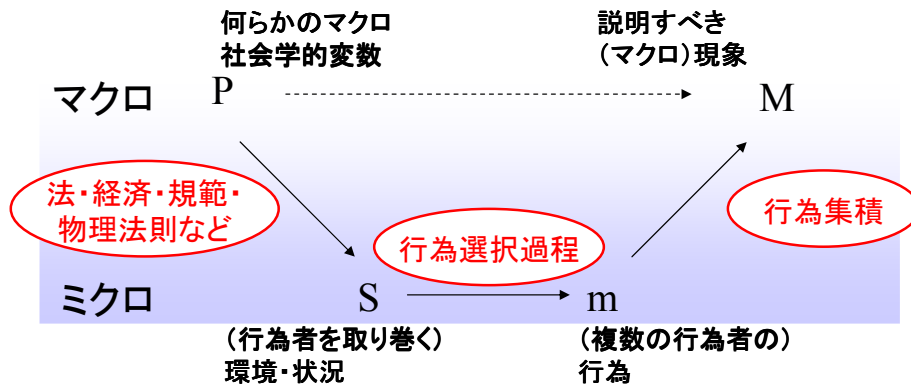


図 1.1. マイクロ マクロ・リンク・ダイヤグラム

マクロの現象を説明する際には，必ずしもミクロからマクロという流れである必要もなく，マクロからマクロという説明もあり得る．これは伝統的なデュルケム流の社会学での立場である．コールマンらの合理的選択理論は，ミクロな要素間になんらかの関係性を仮定し，それらの関係性を考慮しつつ要素が合理的な選択を行う，その行為の集積からマクロな現象が創発するとしている．近年の社会学理論ではこのようにミクロな行為の集積としてマクロな現象を理論的に説明しようとして試みられている．最近のゲーム理論の発展を受けて，既存の社会学における定性的な議論はゲーム理論を利用することによって数理モデル化されてきた．それにより，合理的選択理論を支持する研究者にとっては，ゲーム理論によってミクロな行為の集積としてマクロな現象を説明することができるようになってきた．

1.3.3 本研究の構成要素

図 1.2 に本研究のシステム構成要素を示す．これまで述べたように，個体間関係性を重視し創発現象を説明する立場をとる．さらにその方法としては，実際にそのような環境を計算機上に作り上げ，その様子を観察することによって行うこととしている．この図は構成するシステムと，それぞれの特徴をあらわしている．

システムは先述したように一般的なシステムを想定している．そのため，システムを構成する要素は何らかの下位システムから構成されているかもしれないし，システムは上位システムを構成するかもしれない．

本研究では，システムを構成する要素間の関係は囚人のジレンマゲームとする．囚人のジレンマゲームをプレイする要素をエージェントと呼び，エージェントがゲームにおける行動の選

択方法についてはいくつかのモデルを提案する．要素間を結ぶシステムの構造は近年盛んなネットワーク理論を利用する．

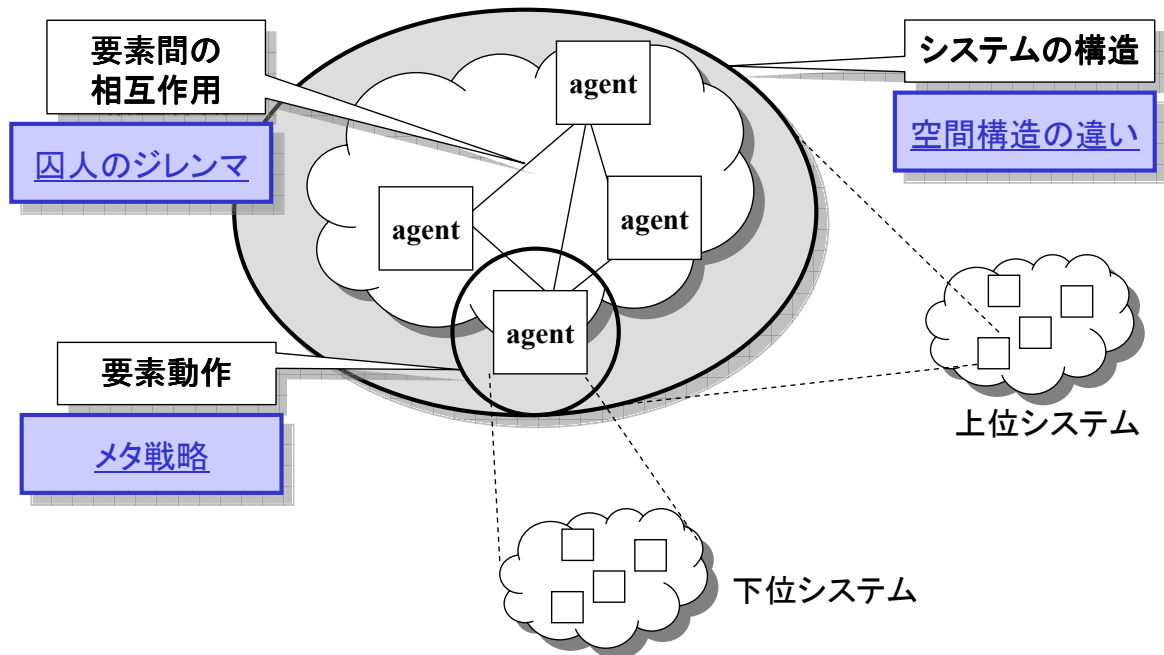


図 1.2. 本研究において取り上げる構成要素

1.4 本論文の構成

本章の最後に、本論文の全体構成を示す．本章では、本研究の背景、目的、および本研究のアプローチについて述べた．

第二章では、意思決定主体が意思を決定する枠組みとなるゲームとは何か、またゲームをプレイするという現象は関連する研究領域の中でどのような位置づけとなるかを独立した章として記す．

第三章では、空間構造上において囚人のジレンマ・ゲームをプレイするエージェント集団の基礎的な振る舞いを考察する．これまで生態学の分野で確認されていた現象が、本シミュレーション実験においても観察された．

第四章では、ゲームと同時に空間構造の変化が発生する場合について調べる．エージェント間の接続をエージェントが操作可能である場合に、どのような空間構造になるかをシミュレーション実験によって確かめた．

第五章では、ゲームをプレイする新たなエージェント・モデルを提案する．エージェントは強化学習を行い、学習パラメータを観察することで、空間構造がエージェントの学習に与える影響を調査する．

第六章では、集団の制御可能性について検討する．空間構造を利用した制御方法を提案し、実験によりその効果を調べる．

第七章では、さらに他エージェントの意図を推測しゲームを行うエージェントを提案する．これまでのエージェントに対する意図を推測するエージェントの強さを調べ、意図を推測する機能の効果を調べる．

最後に、本論文をまとめ今後の課題について言及する．

第2章

個体・集団の意思決定

2.1 はじめに

本章では、本研究に関連する従来の研究を俯瞰する。

本研究は、先に述べたように、多数の要素からなる一つのシステムを想定している。要素は何らかの行動を行い、自らの行動と他の要素の行動との組み合わせから決定される利得を得る。この関係はゲーム理論で記述されるものである。これらの要素の振る舞い、およびシステム全体の様子を観察し知見を得ることを目的としている。

このように本研究自体は何らかの現象に特化したものではなく、様々な現象の基礎となる一般的に共通する部分に焦点をあてている。見渡してみると、このシステム構造自体は、生物学、社会思想、法学、経済学、工学など多岐にわたって適用可能な領域が存在する。異なる現象について、類似した議論が定性的に様々な分野で行われていたが、ゲーム理論によってそれらが接続され、数理的に考察されるようになってきている。また、最近では、ゲームをプレイするプレイヤーに関して、合理的ではない場合についての研究も盛んになりつつある。それは、理論的なアプローチもあれば、構成論的なアプローチもある。このように、全てを統合できる一般システムを考えられるような研究の要素は揃いつつある状況である。本章では、それらの領域に関し概観し、本研究の位置づけを理解しやすくすることを目的とする。それらを統合した新しい見方を提案し、その中で本研究の位置づけを確認する。

本章は次のように構成される。まず、個体の意思決定、集団の意思決定に関しての既存の研究分野をみる。次にそれらに共通して関連のあるゲーム理論について説明する。ゲームに関してはゲームを構成する様々な概念があるため、後の章の理解を助けるためにそれらを整理しておく。次に、具体的な個体の意思決定モデルを考える際に必要となるいくつかの概念を説明する。最後に、それらをまとめ、本研究の位置づけを確認する。

2.2 個体の振る舞い

集団の振る舞いはマクロに統計的に見ることも可能であるが、先に述べたように本研究ではミクロな行為の集積としてのマクロな現象という立場をとる。そこで、まずはミクロな行為はどのようにして意思決定されるのかという点について整理する。古くから人間の振る舞いに関する研究は盛んであった。ここでは、脳科学と人工知能、行動経済学をピックアップしその概要を述べる。

2.2.1 脳科学と人工知能

人間の意思決定において最も重要な人の器官は明らかに脳である。その脳に関する研究、いくつかの種類がある。例えば以下のように分類することが可能である [松本 96]。

1. 脳が獲得したアルゴリズムの研究
2. 脳がアルゴリズムを自動的に獲得するための「戦略」の研究
 - (a) 脳の発生・分化
 - (b) 成熟した大人の脳での学習や学習制御の分子・細胞レベルからの脳研究

医学系の研究者は全領域に渡ってその活動を続けているが、近年の計算機の発達により計算機科学の研究者の参入も増えてきた。計算機上でのドライな研究では、主に以上のうちアルゴリズム自体、つまり獲得したアルゴリズムとそのアルゴリズム獲得戦略に関わり、基本的にはウェットな有機的な部分は省かれている。また、人間には判断系は、主に前頭前野における知的な判断系と、大脳辺縁系における情動的な判断系が存在する。これまでの計算機上での研究で、主に現在扱われているのは前者であるともいえる。

デビッド・マーによれば脳を理解するために必要な三つのレベルがあるという(表 2.1(「ビジョン」[Marr 82] より))。ここではアルゴリズムとともに、ハードウェアによる実現が含まれており、上の分類と脳科学の研究者のビジョンが共有されている。

表 2.1. 脳の理解レベル

計算理論 (computational theory)	計算の目標は何か、なぜそれが適切なのか、そしてその実行可能な方略の論理は何か
表現とアルゴリズム (representation and algorithm)	この計算理論はどのようにして実現することができるのか。特に入力と出力の表現は何か、そして変換のためのアルゴリズムは何か
ハードウェアによる実現 (hardware implementation)	表現とアルゴリズムがどのようにして物理的に実現されるか。

近年では、脳科学の発展が著しく、計算機上での学習モデルと実際の脳の働きが一致する例などもでてきた。そこで、計算機を利用した脳研究を更に進める流れが出てきている。例えば、川人らは、「脳の理解を、その機能を脳と同じ方法で実現できる計算機のプログラムあるいは人工的な機械を作れる程度に、深く本質的に理解することを目指すアプローチ」[川人 96]である計算論的神経科学を提唱している。

また、銅谷による役割分担仮説では、脳の3つの部位を「なにをするのか」によって分類するのではなく「どのように学習されるか」によって分類する。この仮説によれば、大脳皮質では「教師なし学習」が行われ外界の統計的性質を反映した表現を作り、小脳では誤差信号によって学習が進む「教師付き学習」が行われ、大脳基底核では、報酬による「強化学習」が行われる [Doya 00]。これは計算機上での学習モデルと脳の部位を対応させたものであり、計算機科学の分野からの脳科学への貢献も盛んである。

2.2.2 経済学からのアプローチ

経済学では合理的経済人が仮定される。認知や判断処理に関し完全であり、意思決定における基準は物質的利益の追求である。しかし、実際の人間は必ずしも合理的経済人のように振舞うことはありえない。経済学への批判として、この点は古くから指摘されてきた。

そうした中で、合理的経済学派における合理的経済人の仮定を排し、経済学の分野も生まれしてきた。その一つが行動経済学である。行動経済学とは、人は実際にどのように行動するのか、なぜそうするのか、その行動の結果としてなにが生じるのかといったテーマに取り組む経済学である [友野 06]。

行動経済学は、認知心理学からの影響が大きい。また、進化論的なものの見方も導入されるなど、種々の分野と関連性がある。これは人間の行動という基礎的な研究対象であることが大きい。一つの基準としてこれまでの行動経済学では、既存の合理的経済人に反する事例に着目し、それらの証拠を系統的に収集してきた。有名な研究はカーネマンとトヴァルスキーによるプロスペクト理論である [Kahneman 79]。彼らはヒューリスティクスとバイアスの研究により、合理性という非現実的な観念を否定した。

特に最近では神経経済学という分野もできつつある。これは神経学的見地から経済学的行動を研究するものである。脳をモニターしながら、何らかのゲームをさせるなど、これ以前では表面的な人の行動を見る事しか出来なかったところを、脳の状態を見るところまで進めたものである。例えば、ゲームにおける協力行動はドーパミンの放出と快行動とに関わる領域を活性化することがわかっており、美德そのものが報酬となっていることを示唆している結果も示されている [Rilling 02]。

また、人間行動に関しては、金融論の研究者のマイケル・ジャンセンは5つのモデルとして、社会学、心理学、政治学、経済学、そして新たなモデルを挙げた。最初の三つである社会学、心理学、政治学に基づくモデルは切って捨てる。経済学を土台にした第四のモデルは否定せず、第五のモデルに統合している。「知恵があり、評価を下し、最大化を目指す存在として

のモデル (REMM: Resourceful, Evaluative, Maximizing Model) [Jensen 94] を提案している。このモデルでは人間は、様々な欲求を持っており、ある欲求を他の欲求と比較したり別の欲求を満たしたり、といった行動をとる、行動の根源は欲求であり、それを満足することはなく最大化を目指すというものである。

2.2.3 個体の振る舞いのまとめ

以上、実際に意思決定を行っている器官を対象とした研究と人間の意思決定の行動原理を説き明かそうとする研究を説明した。前者は、さらに意思決定アルゴリズムやアルゴリズム獲得、ハードウェアへの実装までを射程とする。後者は、ゲーム的な状況を用意し、合理的な意思決定から外れた部分にフォーカスしている。また、ゲーム的な状況は意思決定を考える際に、条件が整理されたものとして有益であることから、前者の脳科学と融合した試みも始まっている。

2.3 集団の振る舞い

集団としての振る舞いを調べる場合、これまでどのような研究がなされてきたかをまとめる。ここでは、多数の要素からなるシステムについての既存研究分野を概観し、その本質的な部分を抽出しまとめる。

ただし、必ずしも上記の個体の振る舞いのモデルとは直接接続されてはいない。

2.3.1 数理生物学

生命現象や生物の生態に関しては、人間自体も生物であること、生物の存在が非常に身近であることから古くから研究が行われており、またその射程とする範囲も広い。古くは自然史学といった様相であったが、19世紀以降はダーウィンによる進化論の提唱や、メンデルの遺伝法則の発見があり実験科学の面が増し、DNAの発見から現代では実験が主となってきている。

また近年、集団遺伝学や個体群生態学の分野で数学モデルを用いた解析の有効性が広く認識されるようになり、数理生態学・数理生物学といった分野が形成されてきた。これら研究分野では、生物の種としての振る舞いを数理モデルとして記述することにより、解析的に現象を説明することができるようになった。生物学における数理モデルの利点は、例えば生物群の増殖に関するロジスティック・モデル、二種以上の生物が互いに影響を及ぼし合うロトカ・ヴォルテラモデルなどの利用により、いくつかのパラメータによって個体数の増殖パターンが簡潔に説明できるようになる点にある。

2.3.2 自己駆動粒子

ニュートンの力学は、慣性の法則、作用・反作用の法則、運動の法則という三つの基本原理に基づいている。これらの法則を満たす要素が自然界には多いため、ニュートン力学で説明・予測できる現象が多かった。また気体などの多くの要素を考える場合には、個々の法則を追うのではなく要素の集団全体としての統計的な性質に関する統計力学が用いられた。

ニュートン力学や統計力学では、基本的には要素として上に挙げた三つの法則を満たすことを前提としている。しかし、近年では何らかのシステムを考えるとき、システムの規模が大きくなってきたことや、システム概念の及ぶ範囲が広がってきたことから、要素としては人や車といったものを想定することも増えてきた。これら上記法則を満たさないものを想定するとき、ニュートン力学以外の理論が必要となる。

主に物理系の研究者たちによる取り組みによって、このような要素のことを自己駆動粒子 (self-driven particle) と呼び近年研究が始められつつある。ニュートンの三つの基本原理を満たさない要素からなるシステムにおいて、そのシステムがどのような振る舞いをするか等を調べる研究である。例えば、パケット・ネットワークの輻輳、交通システムにおける渋滞、森林火災の様子などである。また、この領域は数理的に現象を扱うため、類似した領域からのモデルの流用も多い。数理生物学者によって生体内のタンパク質合成のモデルとして考えられた非対称単純排除過程 (ASEP: Asymmetric Simple Exclusion Procedure) や数学者ノイマンによって格子状のセルと単純な規則から構成されるセルオートマトンが渋滞現象を解明する用途に利用される。

2.3.3 思想，社会学

ホブズは、ルールがなければ社会はどうなるのか、という考察を行った [Hobbes 51]。法律、制度、規範といったいわゆるルールがなく人間の意思決定に拘束が存在しない状況を仮定し、そのとき「万人の万人に対する闘争」となる自然状態となると導き出した。ホブズ問題とは「自然状態における戦争状態を回避し、社会秩序をどのように形成するか？」という問いである。ホブズ問題は秩序問題とも呼ばれる。

ホブズは強大な強制力を持つなんらかのシステムの存在を要求した。この対する考え方が発展し、ロックやルソーを代表とするような社会契約説を生んだ。この説では人々の行動を制限する統治システムを仮定する。人々はその統治システムの権利と責任に関する契約を仮想的に結ぶことにより統治システムが正当化される。

もう一つのアプローチは、自発的に秩序形成が発生する創発的な現象を仮定するアプローチも存在する。このような考え方は、特に資本主義経済に関しよく用いられ、ミクロ経済学のパレート最適市場などが具体例として挙げられる。ただし、秩序形成の対象となるものはある程度限定されており、自然状態からの秩序形成に関していえばもう少し強力な仮定が必要ではないかと思われる。

さて、現在の社会を考えてみれば、自然状態のような無秩序が支配しているわけではない。社会学は、社会のほぼあらゆる現象を射程とするが、古くからホップズ問題については取り組まれ、この問題が今なお社会学の根底に横たわっている。また、ホップズ、ロック、ルソーといった人物が、一般には社会思想の文脈で取り上げられることも多く、思想・哲学の分野においても大きく関連する問いである。特に近年では、正義論で知られるロールズも秩序ある社会について語っている [Rawls 71]。ロールズによれば、秩序ある社会とは、社会の構成員全員が同一の行動ルール、すなわち正義の諸原理を拘束力のあるものとして承認し、また、他の人々も承認することを知っている社会である。さらに、基本的な諸社会制度がその正義の諸原理を満たし、また満たしていることが一般的に知られている社会である。思想的なこの分野では、なんらかの仮定をおき、社会の発展段階を説明したり、何らかの結論を導く、といった取り組みがあるように思われる。

また近年では、この問題に対し数理社会学の分野においてゲーム理論を用いて理論的な考察も進んでいる。囚人のジレンマゲームを用いての道徳原理の説明の試みや、集団のジレンマの理論的な解決の説明などである。

近年の数理社会学に先駆けるものとして、第一章でも触れたコールマンの合理的選択論がある [Coleman 98]。マクロな現象をミクロな行為の集積として考えるものであり、個人は合理的経済人を仮定し、ゲーム的な状況において合理的な選択を行う。社会における現象はそれらの行為を観察することによって説明できるとの立場である。

2.3.4 集団の振る舞いのまとめ

集団を形成する個体としては、様々なものがある。数理生物学においては原始的な生命から高度な知性をもつ人間まで、自己駆動粒子ではパケットのような情報エンティティから人間、車などかなりのバラエティを持つ。それぞれ集団としての振る舞いを記述・説明するべく数理モデルが作られている。また社会という非常に大きな研究領域において、個人間の関係性と個人の合理的選択をベースとして社会の現象を説明する試みがある。

2.4 ゲーム

ここまで見てきた様に、多数の要素からなるシステムにおける現象は、要素の行為の集積として考えることができる。そして、その要素の行為に影響を与えるものは、要素間の利得構造である。ここでは、その利得構造とそれに依存する要素の行為に注目し整理することとする。

複数の意思決定者が存在するとき、それぞれの意思決定はどのようになるか、といった問題を考えることは重要である。1928年にフォン・ノイマン (John von Neumann) とオスカー・モルゲンシュテルン (Oskar Morgenstern) が後のゲーム理論となる論文を独立にしている。その後、プリンストン大学において二人は共同研究を開始し、1944年にフォン・ノイマンとモルゲンシュテルンによる「ゲームの理論と経済行動」 [Neumann 44] によりゲーム理論は始まる。岡田によれば、ゲーム理論とは「経済社会におけるさまざまな意思決定の相互依存

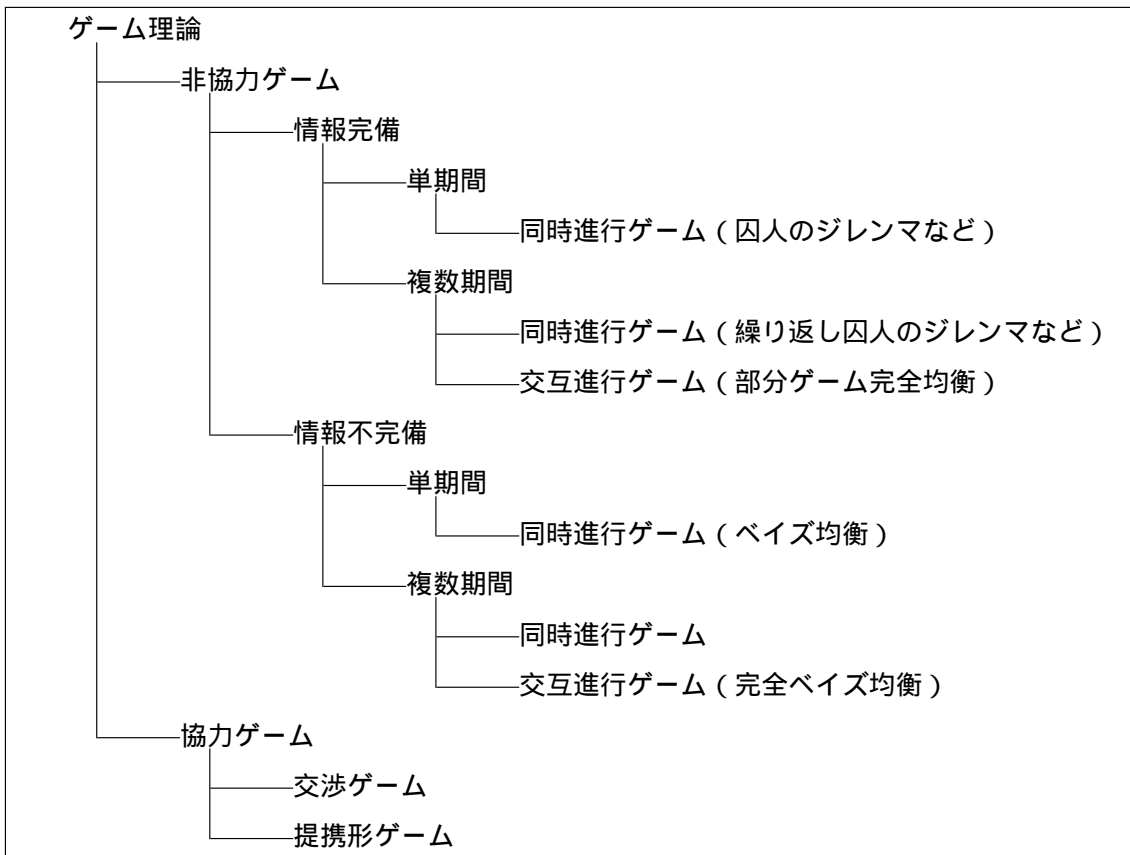


図 2.1. ゲーム理論の分類

関係を数理的な厳密な方法論を用いて分析する理論」である [岡田 96] .

ゲーム的状况の問題構造は、多くの学問分野を横断する学際的で総合的な理論とみなされており、先に挙げた研究分野を考える上で非常に強力なツールとなる。

2.4.1 ゲームの種類

ゲームには、ゲームをプレイするプレイヤー同士の関係、プレイヤーの持つ知識の範囲、プレイするタイミングなどによっていくつかの種類に分類することが可能である。図 2.1([鈴木 99]を参考とした) にゲームの分類を示す。

一般にゲーム理論は、非協力ゲームと協力ゲームに大別される。非協力ゲームでは、自らの利得を最大化することを目的として、プレイヤーがそれぞれ独立に他のプレイヤーとは協力せず戦略を決定する。協力ゲームでは、個々のプレイヤーは、他のどのプレイヤーと協力するべきか、またその際協力して得られた利得はどのように分配するか、などに焦点をあてている。

実際の社会を観察すると、非協力ゲームも協力ゲームもどちらの場合もありうる。互いに協力を前提としない非協力ゲームから、互いの協力を行う協力ゲームへの接続は一つの研究対象である。しかし、実際問題としては、協力ゲームはある程度の協力の前提をおかなければ、な

かなかそのような問題設定が肯定される状況にはなりにくいのは事実である．生物などの原始的な環境や，法が確立するまえの社会などは非協力ゲームであるだろう．

前提条件の多寡という意味においては，非協力ゲームの方がより基礎的な問題であり，その領域においてシステムの振る舞い等を十分に研究した後に協力ゲームへと繋ぐアプローチが良いと考えられる．そのため本研究では，非協力ゲームの枠組みで考えている．また，その他の分類においては基本的には，情報完備，同時進行ゲームを前提としている．

非協力ゲーム

非協力ゲームとは，ナッシュ (John Nash) によれば次のように定義される [Nash 51] ．

1. プレーヤー間のコミュニケーションが不可能
2. 拘束力のある合意が不可能

現在最も受け入れられている非協力ゲームの定義は，「ゲームの展開形表現の中に記述されているものを除いてプレイヤーの間で拘束力のある合意（や他のコミットメント）が可能でないゲーム」である [Harsanyi 88] ．

拘束力のある合意が不可能であるため，コミュニケーションにより相手との口約束があったとしても，それを必ずしも全面的に信じることができない．そのため，基本的には各プレイヤーは各々の利得を独立に最大にするような戦略をとることになる．最終的にプレイヤーがどのような戦略をお互いに取り落ち着くか，というゲームの解を選択する基準がまず課題となり，それを解決することで非協力ゲームについては発展した．これまでの研究においては，非協力ゲームでは理論的な研究は解概念の精緻化を目指すことが多かった．

後に必要となるので，ここでいくつかの解概念について触れておく．

- 支配戦略均衡解
- ナッシュ均衡解
- パレート最適解
- 進化的安定戦略

支配戦略均衡解とは，支配戦略によって均衡にいたる解のことである．相手の選択する任意の戦略に対して，必ず自己の得られる利得が最大となる自分の戦略がある場合に，プレイヤーはその戦略を選ぶべきであり，その戦略を支配戦略という．支配戦略の有無は利得構造が決定されれば，おのずと決まるものであるが，実際には支配戦略が存在する場合は多くないと考えられる．

ナッシュにより提案されたナッシュ均衡は最も有名な解概念である．ナッシュ均衡とは，他のプレイヤーがそれぞれの戦略から離脱しないならば，どのプレイヤーも自分の戦略を変更するインセンティブを持たない状態のことを指す．そのため，一旦ナッシュ均衡に陥れば，プレイヤーが他の戦略に変えることはなく安定である．

パレート最適とは、誰か他のプレイヤーの得る利得を下げることなく、自分の利得を上げることができない状態である。わかりやすく説明すれば、プレイヤーの戦略の組み合わせに対してそれぞれのプレイヤーに返される利得を示す行列があったときに、あるセルの利得組み合わせより望ましいものは存在しない状態である。

進化を考える上で一つの重要な解概念はメイナード・スミスによる ESS (Evolutionary Stable Strategy) である [Maynard Smith 82]。ESS は戦略が進化を行う進化ゲームにおける解概念である。進化ゲームにおいて、ある戦略が他の戦略の侵入に対して安定である場合、それを ESS (進化的な安定な戦略) という。

2x2 ゲームの種類

非協力ゲームにおいて最も有名なものは後述する囚人のジレンマゲームであるが、その理由はプレイヤーの意思決定においてジレンマが存在するからである。最も基本的な、二人のプレイヤーの選択肢が2つの戦略形ゲームの場合については、そのような興味深い状況となるゲームは網羅的に調べられている [Rapoport 66]。同点となる利得要素が存在しないとした場合には、ジレンマゲームは4つしか存在せずそれぞれ、囚人のジレンマゲーム、チキンゲーム、指導者ゲーム、英雄ゲームと名前がついている。これらのジレンマゲームは、政治・経済の状況を説明する場合にしばしば取り上げられる。特に、囚人のジレンマゲームとチキンゲームがそのような状況を表すことが多いといわれている。

囚人のジレンマゲーム

表 2.2. 囚人のジレンマの利得行列

		Player2	
		C	D
Player 1	C	R	T
	D	S	P

C:協調 (Cooperate), D:裏切 (Defect)

また $T > R > P > S$, 且つ $2R > T + S$

1950年代に Albert Tucker によって定式化された囚人のジレンマゲームは、多くの現象を説明する社会生活の本質的なモデルとして扱われることも多く、ゲーム理論において最も有名なゲームの一つである。

囚人のジレンマの説明は以下のようなものである。

ある事件において共犯と思われる二人の人間が捕まった。そこで、警察は二人を別々の部屋において取調べを行う。その際、取調官は彼らに次のような条件を示す。

- 自分が黙秘して、相手も黙秘すれば、懲役1年
- 自分が黙秘して、相手が自白すれば、懲役3年
- 自分が自白して、相手も黙秘すれば、懲役0年（釈放）
- 自分が自白して、相手が自白すれば、懲役2年

ここで懲役年数の数字の絶対値自体には意味はない。しかし、懲役年数の相互の大小関係がジレンマ関係になっていることが重要である。相手がもし黙秘してくれるなら自分も黙秘することによって即釈放となり双方にとって（社会的な倫理はともかく、お互いにとって）好ましい。しかし、だからと言って、黙秘を選択し、相手がもし自白すれば年数が最も長い懲役刑を受けることになる。

この利得の大小関係を保証しつつ抽象化した利得表が表2.2となる。表2.2は、プレイヤーが二人で、プレイヤーが受け取る利得にプレイヤーによる違いがないため、一般に二人対称ゲームと呼ばれる。表においては黙秘をC (Cooperate:協力)とし、自白をD(Defect:裏切)としてある。ここで、それぞれのプレイヤーについてどちらを選択すべきか考えてみる。プレイヤー1となって考えてみると、プレイヤー2がCを選択した場合に得られる利得は、自らがCを選択するならばR、Dを選択するならばTである。このとき $T > R$ の関係があるためプレイヤー1はDを選択すべきである。同様にプレイヤー2がDを選択した場合、 $P > S$ の関係からプレイヤー1はDを選択すべきとなる。Dはこのゲームにおいては支配戦略であり、相手が何を選択しようともDを選択すべきとなる。そして利得構造は対称であるため、プレイヤー2も同様にDを選択する。このようにして、合理的であるならば、両プレイヤーとも裏切を選択する。元の寓話に戻ると、結局このような条件下では、双方とも自白することが最適な戦略となる。

ここでの問題は、もしそれぞれのプレイヤーが合理的であるならば、両プレイヤーとも裏切を選択するため、パレート劣位な唯一のナッシュ均衡(D,D)となり、他にもっと好ましい状況が存在するにも関わらずそれを社会的には選択できない、というジレンマが発生する点にある。

しかし、このように理論的には裏切同士が安定的なナッシュ均衡解であるにもかかわらず、現実の社会に目を向けると協調行動が観察されるのはなぜかという疑問が発生する。この点において囚人のジレンマは主に研究されてきた。

また、ゲームそれ自体も興味深い対象ではあるが、それを利用して他の現象を説明するために用いられることもある。

n 人囚人のジレンマゲーム

二人対称ゲームを既に見たが、ここでは n 人のゲームに拡張した場合を見る。基本的なジレンマ構造は同様であるが、 n 人であるため以下のような工夫がこらされている。

各エージェントの利得 f は、当エージェントの行動 $X \in C, D$ と、他の $(N-1)$ 人のエージェントの中でCを選択しているエージェントの人数 c によって決まり、 $f(X, c)$ であらわさ

れる。

n 人囚人のジレンマゲームは以下の条件を満たす。

1. $0 \leq c \leq N - 1$ の任意の c に対して, $f(D, c) > f(C, c)$ が成立する
2. $f(C, N - 1) > f(D, 0)$ が成立する
3. $f(C, c)$ と $f(D, c)$ はともに c の単調増加関数

これらの条件を満たすことにより, プレーヤー間にジレンマ構造が生まれる。このような多プレーヤーによるゲームも社会においては多い。確かに多数の意思主体がプレイする一つのゲーム構造で考えられる創発現象ではことも多いと思われる。しかし, その考察は二人ゲームの次の段階と考え, 本研究では, 二人ゲームに注目することとする。

協力ゲーム

協力ゲームでは, 基本的には, 他のどのプレーヤーと協力するか, 協力して得た報酬をどのようにプレーヤー間で分配するか, が研究対象である。

本研究では, 非協力ゲームとの関連を除いて, 具体的な内容については触れない。

非協力ゲームと協力ゲームの関係

ナッシュの論文 [Nash 51] の最後では, 非協力ゲームの枠組みで交渉問題を分析できることが触れられ, 次の論文 [Nash 53] でそれが実際に可能であることを示した。この論文では, 交渉を2段階の非協力ゲームとして定式化し, このゲームの均衡点としてナッシュ交渉解を実現している。ナッシュ均衡の拘束力が協力ゲームの前提の拘束につながるというわけである。このように非協力ゲームから協力ゲームに接続する試みをナッシュプログラムと呼ぶ。非協力ゲームの定義では, 拘束力のある取り決めはできない。そのため, 理論的にはナッシュが示すような方法で協力ゲームへの変換を行わなければならないだろう。

しかしながら, 人間がゲームを行う状況を考えた場合には, 必ずしもその限りではない。例えば, チープトークの導入により協力の可能性が高まることが知られている。ここでチープ・トーク (Cheap Talk) とは, ゼロコストもしくはほとんど無視できるような低コストでの相互情報交換と言う意味で用いられる。

実際に行われた人間による囚人のジレンマ実験の結果をみると, 相手との選択前コミュニケーションを被験者に許可した場合, 許可しない場合に比べて協力を選擇する確率は大きく上昇する。さらに「お互いに裏切らない」という(拘束力のない)約束を許可されている場合, 約束ができない場合より協力の選擇率は高くなる。これらからコミュニケーションにより, 相手が協力するという意図に関する情報が得られるならば, 協力が起こりやすくなる事がわかる [Frank 93]。

このように人間の一種合理的ではない意思決定により, 協力の可能性がある。これは協力が生存の可能性を高めてきたであろう, 進化的な理由によるのかもしれない。

表 2.3. 財の分類

		競合性	
		有	無
排他性	有	私的財	クラブ財
	無	commons	私的財

2.4.2 実社会におけるゲーム

ここまで述べてきたゲームがどのように実際の社会で行われることがあるのだろうか。

例えば、有名なものに共有地の悲劇の問題がある。共有地の悲劇は、生物学者ガーネット・ハーディンが1968年に発表した [Hardin 68] 問題である。表 2.3 に示すように、経済学では財の性質を一般に排他性と競合性で区別する。排他性も競合性もあるものは私的財であり一般の商品などである。排他性も競合性もない純公共財であり、国防警察・消防などが例として挙げられる。そのほかの二つは通常コモンズと呼ばれることがあるが、狭義のコモンズは排他性はないが競合性がある。排他性はあるが競合性はないものはクラブ財と呼ばれる。その他に希少性がないため経済学で扱わない自由財がある

共有地の悲劇と呼ばれる現象は、このうち競技のコモンズにおいて発生する。コモンズである、牛が食べるような草が生えている共有牧草地を想定する。牛を飼う農家がプレーヤーとして存在し、牛を共有地に連れて行き草を食べさせるか、自粛するかを選択できる。その場合、多くの農家が自分の牛を太らせようと共有地の草を食べさせるため、共有地の草がなくなってしまう、というものである。

また、他にも「集団が大規模になればなるほど集団の目標は実現されにくくなる傾向がある」というオルソン問題をゲーム理論を用いて定式化した研究 [木村 02] などもある。基本的に、社会的ジレンマは、集団における協力行動はフリーライダーとの戦いである。

このように社会においては多くのジレンマが存在するが、比較的研究が多いものとしては、国際関係問題がある。国際関係の分析においては、非協力ゲームがそのまま適用されるようなケースが多々ある [Snyder 78]。

本研究では非協力ゲームをとりあげる。その理由は、先に述べたように、非協力ゲームの方がより基礎的な問題であり、その領域においてシステムの振る舞い等を十分に研究した後に協力ゲームへと繋ぐアプローチが良いと考えられるからである。

2.4.3 情報

さて、ゲームの形式は上に述べたような非協力ゲームと協力ゲームが存在するが、各プレーヤーが戦略を選択する際にどのような情報に基づいてその戦略を選択するのか、という点につ

いても整理が必要である。

大きくは次のように情報概念は分類することができる。

完全 (perfect) 情報集合はすべて一点集合である

完備 (complete) ゲームのルール(ことに利得)が共通知識

対称 (symmetric) どのプレイヤーも他と異なった情報を有しない

確実 (certain) 偶然手番はない

完全情報とは、展開型ゲームについての情報概念である。正確には情報集合と呼ばれるもので定義されるが、簡単に言えば、それまでに発生したゲームの結果をプレイヤーは全て知っていることを指す。

すべてのプレイヤーはゲームのルールを完全に知っていて、さらに「他のプレイヤーもゲームのルールを知っている」ことをすべてのプレイヤーは完全に知っている、そのときゲームのルールは全てのプレイヤーの共有知識となっている。そのような状況を情報完備という。

対称情報とは、ゲームに参加するプレイヤーは全て同一の情報を持っていることである。したがって、戦略を決定する際に特定のプレイヤーが有利であるということはない。しかし、実際には情報の非対称性がほぼ全ての場合に発生する。何かをする場合に、自分と他の誰かがその対象について知っていることが異なることは多い。それによりそのゲームに参加するプレイヤーには全く違ったゲームに見えることになるかもしれない。特に経済行為は情報の非対称性を利用して活動を行うことが多く、経済学の分野では近年盛んに情報の非対称性に関する研究が行われている。

偶然情報とは、展開型ゲームにおいて、ある手番がプレイヤーによって選択されず、偶然に支配されることを指す。そのような偶然の有無を示している。

本研究においては、情報は、完全・完備・対称・確実であることを仮定している。

2.4.4 合理性

ゲーム理論においては、ゲームをプレイするプレイヤーは合理的経済人が仮定されている。ではその合理性とは何かという点については議論がある。幅広い分野で活躍したハーバート・サイモンは合理性についての以下のように分類した [Simon 76]。

実質的合理性 (substantive rationality) 与えられた制約条件の下で、目的に照らして行動が適切であるとき、その行動は実質的に合理的である。

手続き的合理性 (procedural rationality) 行動が十分に考えられて計画されているとき、その行動は手続き的に合理的である。

また同時にサイモンは、人間は完全に合理的ではありえないとし、限定合理性 (bounded rationality) を主張した。そして、合理性は最大化ではなく満足化の原理であるとみなした。

また、アカロフは近似合理性 (near rationality) という概念を提案した [Akerlof 85]。近似合理性は「経済主体が自らの効用あるいは利益を最大化しないという意味で非合理的ではあってもこの主体が合理的に行動しないことによる損失はきわめて小さい」というものである。この仮定のもとでは、惰性や経験に基づいて完全に最適な行動をとらない意思決定者でも、ゆっくりと再考を重ね行動すれば、最適行動と比べれば損失を被るが、その損失の大きさは大変小さい可能性があることを指摘した。

主流経済学においては合理的経済人が仮定されているが、このように人間は合理的足りえず、いくらかの程度の違いはあれ合理的でない行動をとることが認識されていた。その流れを汲むものが行動経済学などである。

さて、以上では合理性とは言いつつも、その合理性自体は明確とは言いがたい。サイモンの定義では、何らかの目的を前提とし、それに沿った行動が合理的であることを意味している。つまり、サイモンは合理性を行動の属性と限定している。しかし、合理性概念は必ずしも行動の属性とは限らない。

例えば、プレイヤーの属性、プレイヤーの選好、プレイヤーの計算能力等、様々なものが考えられる。金子は、プレイヤーの計算能力、論理・推論能力を合理性に限定し、そのほかの問題は派生的な意味とすることを提案した [金子 05]。他の問題は、基本的には何らかのメトリックであり、プレイヤーが十分合理的ならば、その問題も合理的になる、というスタンスである。

本研究においては、仮に合理性をどの属性と定義した場合にもプレイヤーは合理的ではない。本研究では構成論的アプローチをとり、エージェントは遺伝子などから決定される意思決定機構をもつ。そのため、必ずしも上で説明したような合理性は持たない。もちろんプレイヤー自身はその意思決定が生存につながると信じて行うわけであるが、意思決定の際、生存という目的に照らしてその妥当性のチェックは行わないし、恐らく妥当ではない決定は多々ある。

2.5 意思決定モデル

意思決定主体はどのように意思決定をするかの枠組みをここでは述べる。基本的に意思決定を行う際の具体的なアルゴリズムは、後の章で必要となった際に説明している。

2.5.1 評価軸

人間にはあらゆる欲求があり、異なる欲求の間で比較や選択を行いながら人間は行動するという考え方がある [Jensen 94]。人間の行動モデルを考える上では、欲求をどのように扱うべきだろうか。欲求の根源には様々なものがあると考えられるが、経済学においては、便宜上、その欲求を貨幣を得ることに収束させることによって、数理モデルとして扱うことが可能となっている。恐らくは n 次元欲求を何らかの変換で 1 次元に投射しなければ、有形無形に関わらず財の贈与が、言い換えれば他の人間とのコミュニケーションが困難になるからであろう。

さて、一次元に投射したところでそれはどのような性質を持つかという点についても議論は残る。選好と効用という概念である。

効用には序数的効用と基数的効用があり、前者が選好に対応する。序数的とは、 a と b を比較して a の方が好ましい、というように物事を判断していくことである。基数的とは、 $a = 100$ というように数値的に嬉しさ・好ましさをあらわすことを指す。選好に関しては、次のような条件を満たすことで、例えば消費者が合理的である、というような仮定をおくことになる。ただし、必ずしも満たす必要はなく、扱う対象によって必要となる条件は異なる。

完備性 すべての選択肢 a, b について、 $a \leq b$ または $b \leq a$ が成り立つ。

推移性 すべての選択肢 a, b, c について、 $a \leq b$ かつ $b \leq c$ ならば、 $a \leq c$ が成り立つ。

完備性とは、意思決定主体がすべての選択肢について嗜好を表明できることを意味し、推移性は選好関係が首尾一貫していることを意味する。一般に効用概念は、上記二つの条件を満たすので選好概念の方が概念としては大きい。

効用関数とは

$$\text{選好関係 } a \leq b \text{ ならば } u(a) \leq u(b) \quad (2.1)$$

となる関数として定義される。

基本的に工学的な視点でモデルが作られる場合、おおよそ効用関数を暗黙のうちに前提としていることが多い。強化学習の報酬などがその一例である。効用関数の性質としては、工学分野においては線形関数がほとんどであり複雑な関数はなかなか見られないが、経済学分野では、効用関数としては様々なものが知られている。例えば、効用関数に関する仮定として、単調性、凸性、微分可能性などを設定し、それを満たす関数を選択することになる。

本研究では、一次元の基数的な効用と線形な効用関数を仮定し、意思決定主体はその効用を元に意思決定するものとする。

2.5.2 不確実性

なんらかの行動を選択するとき、「確かではない」情報をもとに選択する必要にせまられることがある。例えば、ある行為に対して得られる利得が確率的にわかる場合、期待値を算出することによって最も期待値が高い行動を決定することができるかもしれない。しかしそのような計算が不可能なこともあるだろう。これは情報の不確実性に依存する問題であり、情報の不確実性についてもここで整理しておく。

フランク・ナイトは、不確実性とリスクの違いを明確にした [Knight 21]。日常生活では不確実性とリスクは同義で使用されることが多いが、厳密には、不確実性とはこれから起こることが確実でないことであり、リスクとは将来起こりうるものが確率的に予測できることである。

ナイトの分類によれば確率も三つに分かれる．純粹に数学的組み合わせによって求めることができる先驗的確率，蓄積されたデータに基づく統計的確率，もう一つは推定である．ここで推定は個々の個人に依存するものであり，ベイズ推定を行うなどして得られる主觀的確率に相当する．

本研究では，一つのエージェントモデルを除いて確率を利用した意思決定は行わない．一つのエージェント・モデルでは，ベイズ推定を利用した主觀的確率を元に意思決定する．

2.5.3 意思決定の基準

意思決定のための選択肢の評価軸として効用があること，不確実性を確率として扱うことができることがわかった．そのとき行動をどのように選択するか，その基準もまた様々なものがある．以下にそのいくつかを示す．

自らの選択する行動に対してどのような結果がどのような確率で得られるか，という情報がある場合には以下の基準がある．

- 期待値基準
- 満足度基準

期待値基準は，ある結果が発生する確率と得られる利得から期待値を算出し，最も高い期待値を得られる行動を選択するものである．満足度基準は，ある行動をとった時に満足する基準を得られる確率を元に選択する．具体的には，次のような満足度を満たす確率を各行動ごとに求め，その最大となる行動を選択することになる．

$$\text{満足度(行動)} = \sum_{\text{考えうる状態}} P(\text{効用} > \text{満足度基準}) \quad (2.2)$$

自らの選択する行動に対して得られる結果が「不確実」である場合，次のような基準がある．

- ラプラス基準
- ワルド基準(マクスミン基準)
- ハーヴィッツ(フルビッツ)基準
- サヴェッジ(サベージ)基準

事象が「不確実」な場合，得られる結果が発生する確率が不明なため期待値を出すことはできない．そこで，全ての事象が同じ確率で発生するとみなし，利得の平均の最大の戦略をとるのがラプラス基準である．

ワルド基準は，その行動を選択した場合に最悪となる利得を，各行動について比較し，その値が最も大きくなるものを選択する基準である．最悪ケースを考えそれを緩和するものを選択するポリシーである．これに似たものとしては，マックスマックスのように，最良ケースを選択する基準などもある．

ハーヴィッツ基準は、ある行動を選んだときに得られる利得のうち最大のものと最小のものとの重み付け平均値を最大とする戦略である。この重み付けの値によって、マクスミンとマクスマックスの間をとる。

サヴェッジ基準とは、全ての行動に対して、それぞれ得られる結果に対し最大のものを0とし、他のものを最大値からの差とし、それを残念度 (regret) とする。最大残念度を最小とする行動を選択する。

本研究では、一つのエージェントモデルに関しては期待値基準を用いる。その他に関しては、得られる結果を元に意思決定するモデルではない。

2.6 集団の意思決定

集団における意思決定についてここでは述べる。集団において何らかの意思決定を行うということは、その意思決定に従うという合意が必要である。そのため、非協力ゲームのような状況では集団での意思決定はありえない。また、協力ゲームのように集団の構成や利益の分配というフェーズともまた異なり、一旦何らかの理由で作られた集団において、意思決定が行われる。

集団での意思決定は、社会選択理論という分野でこれまで研究されてきた。社会を構成する各個人はありうる社会状態についての選好を持ち、社会的な決定とは選好順序の集約という位置づけである。選好順序の集約方法は社会的厚生関数と呼ばれる。

2.7 まとめ

本章では、まず初めに、個人の意思決定や集団の振る舞いに関連する研究分野をみた。本章で触れた領域において通底するのは、第一章にも示したような一般システムの構造である。多くの要素から構成されるシステムにおいて、要素間にはゲーム的状况が設定されており、そのゲームに対して要素は行動を行う。そのような問題構造において、創発的な現象を構成論的アプローチによって調べるために必要となる、ゲーム的状况、要素の行動モデルの整理を次に行った。ゲーム的状况は社会において多々存在するが、存在する4つのジレンマゲームのうち囚人のジレンマゲームを特に説明した。次に個人の意思決定を考える場合の基礎となるモデル概念を概観した。

本研究の位置づけは以下のようなになる。工学的な応用などを考えれば、協力ゲームや社会的選択を扱うのが良いと考える。しかしながら、論文の目的は創発現象に関する知見を得ることであり、その目的においては、社会の発展過程等を鑑みるにより原始的な段階であられる非協力ゲームの方がより基礎的な問題であり、題材としてふさわしいと考える。非協力的な状況に関してシステムの振る舞い等を十分に研究した後に協力ゲームへと繋ぐアプローチが良いと考えられるからである。

特に、空間構造上における囚人のジレンマゲームについて今後論じる。囚人のジレンマという社会において頻発するゲームを題材にして、社会に存在する要素間の構造がどのように要素に影響し、また影響されるのか、という点について見ていく。

以上の条件から本章で整理した各点についてここでまとめると次のようになる。本研究では、非協力ゲームである囚人のジレンマゲームをとりあげる。ただし、完全・完備・対称・確実情報とし、プレイヤーは必ずしも合理的ではない。

プレイヤーは、一次元の基数的な効用と線形な効用関数を仮定し、その効用を元に意思決定するものとする。以降の章ではいくつかのゲームをプレイするプレイヤーであるエージェントモデルを提案するが、一つのエージェントモデルを除いて確率を利用した意思決定は行わない。その際、主観確率と期待される利得から計算される期待値基準を用いる。

第3章

空間構造上における囚人のジレンマゲーム

3.1 はじめに

本章では、空間構造上における囚人のジレンマゲームの基本的な部分について触れる。

先にも説明したとおり、実社会においてプレイヤー間の利得構造を整理してみると、多くの現象が囚人のジレンマゲームと判明することもあり、多くの研究分野でこのゲームは知られている。研究の興味の対象となっているのは、多くの場合にはナッシュ均衡である裏切手番を選択せずに、協力手番を選択するようになるメカニズムである。

よく知られるのは繰り返しゲームにすることによる説明である。合理的なプレイヤーは繰り返されるゲームの中で将来の利得を考えるため、協力の方が合理的に考えて利得が多くなるというものである。

また、別の説明が空間構造の導入である。ゲームを行うプレイヤーを空間上に配置し、近傍のプレイヤー同士で対戦させ、プレイヤー集団を進化させていくことにより、初期には裏切戦略が支配的であった集団の中に協調戦略が発生する。空間構造を扱った従来の研究には、二次元格子平面を扱ったものが多い。例えば、2種類の単純な戦略 (All C, All D) を行う個体によるもの [Nowak 92]、メタ戦略を持つ個体によるもの [Lindgren 94]、模倣の方法による差異を観察したもの [Nakamaru 97]、ノイズに注目したもの [石淵 00]、N 人囚人のジレンマに関するもの [Suzuki 03] などである。近年、スモールワールドネットワークなどのネットワーク理論に関する研究が盛んになってきてからは、ネットワーク上でのゲームも研究されるようになってきた (例えば [Watts 99][Masuda 03])。

このような多くのプレイヤーによってゲームが行われる場合、解析的な研究はなかなか困難であるが、空間構造が比較的整然としたもの場合はそのような研究も存在する。例えば、各頂点が同一リンク数を持つ正則グラフ上でのゲームについての解析的な研究 [Ohtsuki 06] などである。

囚人のジレンマゲームは、それ自体の協力メカニズムが主な研究対象ではあるが、それを利

用して他の現象を説明するために用いられることもあり、例えば、baldwin 効果の確認したものの [Suzuki 04] などである。このような研究においては、個体の一般的な活動として囚人のジレンマをとらえ、そのゲームを行うことによって他の現象がどのように影響をうけていくのか、という視点をもつ。

本研究には、協力現象を説明するために空間構造がゲームをプレイするエージェントに与える影響と、空間構造上で囚人のジレンマゲームがエージェントに与える影響との二側面がある。

3.2 空間構造

本研究では、空間構造上にエージェントが存在することを前提としている。そのためここでは、空間構造についてまとめておく。

3.2.1 スモールワールドネットワーク

スモールワールドとは、もともとは、出会った二人の人間の間に共通の知人がいる場合が多いことを指す^{*1}、社会心理学の分野の概念であった。歴史的には、六次の隔たりの実験のミルグラム [Milgram 67]、共同体の中での情報の伝播に関する考察グラノヴェッター [Granovetter 73] などに端を発する。

1998年に Duncan Watts はパス長とクラスタ係数の二つの特徴量を用いて、それまで経験的に知られていたスモールワールドネットワークを定式化した [Watts 98]。パス長とクラスタ係数は次のように定義されている。

パス長 全ての2ノードの組み合わせにおける最短パスの長さ(ホップ数)の平均値

クラスタ係数 あるノードが n 個のノードとリンクで結ばれている時、それら n 個のノード間のあり得る組み合わせのうち実際にリンクで結ばれている割合の平均値

ただし、ここでノード同士を結ぶものをリンクと呼んでいるが、エッジと呼ばれることもある。本論文ではリンクに統一している。

ここで、ノード数とリンク数を固定し、ランダムさのみを変化させると、ネットワークは図 3.1 のようになり、次のような特徴を持つ。

それぞれのノードが近傍のノードとのみ結ばれている、ランダムさが全くないネットワークでは、パス長、クラスタ係数とも大きい。逆に、全てのリンクがランダムに決定される、完全にランダムなネットワークでは、パス長、クラスタ係数とも小さい。スモールワールドネットワークはこの中間領域であり、パス長は小さく、クラスタ係数は大きいという特徴がある。

このようなスモールワールドネットワーク構造は、これまでに様々な分野で観察されている。

^{*1} 友人との会話の中で共通の知人がいることを発見し「狭い世界ですね (It's a small world!)」という会話

次章の準備のために，ここで数学的な表現を押さえておく．

グラフ G は頂点とリンクを持つ． $V = \{1, \dots, n\}$ は頂点の集合， g はグラフ G の隣接行列である． $g_{i,j}$ は頂点 i と j のペアを意味する $i, j \in V$ を結ぶリンクである．もし $g_{i,j} \neq 0$ ならば，リンクは存在し， $g_{i,j} = 0$ の場合にはリンクは存在しない．

ここで， $d(i, j, g)$ を i と j を結ぶ最短パスの長さとする．パス長 L はネットワーク上全ての頂点のペアの平均である*2． L は正確には式 (3.1) で表される．

$$L = \frac{1}{n(n-1)} \sum_i \sum_{j \neq i} d(i, j, g) \tag{3.1}$$

$N^1(i, g) = \{k \in V | g_{i,k} \neq 0\}$ は i と接続されている頂点の集合である． $E(i, g) = \{(j, k) | g_{j,k} \neq 0, k \neq j, j \in N^1(i, g), k \in N^1(i, g)\}$ は， $N^1(i, g)$ のその間にリンクが存在する頂点の組み合わせ集合である．

簡単のために， $l_i = |N^1(i, g)|$ を i に接続しているリンクの数とする．クラスタ係数 C は式 (3.2) のように定義される．

$$C = \frac{1}{n} \sum_i \frac{|E(i, g)|}{l_i C_2}, \quad l_i C_2 = l_i(l_i - 1)/2 \tag{3.2}$$

スモールワールドネットワークは，ランダムネス p に応じて確率的にリンクの接続先を変化させることによって出現する．ここで，グラフ上の頂点とリンクの数を固定とすると，図 3.1 のように，ランダムネスの変化によりネットワーク構造は変化する．ランダムネス $p = 0$ の時，全ての頂点は近隣に相互接続しあっている．このネットワークはレギュラーネットワークと呼び， L と C はともに大きい．一方，ランダムネットワークと呼ばれる $p = 1$ の時には L と C はともに小さい．その中間領域では， L は小さく， C は大きくなり，スモールワールドネットワークと呼ぶ．

スモールワールドネットワークは，これまで多くの場所で観察されている．

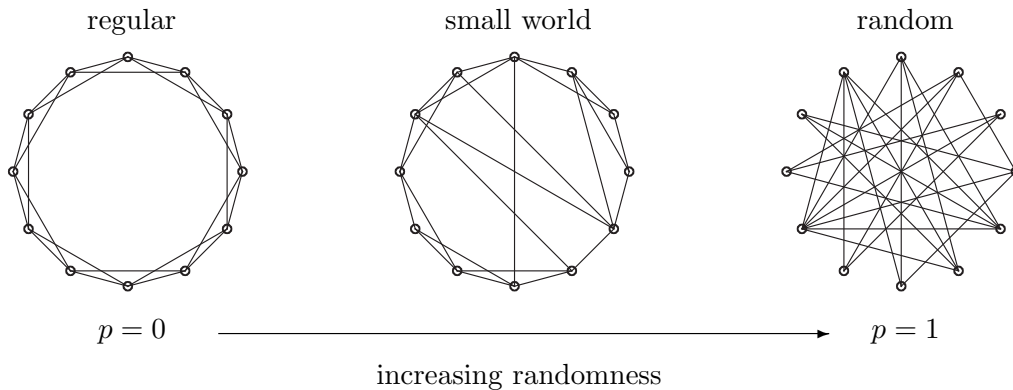


図 3.1. スモールワールドネットワーク

*2 平均あるいはメジアンという定義が存在する [Watts 99]

3.2.2 スケールフリーネットワーク

スモールワールドネットワークモデルが提案されてすぐ、異なるネットワークモデルが提出された。バラバシらによるスケールフリーネットワークである BA モデル（優先的選択モデル）[Barabasi 99] はグラフのリンク分布がべき乗則に従うという特徴がある。これはグラフの成長モデルであり、徐々にリンクを持ったノードをグラフに追加していく。追加の際、既存のノードのうちリンク数の多いノードに優先して接続するというもので、この操作によりスケールフリー性が出現する。スケールフリーとは、各ノードの持つリンクの数とランクの関係がべき乗になる特性のことである。実際のネットワークを観察すると、こちらのモデルの方が当てはまることも多く、スモールワールドネットワークと並びよく参照されるモデルである。その後、スケールフリー性をもつネットワークの作成方法については CDD (Coupled Duplication Divergence)[Vázquez 03]、CNN (Connecting Nearest Neighbor)[Vázquez 03]、LPA (Linear Preferential Attachment)[Oliveira 05] など様々なモデルが提案されている。

本研究では、スケールフリーの特徴であるリンクのべき乗則を満たしつつ、スモールワールドの特徴を満たすモデル [Klemm 02a] をとりあげる。このモデルを利用することにより、リンクの分布の差異の影響が観察可能となる。

このモデルは、同じ研究者によるクラスター化されたスケールフリーモデルである KE モデル [Klemm 02b] とバラバシによるランダムなスケールフリーモデルである BA モデル [Barabasi 99] の組み合わせで作成される。

KE モデルは、成長の過程で以降の頂点がその頂点に接続できるかどうかを示す活性・非活性状態を導入し、頂点を追加することに既存の活性頂点のいずれかを非活性にする。ただし、リンク数の多い頂点は非活性状態になりにくい。この制限を加えることにより、クラスタ化されやすくなる効果が得られる。

KE モデルと BA モデルの組み合わせとは、ある一定のつなぎかえ確率 μ で KE モデルと BA モデルを入れ替える操作を行うことを意味する。この操作を行うことにより、小さい L 、大きな C 、べき乗則を満たすことができる。

μ はスモールワールドネットワークのランダムネス p と同等の役割を果たし、この値の選択によってスモールワールドネットワークの特徴を持たせることができる。

3.2.3 ネットワーク特徴量

このようなネットワーク構造においては、そのネットワークの特徴を示す特徴量がこれまで提案されてきた。

例えば、既に触れた次数 [Barabasi 99]、クラスタ係数 [Watts 98]、パス長 [Watts 98] などと共に、ネットワークの中心に関する、近接中心性 (closeness)[Freeman 79]、媒介中心性 (betweenness)[Freeman 79] などもある。

また，ネットワークにおいて，ある程度の固まったクラスターと呼ばれるサブネットワークも一つの特徴であり，その抽出アルゴリズムもこれまでいくつか提案されている [Girvan 02][Newman 04] .

3.3 エージェントモデル

3.3.1 遺伝子モデル

次に，囚人のジレンマゲームをプレイするエージェントモデルを説明する．先に説明したネットワークのノードがすなわちゲームをプレイするプレイヤーとなり，これをエージェントと呼んでいる．ネットワーク上のノードは，以降ではゲームをプレイするエージェントを表す．

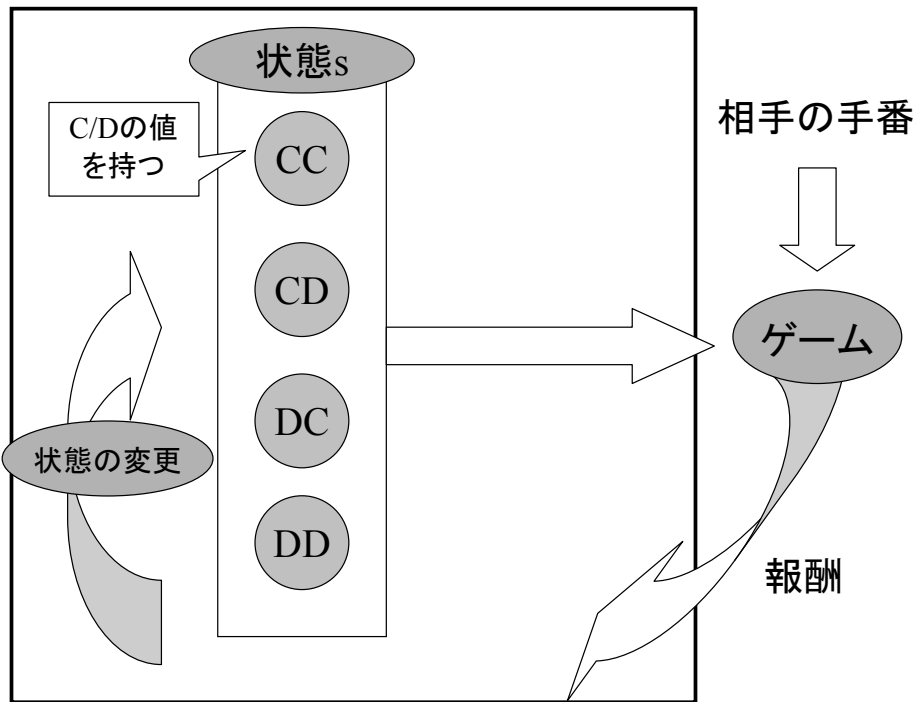


図 3.2. エージェントモデル

エージェントは，遺伝子によって決められたルールに従い，ゲーム中の刺激に対する反射によってゲームをプレイする．エージェントは遺伝子として一次のメタ戦略を一種類持つ．ここで，一次のメタ戦略とは，過去一回のゲームの記憶を保持し，そのゲームの内容を元に次回の手番を決定することを指す．過去一回のゲームにおいては，自分の手番，相手の手番それぞれにつき C もしくは D という 4 通りの場合が考えられる．それに加えて繰り返しゲームの最初の一回の手番をあわせて 5 ビットあれば，一次のメタ戦略は表現可能である．そこで，戦略は表 3.1 のように遺伝子コード化される．また初回の手も遺伝子により決定される．

以上をまとめたエージェントの手番選択プロセスを図 3.2 に示す．各状態について C/D の

表 3.1. メタ戦略例

前回の 自分の手	前回の 相手の手	戦略例			
		All C	All D	TFT	Pavlov
C	C	C	D	C	C
C	D	C	D	D	D
D	C	C	D	C	D
D	D	C	D	D	C

どちらかを示す遺伝子が存在し、決定論的に手番が選択されることになる。

3.3.2 戦略の表現可能性

これまでの研究で、囚人のジレンマをプレイする戦略で有名なものがいくつかある。ここでは説明し、上のモデルで表現可能であるか検討する。

TFT

TFT (Tit-For-Tat : しっぺ返し) は囚人のジレンマゲームにおいて最も有名な戦略である。この戦略は有名なアクセルロッドのトーナメントの勝者となり [Axelrod 85] 知られるようになった。動作は極めてシンプルで、繰り返しゲームにおいて、前回相手のプレーヤーが出した手番を次のゲームで自分が模倣するというものである。これはシンプルでありながら、相手との協調関係をうまく導出するモデルとして評価されている。

Pavlov

Pavlov 戦略 [Nowak 93] は "win-stay, lose-shift" としても知られる。この言葉が表すように、順調な時には戦略を変更せず、そうでない時に前回の自分の手番をひっくり返すというものである。具体的には、例えば、相手が協調することを期待して自分が協調を出したときに裏切られた時、あるいは相手が協調を出さずと思って裏切りを出したときに相手も裏切ってきた時、は自分の期待通りいかなかったケースであるため、前回の自分の選択が誤りであったことを認識し、次回は自分の手番を変えるというものである。逆に、それ以外の場合は自分の思惑通りいったケースであるので、前回の自分の手番と同じものを選択する。

トリガー戦略

他に有名なものとしてトリガー戦略がある。これは、繰り返しゲームにおいて初期は協力を出し、お互い協力が続いている間は協力を出し続けるが、一度相手が裏切った後は、裏切りを選択するというものである。

以上、三つの有名な戦略を取り上げたが、一次のメタ戦略で表現不可能なものはトリガー戦略のみである。また、二回以上前のゲームの記憶を持つ二次以上のメタ戦略でなければ表現不可能な戦略はもちろん存在する。しかし、これまでの研究から TFT や Pavlov 戦略の強さは一定に評価されており、本研究においても一次のメタ戦略で十分であると考え、以降はこの戦略表現を基本として扱う。表 3.2 に TFT と Pavlov が対戦した場合の対戦例を示す。ここでは、初期手番を C と D と出している。それ以降は上述の戦略に従い手番は自動的に決定される。

表 3.2. 対戦例

	1	2	3	...
TFT	D	C	D	...
Pavlov	C	D	D	...

3.4 ネットワーク構造の影響

3.4.1 シミュレーション内容

ここで想定するものは初期の非協力的な状況である。その場合、少なくとも最初は身近な個体との接触となるため、スケールフリーネットワークではなく、スモールワールドネットワークが適当であると考えられる。

そこで、スモールワールドネットワークモデルにおいて、ネットワーク構造がエージェントの戦略に与える影響を調べるために、以下のようなシミュレーションを行った。

エージェントを囚人のジレンマにおけるプレーヤーとして、ネットワーク上においてリンクにより結ばれたエージェント同士を対戦させる。

ここでは次のように世代交代を行うものとする。ある世代に、全てのリンクについて対戦が終了すると、その世代において取得した利得に応じて、2割のエージェントが確率的に死亡し、近傍のエージェントのうち最も利得の大きいものがコピーされる。ただしコピー時には0.2%の割合で突然変異が発生する。以下にシミュレーション結果を示す。図 3.3 にその様子を示す。

また、表 3.3 にシミュレーション諸元を示す。ただし、表 2.2 において $(T, R, P, S) = (5, 3, 1, 0)$ とした。

3.4.2 シミュレーション結果

ランダムネス $R=0/1$ とした場合の、平均利得の推移を図 3.4 に示す。図を見ると、最初に一旦利得が落ち込む期間が存在する。これは裏切を基調とする戦略が優勢であることに起因するが、裏切同士の対戦では高利得が得られないため、ある程度の時間が経過すると死滅し始め

表 3.3. シミュレーション諸元

パラメータ名	値
集団サイズ	400
エージェントの持つ平均リンク数	6
対戦あたり繰り返しゲーム数	100
ジェネレーションギャップ	0.1
突然変異率	0.002
世代数	2000
ノイズ	0.01

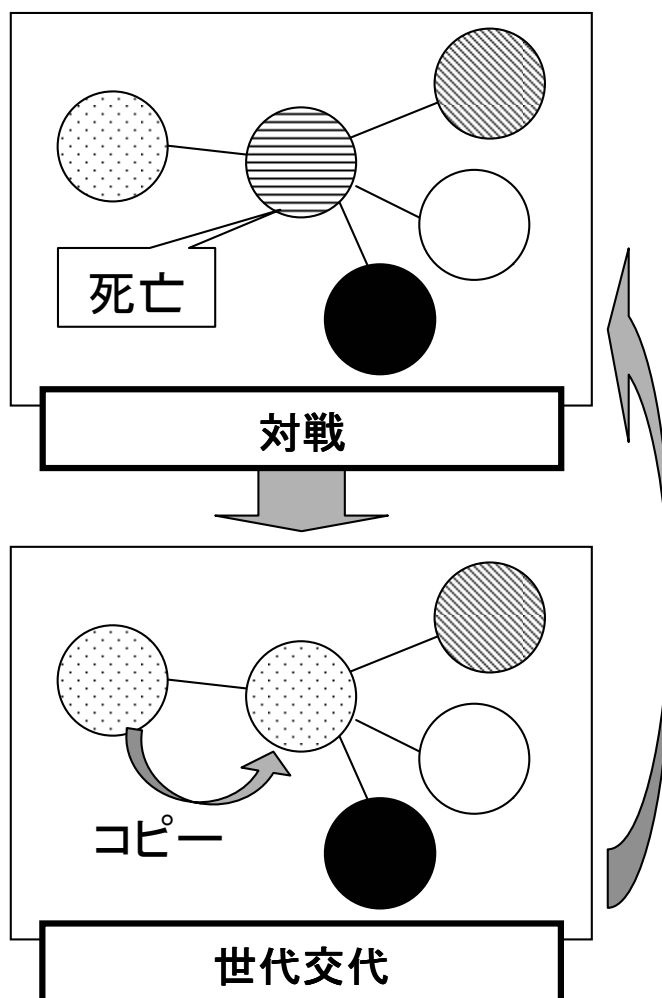


図 3.3. 各世代における手順

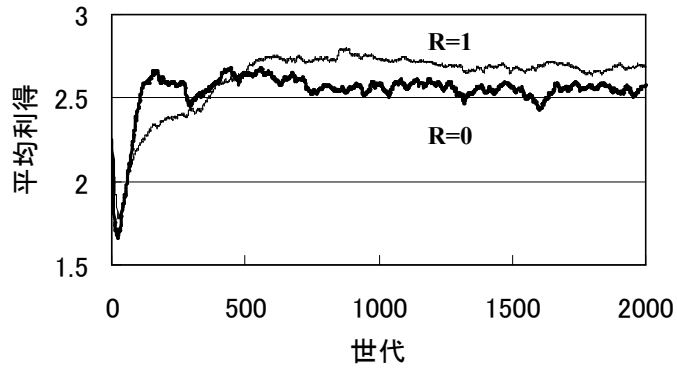


図 3.4. 平均利得推移 (R=0/1)

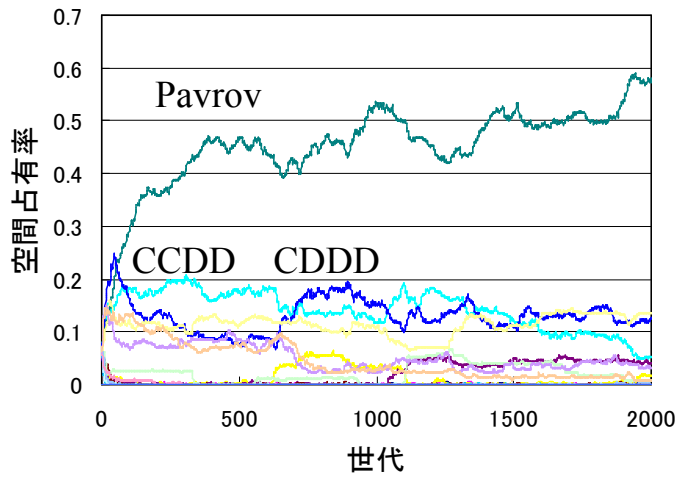


図 3.5. 世代交代例 (R=0.2)

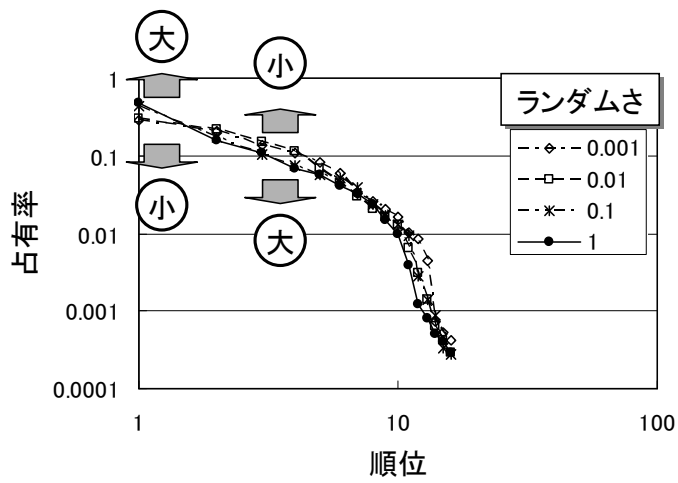


図 3.6. 時空間占有率と順位の関係

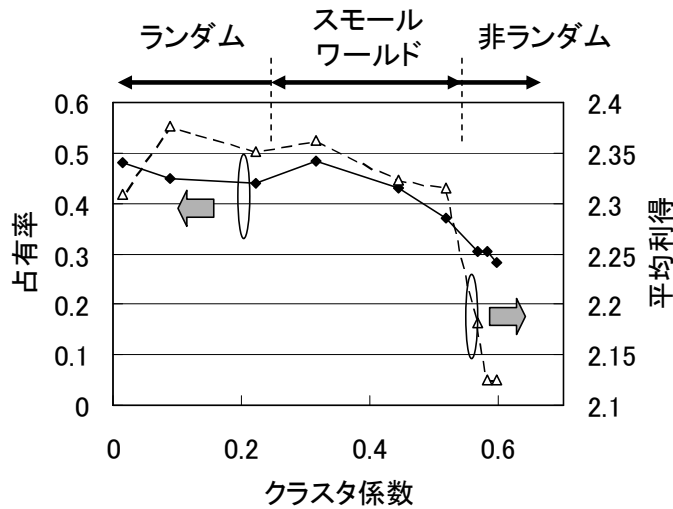


図 3.7. クラスタ係数に対する支配種の時空間占有率，平均利得の変化

ることになる．その後，協調を基調とする戦略が優位になり，平均利得は上昇し一定の値に収束した状態でそれ以降推移する．ただし，収束の際の振る舞いがネットワークのランダムネスに依存して異なっており，ランダムネスが大きくなるに従い収束は遅くなる．

それぞれの種がどのように推移するかを示す進化ダイナミクスの例を図 3.5 に示す ($L=6$, $R=0.2$)．この図にも表れているが，全ての条件において Pavlov 戦略が生き残りやすく，協調戦略が発生していることがわかった．

次に，(世代数) × (エージェント数) のシミュレーション時空間における種の占有率 (以降，時空間占有率) へのネットワーク構造の影響を調べた．

まず，ランダムネス R を固定し L の値を変化させた．図には示さないがリンクの数が少なすぎる場合を除きほぼ似た傾向であった．そこで次に，ランダムネスを変化させた場合の種の時空間占有率を調べた (図 3.6)．全てのネットワークにおいて各種の時空間占有率順位と時空間占有率の関係は，存在する 16 種の内 10 位程度でカットオフされる冪乗則に従う事がわかった．

ただし，ネットワーク構造によって若干傾きが異なる．特に上位の種においてネットワーク構造の違いによる差が顕著に出ていることに注目し，ここでは最も占有率が高い支配種に関しクラスタ係数と時空間占有率の関係を調べた (図 3.7)．これを見ると，クラスタ係数の低いネットワークの方が支配種の時空間占有率が高いことがわかる．図 3.7 に平均利得の値を重ねてみると，クラスタ係数の高いネットワークは平均利得が低いことがわかる．スモールワールドネットワークの範囲に注目してみると，減衰が始まる領域に相当する．これから，支配種の時空間占有率，平均利得が最も大きくなる範囲で，空間を越えるコストを最小にするネットワークともいえる．

3.5 まとめ

本章では、これまでネットワーク上でゲームをプレイするプレイヤーについての生態学的知見はなかったが、本研究では囚人のジレンマゲームにおける、エージェントの戦略へのネットワーク構造の影響について調べた。その結果、これまで生物において観察されたような、各生物種のランクと個体数はべき乗則に従うことが示された。また、Pavlov 戦略が構造に関わらず生き残りやすいこと、クラスタ係数が小さくなると一種類の戦略が支配的になり、平均利得が高くなることがわかった。

第4章

ネットワーク形成

4.1 はじめに

第三章においては、空間構造はスモールワールドネットワークと固定して、ネットワーク上における囚人のジレンマゲームを行った。しかし、空間構造として必ずしもスモールワールドネットワークとは限らず、他のネットワークモデルが適当かもしれない。また、静的なネットワークではなく、動的にネットワーク構造が変化することも考えられる。

そこで、本章では、ネットワークの動的な変化に着目し、囚人のジレンマゲームをプレイするプレイヤーの自律的な動きにより、ネットワークがどのように変化するかという点を観察する。

エージェントとリンクの数が固定された場合に存在するネットワークの種類について考えてみると、 n 個のエージェントから構成される無向ネットワークは $2^{n(n-1)/2}$ 通りの可能性がある。しかしながら、実際の世界では n が大きく莫大な数の可能性があるにも関わらず、多くのネットワークでスモールワールド [Watts 98]、スケールフリー [Barabasi 99] といった性質を示すことが多い。これまでスモールワールド、スケールフリーの性質を示すネットワークとなるようなネットワーク形成方法はいくつか提出されてきた。しかし、ネットワークを構成するエージェント戦略の変化は同時に考慮されてこなかった。

一方、ネットワーク形成を戦略とし自分の利得を最大にするようゲーム理論的に考察すると、完全合理性を持つプレイヤーにより形成されるネットワークは、ナッシュネットワークと呼ばれるホイールやスター型のネットワークになることが知られている [Bala 00]

プレイヤーの戦略とネットワーク形成の統合的な研究はこれまで十分でない。本稿では、囚人のジレンマゲームにおいて、エージェントの戦略変化を包含するネットワーク形成モデルを提案し、戦略とネットワークのダイナミクスを調査する。

ネットワークの形成においては、エージェントとリンクの数自体の増減を考えることもできるが、その場合ネットワークの構造と規模の要素が混在することになるため、以降ではネットワークを構成するエージェントとリンク数は固定としている。

4.2 自由生成条件におけるネットワーク形成

ここでは自由生成条件におけるネットワークの変化を調べる．自由生成条件とは，生成されるエージェントは死亡したエージェントとは完全に独立で自由であることとする．図 4.1 に各世代における手順を示す．

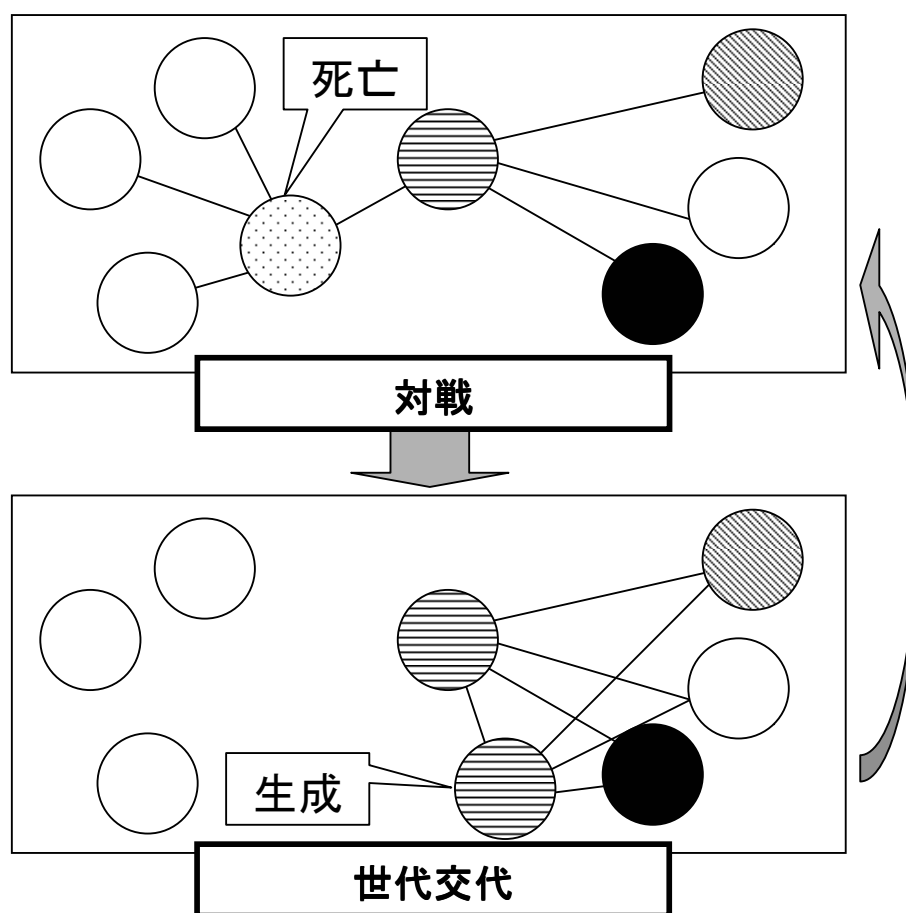


図 4.1. 各世代における手順

基本的な条件は前章のシミュレーション実験と同様である．まず，初期ネットワークを作成する．その後，各世代ごとに次のような手順をおこなう．まずリンクでつながれた全てのエージェント同士は囚人のジレンマゲームを行う．次に，その世代において取得した利得に応じて，死亡するエージェントとして2割のエージェントが確率的に選択される．エージェント生成においては，死亡したエージェントとは無関係に，ネットワーク中のエージェントの中から利得に応じて確率的に選択し，そのエージェントをコピーする．その際，リンクは，コピー元のエージェントにリンクされているエージェントの中からランダムに選択し，ネットワーク全体に存在するリンク数が各世代で一定になるように数を制限し，作成する．

以下にシミュレーション結果を示す．

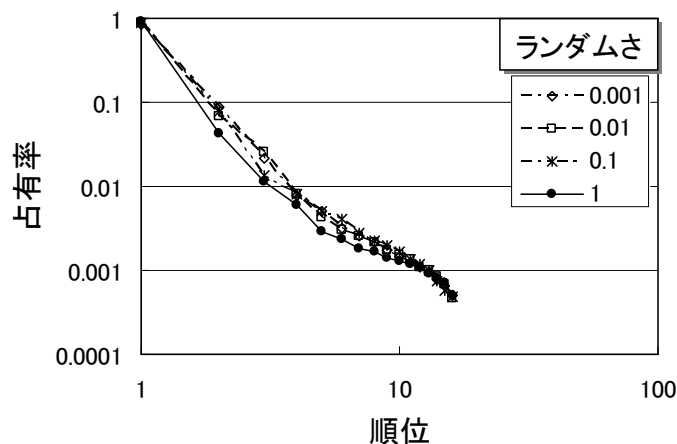


図 4.2. 時空間占有率と順位の関係

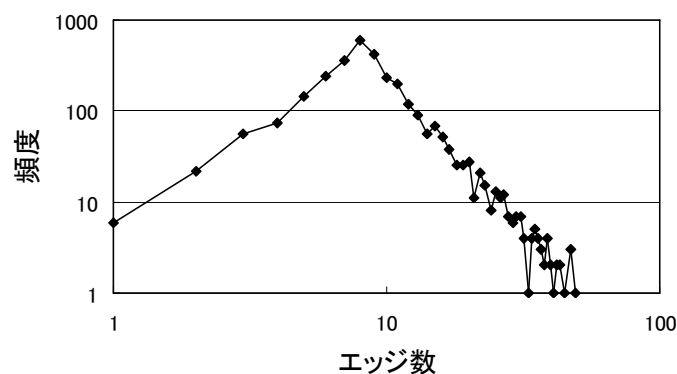


図 4.3. リンク数とエージェント頻度 (エージェント数=3000)

初期条件として、スモールワールドネットワークにおけるランダムネスの値を変化させて試行を行った。しかし、ネットワークの構造が動的に変化するため、結果としては、初期値の値にはそれほど影響を受けずにパス長は 3.5~7、クラスタ係数は 0.5~0.6、平均利得は 2.9 を超える範囲と、ある程度の範囲に収束した。

次に、種の時空間占有率順位と時空間占有率の関係を図 4.2 に示す。ネットワーク構造が固定された場合が上に凸だったのに対し、この場合にはほぼ直線もしくは下に凸の形となり、上位種がさらに支配的となっていることがわかる。

新エージェント生成方法の特徴から、各エージェントのリンク数は世代が進むにつれ変化する。図 4.3 に各エージェントのもつリンクの数とエージェントの頻度の分布を示す。グラフは冪乗則に従っていることが明らかである。ただし、べき乗則であることを確認するために、このときのみ集団のサイズを 3000 と他のケースに比べて大きめにしている。これは先に説明したスケールフリーと呼ばれる特徴 [Barabasi 99] であり、様々な場所で観察されている。

4.3 固定生成条件におけるネットワーク形成

次に、固定生成条件下におけるネットワークの変化を調べる．自由生成条件では、生成されるエージェントは死亡したエージェントと無関係であった．ここでは、死亡したエージェントのネットワーク上の位置にエージェントは生成されるものとする．そして、世代交代が終わった後に、リンク変更を行うこととする．

まず、提案するネットワーク形成モデルを説明し、その後シミュレーションについて説明する．

4.3.1 ネットワーク形成モデル

提案するネットワーク形成モデルを示す．エージェントは世代交代が終わった後で、リンク変更を行うことになるが、その作業をどのように行うかを個々では説明する．

ただし、エージェントのネットワーク知識に関する限定された知識を仮定し、2 ホップ先までのエージェント情報までを知ることができるとしている．

$N = \{1, 2, \dots, n\}$ を n 個の意思決定主体であるエージェントとする．以下、 $i, j \in N$ とする．各エージェントは m 本のリンクを所有し、所有するリンクに関して他のエージェントへの切断・接続の操作を自由に行えることとする．世代 t の時点のネットワーク $g(t)$ は、 i, j の全ての組み合わせ $i, j \in N$ におけるリンクの有無を示す要素 $g(t)_{i,j}$ からなる (式 (4.1)) ．

$$g(t)_{i,j} = \begin{cases} 1 & j \text{ is linked by } i \\ 0 & \text{absence} \\ -1 & i \text{ is linked by } j \end{cases} \quad (4.1)$$

値 1, -1 の違いはリンクの所有エージェントの違いのみをあらわし、リンクを通してなされる相互作用自体には影響しない． i が j に接続していれば $g(t)_{i,j} = 1$ 、逆に j が i に接続していれば $g(t)_{i,j} = -1$ と表される． i に接続されているリンク数を l_i とする．

ここで、 i の隣接エージェント集合は $N^1(i, g(t)) = \{k \in N | g(t)_{i,k} \neq 0\}$ である．また i の所有するリンクで結ばれた隣接エージェント集合は $N_o^1(i, g(t)) = \{k \in N | g(t)_{i,k} = 1\}$ となり、 $N^1(i, g(t))$ の部分集合である．更に 2 ホップ先のエージェント集合を考えると、隣接エージェント $j \in N^1(i, g)$ の隣接ノードと考えて、 $N^2(i, g(t)) = \{k \in N | k \in N^1(j, g(t)), j \in N^1(i, g(t))\}$ となる．

ネットワーク上では、リンクで結ばれているエージェント同士は、2 人対称ゲームを行うこととし、全てのリンクについてゲームを行う． i が j との対戦の結果に得る利得を $p_{i,j}$ とすると、 i の平均利得は $\bar{p}_i = \sum_{j \in N^1(i, g(t))} p_{i,j} / l_i$ となる．

対戦後、全てのエージェントについてリンク張替え操作を行う．この操作は概念的には物理的な移動に近い．エージェント i についての操作は、まず最初に接続エージェントの切断 (式 (4.2))、次に新規エージェントの接続 (式 (4.3)) からなる．

$$N_o^1(i, g(t+1)) = N_o^1(i, g(t)) - N_i^{del}(\in N_o^1(i, g(t))) \quad (4.2)$$

$$N_o^1(i, g(t+1)) = N_o^1(i, g(t)) + N_i^{add}(\in N^2(i, g(t))) \quad (4.3)$$

つまり，所有リンクで接続されたエージェントの中から N_i^{del} を選択し切断し，2 ホップ先のエージェントの中から N_i^{add} を選択し接続する．エージェントは，切断，接続するために評価する必要がある． i が評価する主体のエージェント， j を i から直接接続されたエージェント， k を j に接続されたエージェントとすると，評価指標は次のものがあるとする．

- 隣接エージェントの指標

$$U(i, j) = \{l_i, l_j, p_{i,j}, p_{j,i}, \bar{p}_i, \bar{p}_j\}$$

- 1 ホップ先のエージェントの指標:

$$V(i, j, k) = \{l_i, l_j, l_k, p_{i,j}, p_{j,i}, p_{j,k}, p_{k,j}, \bar{p}_i, \bar{p}_j, \bar{p}_k\}$$

さらに，上の指標だけでは不十分であると考え，その差分までを導入する．評価関数は同様に以下のように表される．

- 指標の差分も考慮した指標:

$$W(U) = \{r - s | r \in U, s \in U \cup \phi\}$$

$$W(V) = \{r - s | r \in V, s \in V \cup \phi\}$$

- 指標の評価関数:

$$H = \{\text{argmin}, \text{argmax}, \text{random}\}$$

また，エージェントは $h \in H, w \in W$ を遺伝子コード化し保持し，リンク張替え操作は遺伝子に従って行うとする． N_i^{del} は式 (4.4) に従い選択される．

$$N_i^{del} = H_x (W(U(i, x)) | x \in N_o^1(i, g(t))) \quad (4.4)$$

N_i^{add} は，2 ホップ先のエージェント集合を一度に評価する式 (4.5)，または隣接エージェント集合評価を2度行う式 (4.6) のいずれかに従い選択される．ただし，式 (4.6) に従う場合，エージェントは $h' \in H, w' \in W$ も保持する．

$$N_i^{add} = h_x (w(V(i, j, x)) | x \in N^1(j, g(t)), j \in N^1(i, g(t))) \quad (4.5)$$

$$N_i^{add} = h'_x (w'(U(j, x)) | x \in N^1(j, g(t)))$$

$$\text{where } j = h_j (w(U(i, j)) | j \in N^1(i, g(t))) \quad (4.6)$$

ただし， N_i^{del}, N_i^{add} に複数の要素が含まれる場合，および $h = \text{random}$ の場合は，集合の要素から一様ランダムに一つ選択することとする．

以上のネットワーク形成操作を， $t = 0$ の時にランダムネス 0 のスモールワールドネットワークから開始し，各世代，全対戦後，全てのエージェントについて行う．

4.3.2 シミュレーション実験

各エージェントは、自分と相手の戦略を1回記憶する5bitのメタ戦略と、上述のリンク張替え操作遺伝子を持つ。エージェントはそれぞれ独立にメタ戦略、リンク張替え操作を行う場合を調べた。

表4.1にシミュレーション緒元を示す。

表4.1. シミュレーション緒元

パラメータ	値
集団サイズ	400
エージェントの持つ平均リンク数	6
対戦あたり繰り返しゲーム数	100
ジェネレーションギャップ	20%
突然変異率	0.02%
世代数	2000
囚人のジレンマ利得パラメータ (T,R,P,S)	(5,3,1,0)

図4.4に固定生成条件における各世代の手順を示す。基本的には、前章の実験手順と同様であるが、世代の最後にリンク張替え操作がある点異なる。

まず、リンク張替え操作を行う場合と行わない場合を比較したものを図4.5に示す。両ケースにおいて、一度非協調的な戦略が栄えた後、協調的な戦略が栄えるというシナリオは同じである。しかし、リンクを張り替える場合には、張り替えない場合に比べて収束が早いことがわかる。この過程においては、非協調的な戦略は協調的な戦略を求めてリンク張替えを行い、協調的な戦略は協調的な戦略同士でリンクし安定した集団を作ろうとする。ただし接続の嗜好性から、非協調的な戦略は、全リンクは自分の所有するリンクのみであるのに対し、協調的な戦略は被リンクも存在する。このようにして最終的には被リンクの分だけ利得を稼ぐ協調戦略が栄えることになる。

次に図4.6にその他の指標の推移例を示す。

パス長はほぼ単調減少である。この例においては、一度上昇しているが、これはリンク数の多いエージェントが死亡してしまったためである。クラスタ係数は一度落ち込んだあと、徐々に上昇し最終的には1に近づいている。このネットワークでは、リンク数が非常に大きい、少数のスーパーエージェント同士が互いにリンクしあい、その他のエージェントはそれらのスーパーエージェントにリンクする構成となっている。そのためクラスタ係数が1に近い値となっている。

ある時点においてエージェントがそれ以前にリンクしたことがあるエージェント数の平均値を移動距離と定義すると、この例ではクラスタ係数が1に漸近すると同時に30程度に収束

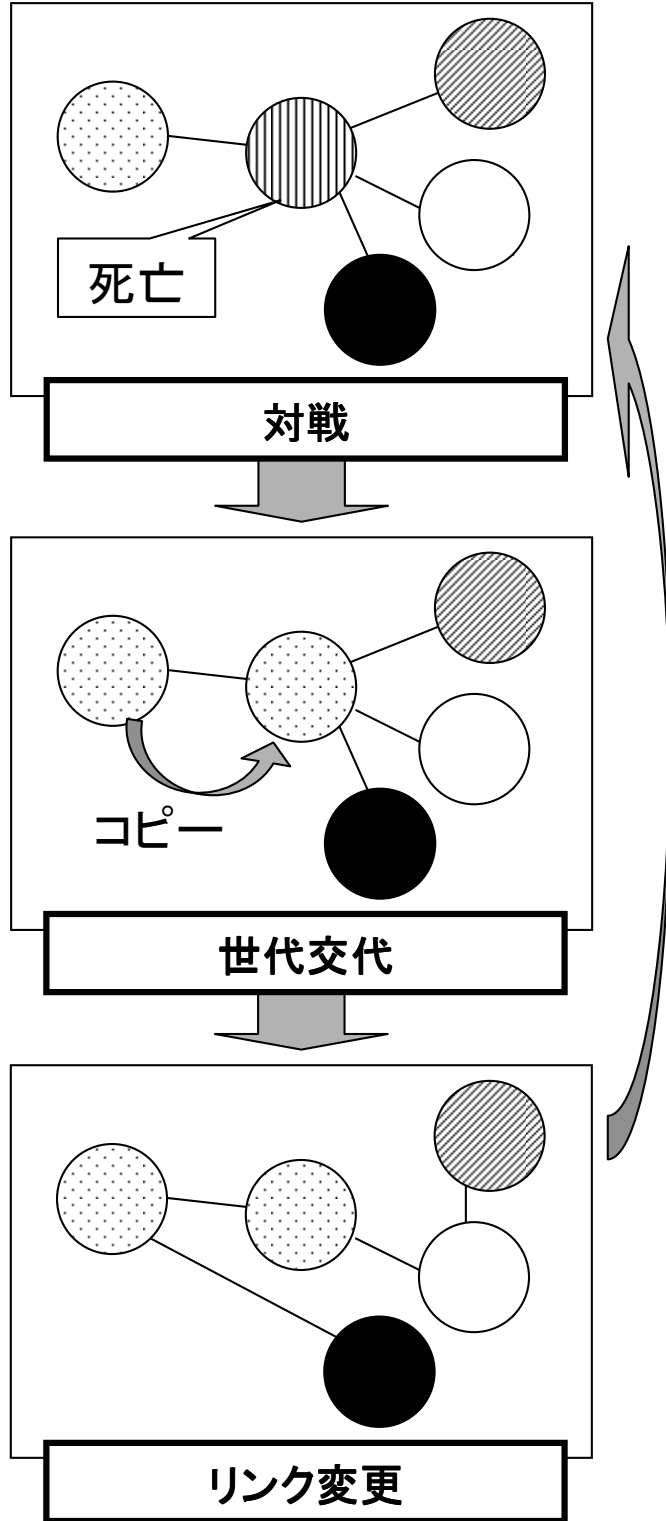


図 4.4. 各世代における手順

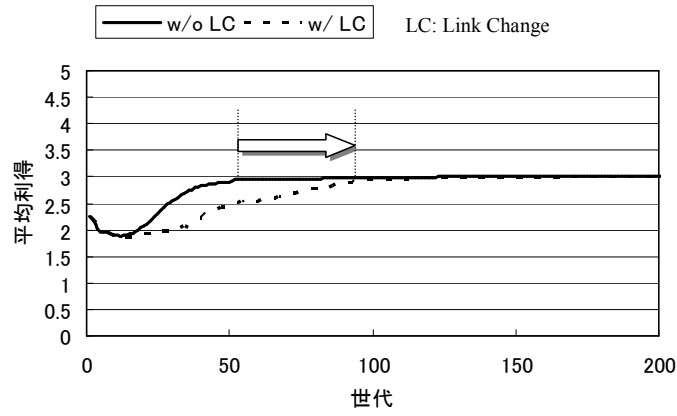


図 4.5. 平均利得の推移

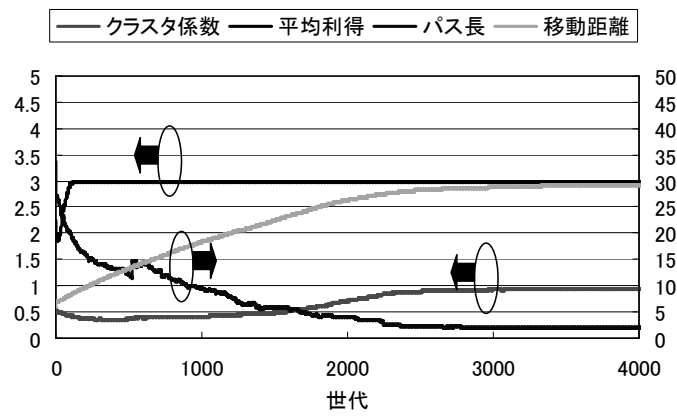


図 4.6. 進化ダイナミクス例 (noise=0, mutation=0)

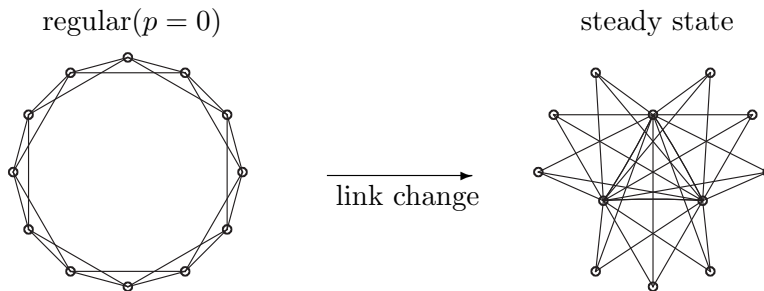


図 4.7. 定常状態におけるネットワーク

した。移動距離が収束する条件は、 N^{add} の選択にリンク数を指標とするエージェントが多いことである。初期においては、当然自エージェントの得る利得を高めるような選択方法が生き残る。しかし、その後はリンク数が多いエージェント＝協調的なエージェントとなるため、利得とリンク数は指標としてほぼ等価となり、リンク数を指標とするエージェントも生き残ることが可能となる。今回のシミュレーション条件では、あえてリンク数を指標とすべき特別な理由は存在しないため、進化の過程において最初に多数を占めた指標が支配的になると考えられる。それぞれの指標の特徴として、リンク数の順位は安定傾向があるが、利得は隣接エージェントに応じて変動があり比較的不安定である。そのため、利得を指標とするエージェントが多数を占める場合には、移動距離はほぼ単調に増加する。また、 N^{del} に関しては、自分の利得を最も小さくするエージェントを選択し切断する操作が生き残ることがわかった。

4.4 まとめ

本章では、世代交代において死亡エージェントと生成エージェントの関係性の差異を意味する自由生成条件、固定生成条件それぞれにおいてネットワーク上での囚人のジレンマゲームを行いネットワークの変化を観察した。

その結果、自由生成条件においてはスケールフリー性が強いネットワークがあらわれた。また、固定生成条件においては、リンク維持コストがない理想的な条件下では、少数のエージェントにその他全てのエージェントが接続する特殊なネットワークが生じることがわかった。

両条件において共通する点として、スモールワールドネットワークから出発したにも関わらず、リンクの変動が存在する場合には最終的にスケールフリー性が観察されることがわかった。

第 5 章

空間構造の学習への影響

5.1 はじめに

生物には、進化と学習という機構が備わっている。進化と学習は環境に適応するためのものであり、それにより自然選択のなかで種として生き残る機会を増加させてきたと考えられる。この機構については 100 年以上前から研究されてきた。進化と学習に対して影響を及ぼす要素は多い。また進化と学習が相互に影響を及ぼし合うとも考えられる。

人間の基本的な能力の一つである学習については、心理学、脳科学、計算機科学等様々な分野で研究が進められた。特に近年の計算機科学分野では、機械に学習させるアルゴリズムが多く提案されその進歩が著しい。

一方、進化は非常に長い年月を必要とする現象であるため、定性的な議論が多かった。しかし、計算機の計算能力が近年急激に高まってきたこともあり、計算機上でシミュレーションとして進化を扱うことが容易になってきた。そのため学習と進化の両方を扱ったものも増えてきており、例えばボールドウィン効果を扱った研究 [Suzuki 04] など近年研究が進んできている。

ゲームの学習という点においては、強化学習の発展と共に研究が進められるようになった。例えば確率的なゲームをプレイする学習エージェント [Hingston 04] や、マルチエージェントシステムにおける学習エージェントが互いの学習が原因で発生する不安定性が示されたもの [Sandholm 96] などである。

しかし、これまで進化と学習を行うゲームをプレイするエージェントについては研究は十分ではない。また、何らかの外部から学習エージェントに与えられる影響を、学習パラメータから調べるようなものは存在していない。そこで、本章では、空間構造が集団において進化と学習に対し、どのような影響を与えるかについて調べている。まず、進化と学習を行うエージェントモデルを提案する。このモデルにより学習に対する影響を観察することが可能になる。次に、空間構造上における囚人のジレンマゲームの計算機シミュレーション実験を行い、その結果について考察する。

5.2 学習エージェントモデル

提案する学習モデルは、強化学習を用いて、ゲームを行う際に利得が得られやすいと考える手番を次から選択できるように学習していく。ここでは、まず強化学習について簡単に説明し、その機構を利用したエージェントモデルを提案する。

5.2.1 強化学習

人間が自然に行っている学習という行為は機械にはかなり困難なものである。古くから学習についての研究は続けられてきたが、近年のコンピュータの発展に伴い、コンピュータに学習をさせる機械学習という研究領域が出現した。機械学習は、教師あり学習、教師なし学習、強化学習と大別される。教師あり学習は知識を持った外部の教師が提供する、例からの学習である。教師なし学習は例が与えられずに出力することが要求される。

強化学習は、ある環境内において、現在の状態を認識し、行動を選択する過程において、試行錯誤しながら最適な行動を学習する振る舞いを指す。教師あり学習とは異なり、行動の正誤を明示的にフィードバックする教師は存在せず、そのかわりに報酬が存在する。学習する主体はその報酬を手がかりに、累積する報酬を最大化するように最適な行動を学習する。

5.2.2 学習エージェント

提案する学習エージェントモデルを説明する。エージェントは進化と学習を行う。エージェントはそれぞれ、生得的に得る初期設定戦略と、それ以降学習を行うためのパラメータがコーディングされている遺伝子を持つ。エージェントは、ゲームを行いつつ学習を行い、以降のゲームではその学習結果をふまえて選択された手番を出力することになる。進化の過程においては、生き残りやすい初期戦略及び学習パラメータが観察されることになる。

繰り返しゲームを行う際、プレイヤーは過去一回のゲームの記憶を保持し、前回の自分と相手の手番の組み合わせによって状態を認識する。プレイヤーは全ての状態・手番の組み合わせについて価値関数を保持し、ある状態 s において行う手番 a を価値関数 $Q(s, a)$ によって評価する。ただし、 s は考えうる状態を表し、この場合は (自分の手番, 相手の手番) = (CC, CD, DC, DD) と初期状態の 5 通りに限定している。基本的には a_{t+1} は式 (5.1) のように価値関数によって選択される。例えば、 $Q(CC, C) \geq Q(CC, D)$ という条件下ならば、状態 CC において次の手番は C が選択される。

$$\text{next move} = \begin{cases} C, & Q(s, C) \geq Q(s, D) \\ D, & Q(s, C) < Q(s, D) \end{cases} \quad (5.1)$$

選択した手番に対しては、得られた利得を報酬として価値関数が更新される。更新のルールは式 (5.2) に示されるとおりである。 $Q(s_t, a_t)$ は t における価値関数を表し、 $Q(s_{t+1}, a_{t+1})$ は更新された後の価値関数を表す。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \quad (5.2)$$

ただし、 α は学習率、 r_{t+1} は手番 a_t に対する報酬、 γ は割引率である。学習モデルとしては本研究では sarsa[Sutton 98] を採用した。遺伝子には、これら強化学習のためのパラメータと価値関数の初期値が含まれる。

ここで、 $Q(s, a)$ が大きいほど a は優れていると評価された手番であり、プレイヤーは優れた手番を選択すべきである。ただし、一般にプレイヤーは行動探索的な機構も備えているべきであり、本モデルでは次の手番 a_{t+1} はある確率 ϵ によってランダムに選択する ϵ グリディ法によって選択する。例えば、発生させた乱数の値が ϵ 未満であれば C/D はランダムに選択され、 ϵ 以上の場合には、式 (5.1) のように価値関数によって選択される。

以上をまとめたエージェントの手番選択及び学習プロセスを図 5.1 に示す。

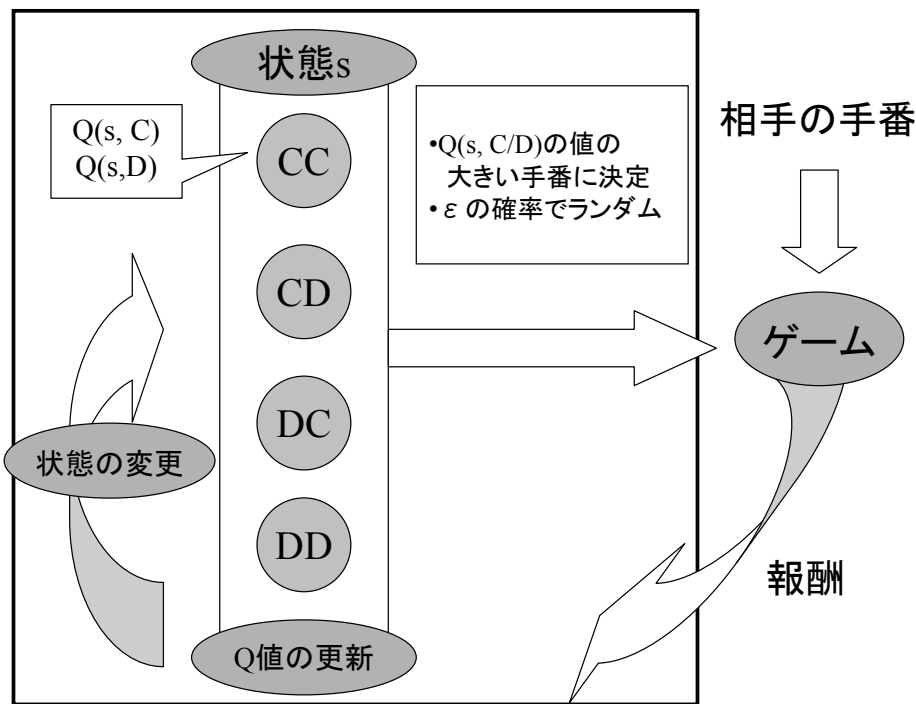


図 5.1. エージェントモデル

更新プロセスは選択毎に行われ、これが繰り返しゲームにおける学習に相当する。戦略は価値関数によって決定され、学習によってその価値関数が増加するため戦略が変化することになる。例えば、遺伝子コード化された初期戦略では TFT だったものが、学習の結果 Pavlov になる、ことが学習による戦略の変化である。

学習プロセスの特性は、強化学習パラメータに依存する。 α は価値関数の変化の速度、 γ は将来得られると考える報酬をどの程度重視するか、 ϵ は探索の度合いである。

以上のモデルは、外部から観察すれば一次のメタ戦略と同等であり、表 5.1 (再掲) のよう

な戦略表現となる。

表 5.1. メタ戦略例 (再掲)

前回の 自分の手	前回の 相手の手	戦略例			
		All C	All D	TFT	Pavlov
C	C	C	D	C	C
C	D	C	D	D	D
D	C	C	D	C	D
D	D	C	D	D	C

5.3 シミュレーション実験

5.3.1 シミュレーション内容

空間構造に依存する集団特性，特に学習に関する傾向を調べた。

以下のように空間構造ごとに少し異なる手順で実験を行った。

ただし，エージェントが行うゲームに関してランダムなノイズを仮定している。ノイズとは，エージェントが手番を選択した後に，エージェントの意図に関わらずある一定の確率でランダムに手番が逆転することを指す。

プールケース

ここでは，空間構造の存在しないケースを仮定している。エージェントは互いに接続されていないため，集団に属する他のエージェントの全てと対戦する可能性がある。

まず， $size$ 個のエージェントを生成する必要がある。

1. サイズ $size$ の集団を生成する (空間構造はないので配置は不要)
2. 以下のプロセスを世代数 gen 回繰り返す
 - i ランダムな順番でエージェントを選択し，(link) 数回，他のエージェントを選択し対戦する。繰り返しゲーム数 $repetition$ 回のゲームから構成されるスーパーゲームを行う
 - ii 全てのスーパーゲーム終了後，その世代において稼いだ利得に応じてランダムにジェネレーションギャップ $gap\%$ のエージェントを殺す
 - iii 死亡したエージェントを補充するために，サイズ 2 のトーナメント選択，2 点交差，突然変異を用いて新たなエージェント集団を再構成する

ここで，ステップ 2 - i において，対戦回数を制限しているのは，他のケースでは空間構造に含まれる (link) 数で対戦回数が制限されるため，他と同等のゲーム数にするためである。

ネットワークケース

空間構造の存在するケースである。

ここで空間構造は、現実世界でも頻繁に観察されるスモールワールドネットワークとスケールフリーネットワークの 2 モデルを取り上げることとする。

1. サイズ $size$ の集団を生成し、作成した空間構造のエージェント部分に配置する
2. 以下のプロセスを世代数 gen 回繰り返す
 - i 空間構造中のリンクをランダムに選択し、そのリンクで直接接続されたエージェント同士で繰り返しゲーム数 $repetition$ 回のゲームから構成されるスーパーゲームを行う
 - ii 全てのスーパーゲーム終了後、その世代において稼いだ利得に応じてランダムにジェネレーションギャップ $gap\%$ のエージェントを殺す
 - iii 死亡した空白エージェントに、そのエージェントから直接接続されたエージェントのうち、その世代で最も利得を稼いだエージェントをコピーする。ただし、コピーの際は、突然変異も考慮する。

表 5.2 にシミュレーション諸元を示す。

上述の強化学習パラメータは三つ存在する (α, γ, ϵ)。しかし、三つを自由に変動可能な状態で観察すると、自由度が高すぎるため結果から定性的な説明をすることが難しい。そのため、三つのうち ϵ を固定し、 α, γ の振る舞いを観察することにした。

シミュレーションは、ノイズ確率および空間構造を初期設定として持つ。空間構造に関するもののうち、ネットワークケースにおいては、ランダムネスパラメータも初期値として設定される。

表 5.2. シミュレーション緒元

パラメータ	値
集団サイズ (size)	400
エージェントの持つ平均リンク数 (link)	6
対戦あたり繰り返しゲーム数 (repetition)	100
ジェネレーションギャップ (gap)	20%
突然変異率 (mutation)	0.02%
世代数 (gen)	2000
囚人のジレンマ利得パラメータ (T,R,P,S)	(5,3,1,0)
強化学習パラメータ ϵ	0.1

5.3.2 シミュレーション結果

ノイズの影響

シミュレーション結果を以下に示す。空間構造のうち、スモールワールドネットワーク、スケールフリーネットワークのケースでは、スモールワールドとしての性質が見られるランダムネスの値 ($p = 0.1$) を設定している。

まずはじめに平均利得に注目する。これは集団のうちどの程度のエージェントが協力しているかを示す指標となる。指標は3に近ければ、ゲームの利得行列の設定をみればわかるように、それは多くの協力者が存在することを意味する。

図 5.2 に空間構造ごとに平均利得を示す。縦軸は平均利得、横軸はノイズとなっている。

プールケースにおいては、学習ありと学習なしの場合の違いは明らかである。学習ありの場合には、ノイズ確率の増減に関わらず、ほぼ一定の値をとり続けている。一方、学習なしの場合には、ノイズが存在しないときにはほぼ全てのエージェントが協力していることを意味する3を示しているが、ノイズ確率が増加するにつれ、値は単調減少していく。学習なし、ありの二つのデータはノイズ確率がおよそ 1.0×10^{-2} の付近にて交差する。これらの結果から、プールケースではノイズ確率という条件が平均利得に大きく影響を与えることがわかる。

スモールワールドネットワークおよびスケールフリーネットワークのケースに関しては、学習なしの場合のデータは明らかに学習ありの場合のデータよりも大きな値を示している。両空間構造において、学習なしの場合には、ノイズ増加に対しては平均利得は単調減少を示す。学習ありの場合は、ノイズに対してほぼフラットである。

以上から、学習なしの場合の方が、学習ありの場合に比較して、限られた条件の場合を除いて平均利得は高いことがわかった。言い換えれば、学習ありの場合には、エージェントは過度に貪欲に利得向上に励むため協力行動が生まれにくいということである。一回限りゲームでの裏切り同士というナッシュ均衡を脱するための説明としてある、繰り返しゲームでは協力を繰り返すほうが互いにハッピーになれるという理論的な結果に反している。この理論からの逸脱は学習による不利益であると考えられる。

ここで、学習の効果をみるために ACR (Action Change Rate) という指標を定義する。エージェントは内部状態それぞれに価値関数の値で手番が決定されており、それらのセットが即ち戦略であるが、生成された状態の時の戦略と学習後の戦略ではどの程度変化しているかを示すものである。図 5.3 に示すように、例えば5ビットのうち、学習によって2ビットが変化していれば $ACR=0.4$ となる。これを集団すべてにとって集計する。

図 5.4 に ACR および強化学習パラメータ α と γ を示す。 α は価値関数の変化の速度、 γ は未来の報酬の重みを意味する。

全てのケースにおいて、ACR は空間構造に依存した値でほとんどフラットである。これは遺伝的な戦略で安定なものは存在せず、絶えず戦略の変化が必要であることを意味する。必ず一定程度は学習能力が必要であることを示唆する。

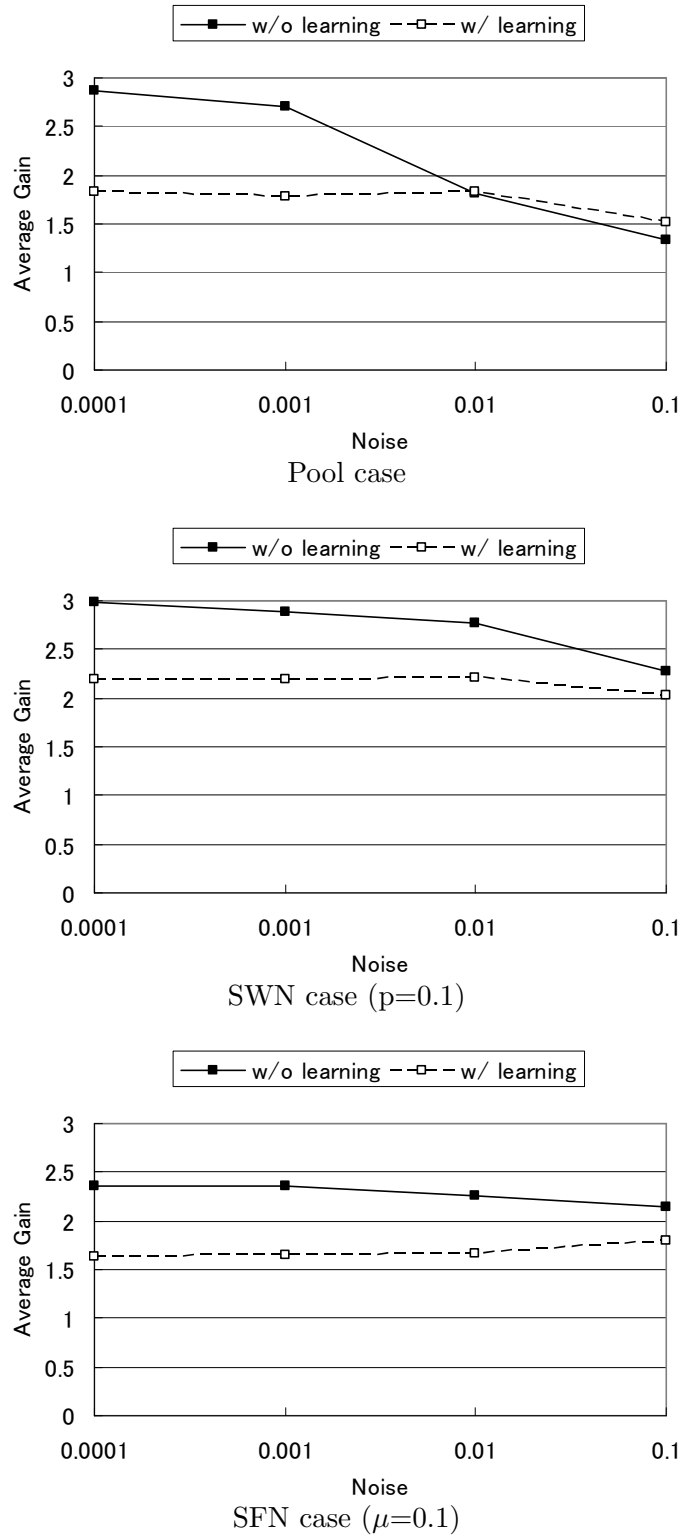


図 5.2. 平均利得

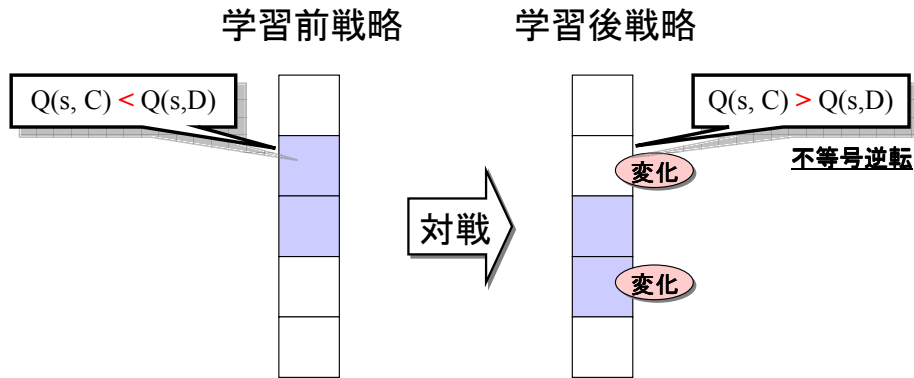


図 5.3. Action Change Rate 定義

エージェントの強化学習パラメータに関しては、 α と γ は、プールケースにおいては 0.2 から 0.5 の範囲である。一方、スモールワールドネットワークケースでは、 γ は α と比較して非常に大きな値をとっている。

スモールワールドネットワークケースでは、エージェントは将来受け取ると期待する利得を重要視し、現在の戦略を変化させる速度は遅めになっている。彼ら是对戦相手が同じ戦略をとってくる可能性が高いと期待し、お互いの関係性を重視する傾向がある。一方、プールケースでは、エージェントはその瞬間に得られる利得を重視し、自らの戦略の変化も早い。彼ら是对戦では全てのエージェントと対戦する可能性が存在するために、関係性を築くことはないためである。

スケールフリーネットワークケースでは、プールケースとスモールワールドネットワークケースの間である。スケールフリーネットワークとスモールワールドネットワークでは、ネットワークポロジの固定という面では似ているが、多数のリンクを集めるハブの有無が異なる。スケールフリーネットワークケースでは、ハブのような中心的なエージェントも確率的に死亡する場合が存在するため、その場合には付近のエージェントにとっては急に異なる戦略をとるエージェントが存在することになる。この効果がリンクの数だけ増幅されることになり、関係性の継続性という意味では、完全にランダムなアクセスが可能なプールケースとほぼ固定的なスモールワールドネットワークの中間の性質を示すと考えられる。

ノイズに関して言えば、ノイズの発生確率が 0.1 を境に振る舞いが異なることがわかる。スモールワールドネットワーク、プールケースでは、 α は増加し、 γ は減少する。ノイズ発生確率が高いということは、ノイズによって相互の協力関係の維持に悪影響を与えやすいということである。そのため長期的な報酬を信頼しにくくなる。そのため、 α の増加、 γ の減少が観察される。これは上述の空間構造に依存する効果と同様であり、未来の予測の容易さといえるかもしれない。以上より、未来の予測の容易さ、ひいては長期的な信頼関係の影響として空間構造およびノイズの両方の効果が確認された。

ただし、スケールフリーネットワークケースでは α 、 γ ともノイズ発生確率に関わらずほぼ一定の値をとり続ける。平均利得を見るとわかるように、そもそもスケールフリーネットワー

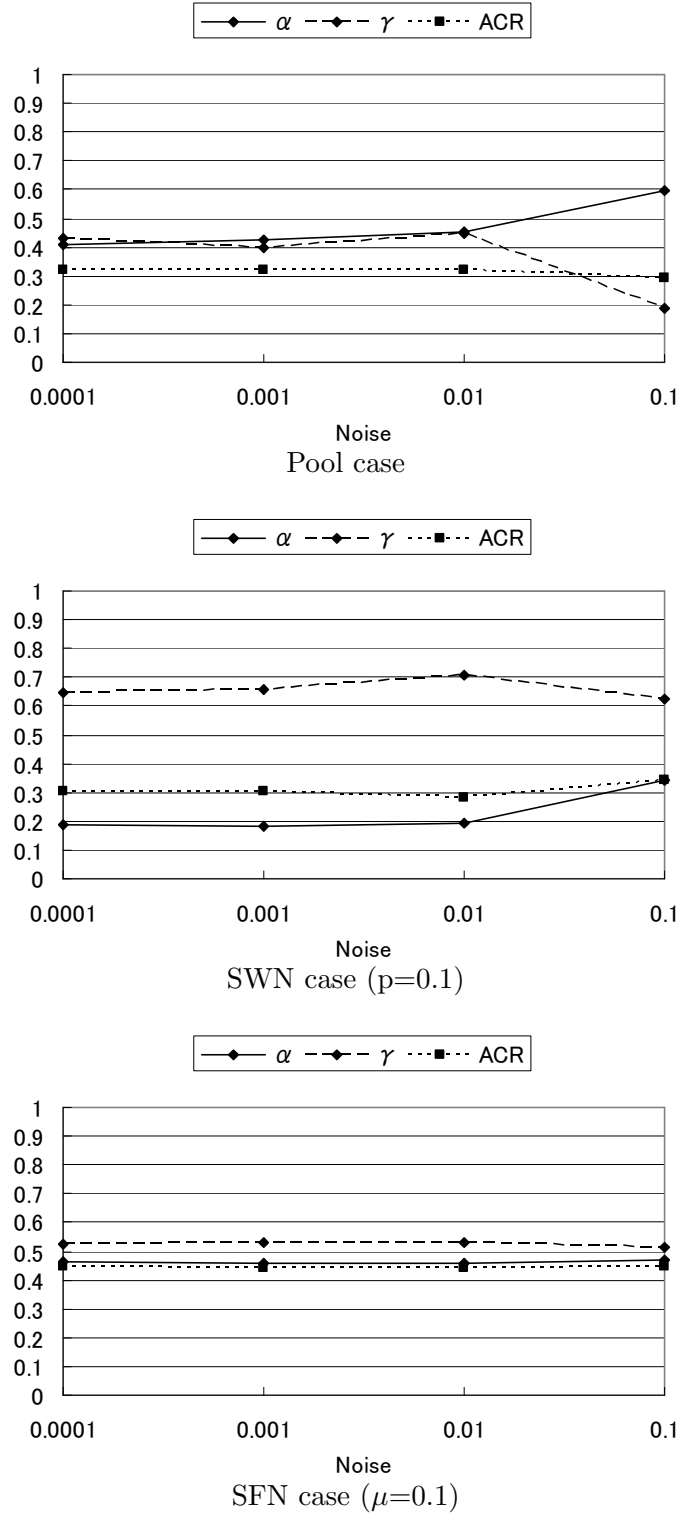


図 5.4. 強化学習パラメータ

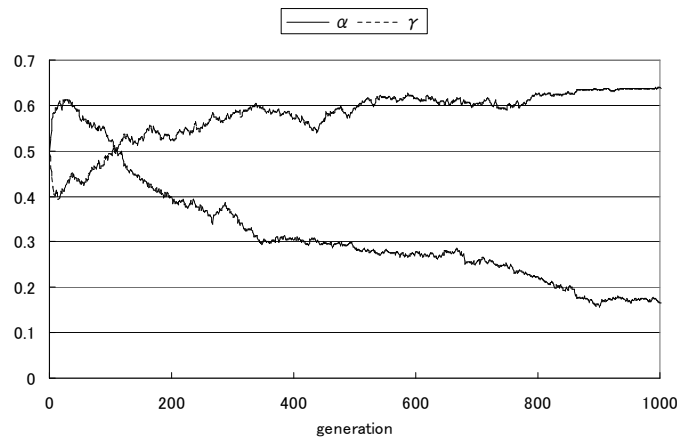


図 5.5. 強化学習パラメータの推移例

クは低い．ノイズの効果よりもハブの死亡による影響の方が高いためこのような結果になっていると考えられる．これをもってスケールフリーネットワークはノイズ耐性が強いということも可能であるが，むしろノイズ以外の影響が顕著なため見えにくいという表現が妥当であろう．

以上の結果から導かれる興味深い点は，未来の予測の容易さが学習プロセスに影響を与え，進化的には学習の仕方が遺伝的にコーディングされてしまう，という点である．信頼できるならば α は低く γ は高い状態で学習が進む．

また，更に進化によってどのように学習が進んだかを見るために一例を示す．5.5 に横軸を世代にとった場合の強化学習パラメータの推移例を示す．これを見るとわかるように，最初に一度 α の値が大きく γ の値が小さくなるような過渡状態を経た後，それぞれ逆転して α の値が小さく γ の値が大きくなる．これは先に見た，非協力的なエージェントが一度は増えるが，増えすぎると逆に非協力的なエージェント同士で生き残りにくくなり最終的に協力的なエージェントが出現するという現象と似ている．

図 5.6 は集団のエントロピーを示している．本シミュレーションでは，32 戦略が存在する．そのためエントロピーの最大値は 5 である．図においては，“L” は学習を表し，“gene” はエージェント生成時のもの，“learned” は学習後のものをあらわす．

全てのケースにおいて，学習後のエントロピーは非常に大きな値をとっている．学習による戦略の変化が存在していることをこれまでにわかったが，戦略は学習によって他のエージェントと同じでなく差異が存在することがわかる．遺伝子のエントロピーの低さは，学習によって戦略は異なるものの新たに生成されるエージェントも遺伝子レベルでは似ているということである．

ランダムネスの影響

ここでは，ノイズの発生確率を 0.01 に固定している．以降の図は横軸をランダムネスにとっている．

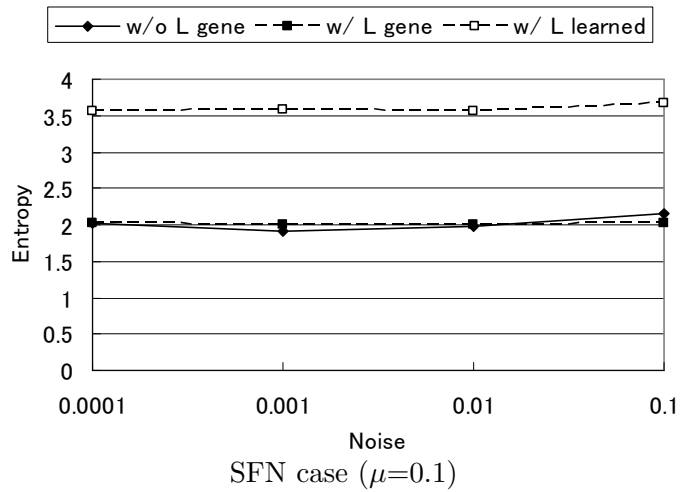
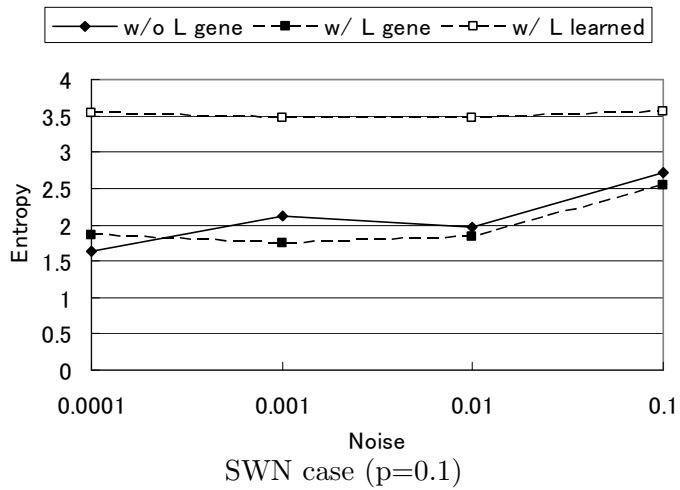
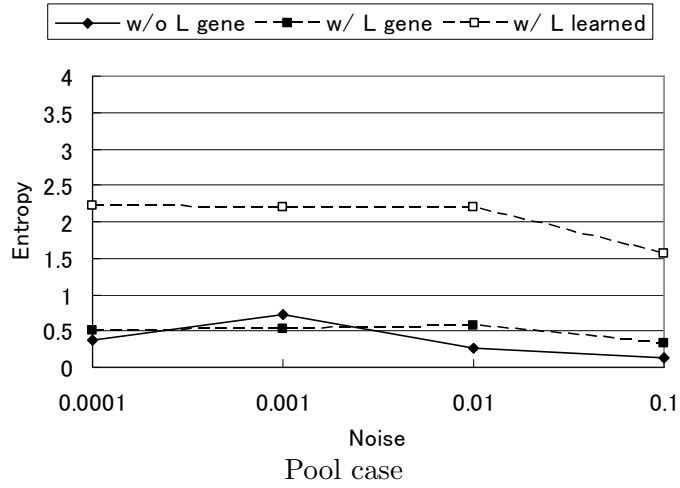


図 5.6. エントロピー

図 5.7 に平均利得を示す．ノイズを変化させた場合と同様に，学習なしの場合が平均利得が高いことに変わりはない．両空間構造において，ランダムネスの影響はそれほど大きくなく，一定の値付近を動かさないことがわかった．

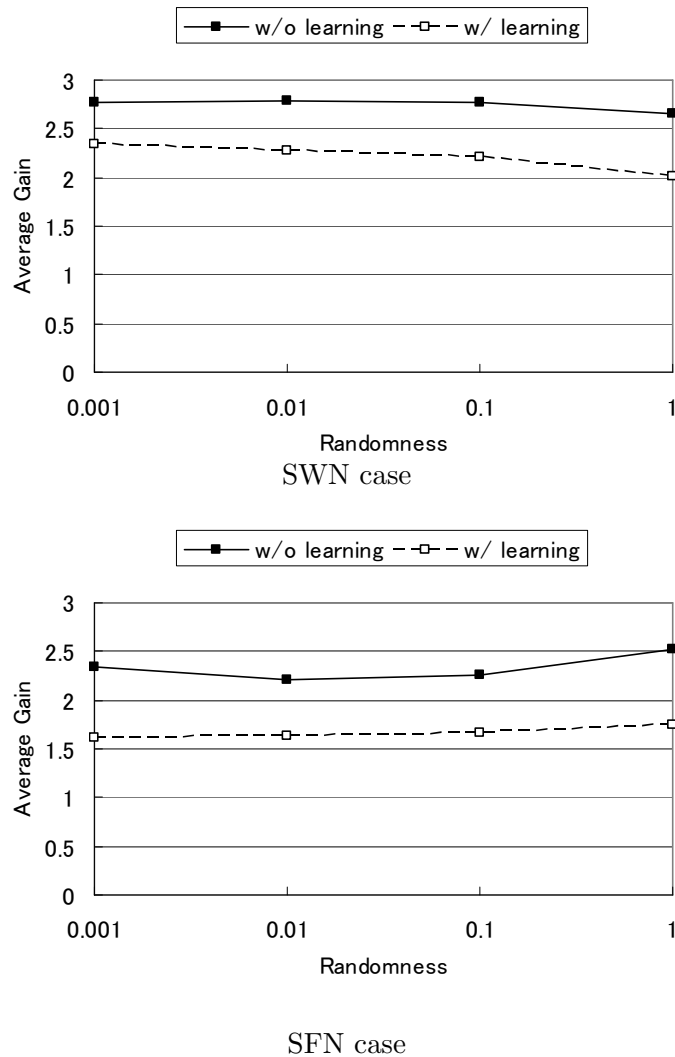


図 5.7. 平均利得

図 5.8 に強化学習パラメータを示す．スモールワールドネットワーク，スケールフリーネットワークともランダムネスが 0.1 以下の場合には，ほぼ一定の値をとっている．また，その値はスモールワールドネットワークでは $\gamma > \alpha$ であり，スケールフリーネットワークでは $\alpha \approx \gamma$ である．ランダムネスの値が 1 になると，振る舞いが異なり，スモールワールドネットワークでは， α は増加し γ は減少する．一方スケールフリーネットワークでは，逆に α は減少し γ は増加する．

図 5.9 は，スモールワールドネットワークとスケールフリーネットワークのケースの集団の

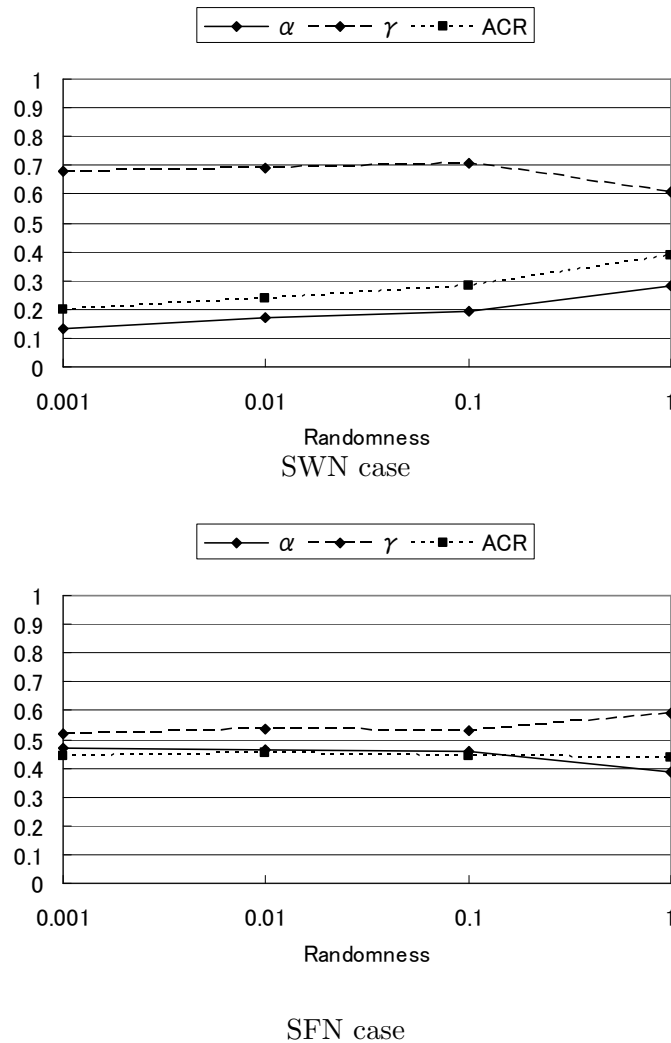


図 5.8. 強化学習パラメータ

エントロピーをあらわしている。ランダムネスが増加すると、スモールワールドネットワークケースでは、全てのエントロピーは増加しているが、スケールフリーネットワークでは、完全にランダムになるときエントロピーが急激に低下しプールケースと同様の状態になっている。

スモールワールドネットワークの場合には、ランダムネスの増加、つまりクラスタ係数の減少に従いエントロピーは上昇する。クラスタ係数が小さくなるということは、安定的でない相手との接触機会の増加を意味するので戦略も多様化しエントロピーが増加する。しかし、スケールフリーネットワークの場合には、ランダムネスが小さい状態においても、ハブの死亡の影響が大きいため必ずしも安定にならずエントロピーも比較的高い。

完全にランダムなとき、パス長、クラスタ係数とも小さくなる。このときスモールワールドネットワークはさらにエントロピーは上昇する。一方、スケールフリーネットワークは遺伝子のエントロピーは減少する。

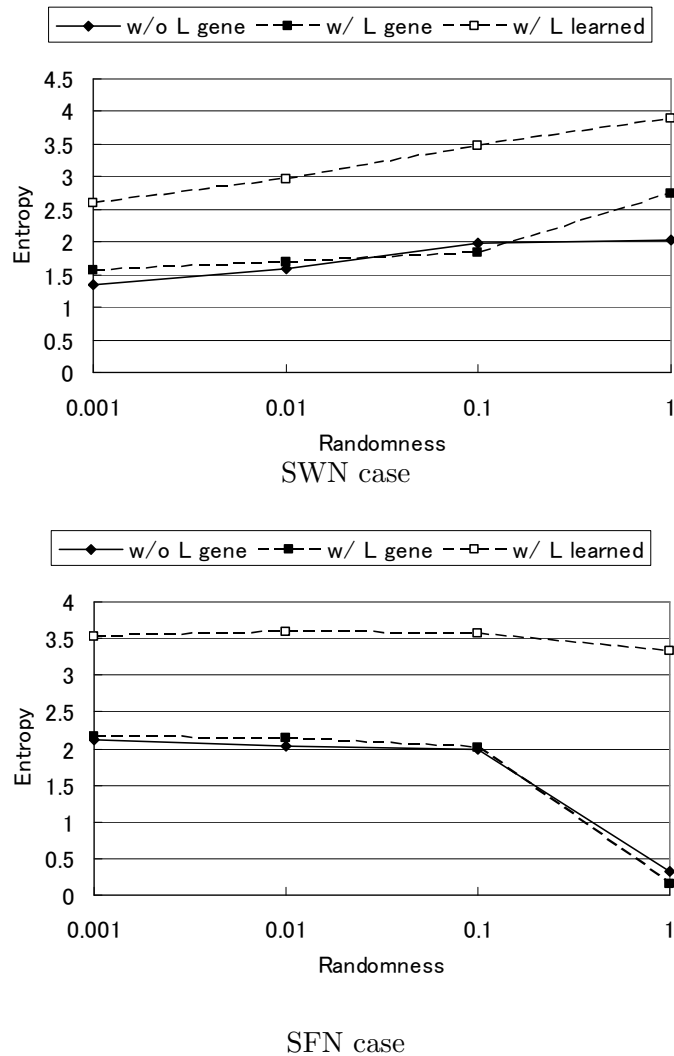


図 5.9. エントロピー

5.4 まとめ

本章では、学習を行うエージェントによる空間構造上での囚人のジレンマゲームを見てきた。まず、進化と学習を行うエージェントモデルを提案し、計算機シミュレーションの結果、空間構造がエージェントの学習に与える影響を確認した。関係性の強い空間構造では長期的な報酬を期待し、逆に関係性の弱い空間構造では刹那的な選択を行うことがわかった。

これはこれまで社会学で行われてきた研究結果を定量的に説明したものとなる。社会学では、社会における流動性と信頼の関係についてはこれまでも考察されてきた。基本的には、社会の流動性の増大は、安定した関係に根ざす相互信頼を低下させる [山岸 98] と言われている。ここでいう信頼は、強化学習パラメータになおせば長期的な報酬を期待することとほぼ等し

く、空間構造と学習の関係を示すことにより、この言説をサポートすることができる。

第6章

集団の制御可能性

6.1 はじめに

本章では、集団の振る舞いを変化させ望ましい性質が発現するような制御が可能であるか検討する。

基本的に、創発的な現象を制御することは困難である。そもそも創発という現象は、全体の振る舞いが個々の要素からは予測困難なものを指すので、定義からして制御が困難であることは明白である。

それを如実に示す例が経済である。どの国の政府も自国の経済を制御しようと経済政策を打ち出し試みるが、必ずしもうまくいくとは限らない。ここで既にある程度手法が整理されている経済分野を参照しておく、経済政策においては大きくはマクロ的な手法とミクロ的な手法に分かれる。マクロ的な手法はマクロ経済の知見に基づく。例えば、財政支出、金利引下げなどマクロ的に関連性のある指標のいずれかを人為的に操作することにより、制御対象となる指標の動きを引き出そうとするものである。一方、ミクロ的な手法は、規制や税制などを操作することにより、要素となる個人や企業の行動の変化となるような政策を行うことにより、ミクロな制御対象である要素の行動の集積として全体として望ましい方向へ誘導するというものである。

人間を規制・制御するものとして、レッシグは法・規範・市場・アーキテクチャの4つであると整理した [Lessig 99]。ここでアーキテクチャとはその世界が従う物理法則のことである。このフレームワークに沿って制御方法について整理してみる。

法による制御を考えると、裏切り行為を選択しないよう、裏切り行為を選択した場合には法による罰を与えるなどとして利得構造を変えることになる。この類の研究としては、モニタリング & サンクションにより、行動メカニズムに影響を与える試みがある [Deguchi 98]。監視者はある確率でゲームを監視しており、エージェントがもし裏切り行為を選択した場合には、その行為に対し罰が与えられるというものである。ただし、この場合、ゲームを監視する中央集権的な存在を仮定する必要がある。

規範では、裏切り行為自体を選択肢として意識に上らないようにする。または、法による罰は

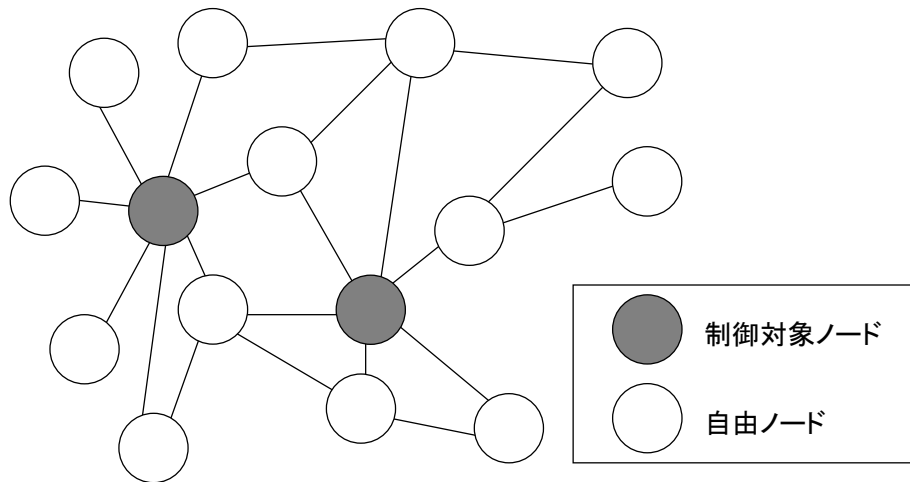


図 6.1. 提案制御方法

与えられないまでも、コミュニティにおいて居辛くなるなどの結果の予測からくる、道徳的な意味において意思決定主体者の認識する利得構造の変化を促す。この例としては、協力行動や制裁行動を利得として認識するエージェントを仮定し、規範による影響を考察した研究 [松本 06] がある。

以上の二つは、ミクロな制御手法であるが、もしマクロな制御対象があるならば同様の効果を期待することができる。

市場とは経済的なインセンティブのことであるが、ここでは囚人のジレンマゲームを発生させないようなアプローチを意味することになる。そもそもこの種のジレンマが発生する以前に予防するというのは良い発想かもしれないが、なかなか難しいことが予想される。また、本研究のスコープから外れるためここでは除外する。

最後にアーキテクチャによる制御は、物理的な制約により裏切をとれないようにすることを指す。これについては、囚人のジレンマ構造をとる具体的な問題に依存するため、なかなか一般的な物理的な制約というの想像しづらい。しかし、そのような直接的な影響ではなく、間接的に影響を及ぼすこともできるかもしれない。それは集団の中での対戦の組み合わせを制御することが一つの例であり、本章の提案方法はこのカテゴリーに属する。

基本的には、何らかの要因でゲームのルールを変化させることによって制御するアプローチが一般的であるが、ゲームをする環境の緩やかな制御による望む方向への誘導を今回試みる。本研究では、対戦の組み合わせは空間構造に強く影響されていると考えこれまで議論を進めてきた。そこで、本章では空間構造を利用した制御方法を提案し、その効果について検討する。

6.2 提案制御方法

ここでは、空間構造を利用した制御方法を提案する。

提案方法は、ネットワークの位置的な差異を利用する手法である。ネットワーク上である種

の特別な位置にあるエージェントに対し、集団に影響を与えるべく何らかの行為をとらせるといふものである。そのエージェントの選択は、最もシンプルな場合にはエージェントが持っているリンクの数で選択することが考えられる。図 6.1 に提案する制御手法を示す。ここでは、保有するリンク数が多い上位エージェントを利用し、それを制御対象とした。図においてグレーのエージェントは保有するリンク数が多く、他エージェントへの影響が大きいのではないかと推測される。そこで、それらを制御対象エージェントとし、制御対象エージェントは戦略を TFT に固定し、進化・学習は行わないこととした。

この仮定は、例えばネットワーク型のサービスを想定したときに、主催者側がある程度注目を集めるポジションを確保でき、そのポジションにおける振る舞いを選択が可能であるというような状況を想定している。

6.3 シミュレーション実験

シミュレーションは、制御対象エージェントを除いてこれまでと同様に行う。制御対象エージェントは、リンク数が多い上位から順番に選択するとし、制御されるエージェント数をシミュレーションパラメータとしている。表 6.1 にシミュレーション緒元を示す。

表 6.1. シミュレーション緒元

パラメータ	値
集団サイズ	400
エージェントの持つ平均リンク数	6
対戦あたり繰り返しゲーム数	100
ジェネレーションギャップ	20%
突然変異率	0.02%
世代数	2000
囚人のジレンマ利得パラメータ (T,R,P,S)	(5,3,1,0)

図 6.2 にシミュレーション結果の平均利得を示す。横軸は制御対象エージェントの数である。

順にみていくと、まずプールケースは提案手法が比較的有効である空間構造であるといえる。制御エージェント数が増加すると、学習なし・あり共に平均利得は上昇する。その後、それぞれ一定の値に収束する。そのとき学習なしの場合の方が利得は高い。注意すべき点は、そもそも他のケースに比較して制御対象エージェントが存在しないときの平均利得が非常に低いことである。そのため、制御対象エージェントが投入されることによって効果が出やすいという一面はあるだろう。プールケースでは、空間構造がないため集団内のどのエージェントともアクセスする可能性があるため伝播が促進される。そのため通常であれば、非常に非協力的なエージェントが集まる集団となりやすいわけである。しかし、その中に、相手の戦略に関わら

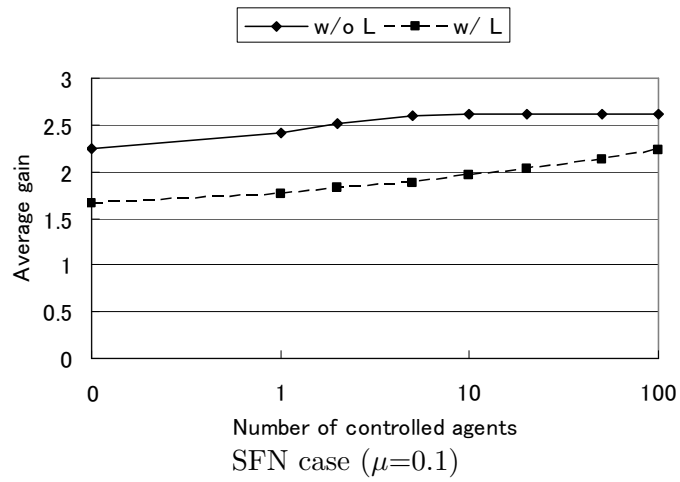
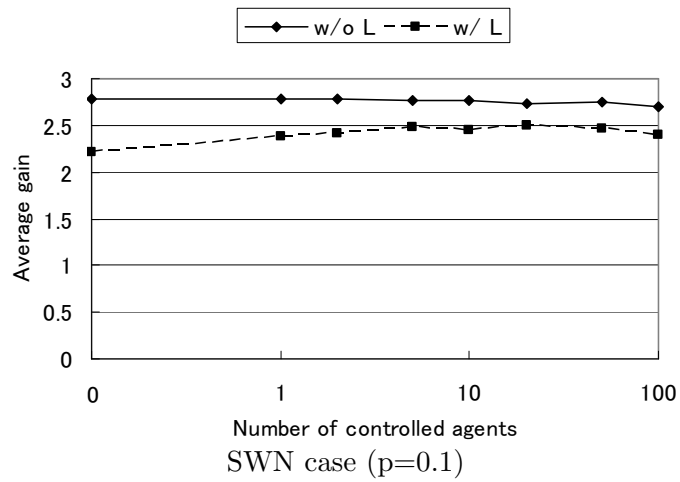
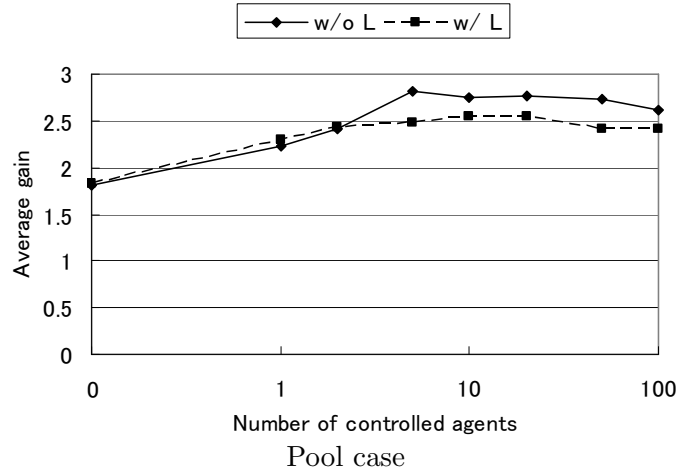


図 6.2. 平均利得

ず自分の戦略を固定する制御対象エージェントが存在すると、その制御対象エージェントの戦略に影響されるエージェントが出現しやすくなり、さらにその影響を受けたエージェントに影響を受けるエージェントが出現し、というふうに、また影響伝播が大きいためである。

スモールワールドネットワークケースでは、制御エージェントの数に関わらず、平均利得はほぼ同じ値をとっており、この提案手法の効果はあまり見られない。これはスモールワールドネットワークという空間構造は、エージェントが持つリンク数にあまり偏りがいないため、伝播の中心となるような位置が存在しないからではないかと考えられる。また、そもそも制御対象エージェントが存在しない場合にも、比較的高い平均利得であるため効果が見えにくいということもあるかもしれない。

スケールフリーネットワークケースでは、学習なしの場合には数エージェントでほぼ飽和状態となり、それ以降の向上はあまりない。一方、学習ありの場合には制御エージェント数が100程度まで継続して向上が見られる。スケールフリーネットワークは、エージェントが持つリンク数に偏りがあるため、空間構造上における位置の違いはスモールワールドネットワークに比べ大きい。その効果が現れているのだと考えられる。

次に、図 6.3 は強化学習パラメータを、図 6.4 は戦略のエントロピーを示す。

プールケースにおいては、制御エージェントを一つ集団に投入する時点で既に大きな変化がある。制御エージェントの戦略は TFT であるため、非協力的なエージェントは罰を受けやすく、死ぬ可能性が高くなる。基本的に、協力的なエージェントは遠い将来の報酬を期待し、手番として協力を選択する。そのため、協力的なエージェントの出現とともに、将来の報酬への期待を示す γ の値が上昇し、学習速度を意味する α の値が減少している。つまり、アクセス自由な環境においては、ある固定の戦略が周りの戦略に影響を及ぼしやすいといえる。エントロピーを見ると、基本的にあまり変化がない。平均利得から非協力的な集団から協力的な集団への変遷があることはわかる。しかし、支配的な種が入れ替わっただけで種の分布はほぼ同じであることを図から読み取れる。また、学習後のエントロピーは 2 以上とある程度高いが、プールケースでは遺伝子レベルではエントロピーは非常に低いことが特徴として挙げられる。

スモールワールドネットワークケースでは、平均利得の場合にもそうであったが、ほとんど変化はない。そもそも制御対象エージェントがいない場合でも他のケースと比較して協力的な集団であったためである。それは α と γ の値にも現れており、プールケースで制御対象エージェントを投入したときのような値が、制御対象エージェントがいない時点で既に達成されている。エントロピーを見ると、制御対象エージェントの数の増加に対して単調に減少している。

スケールフリーネットワークケースでは、あまり学習パラメータへの影響はない一見して不思議な状況である。エントロピーを見ると、他のケースと異なり、制御対象エージェントの増加につれエントロピーが劇的に減少していることが、スケールフリーネットワークにおける一番の特徴として挙げられる。制御対象エージェントを集団に投入することによって、遺伝子レベルでは大きくエントロピーが減少しはじめ、10 エージェント程度で集団のほとんど全てのエージェントが一種類になってしまうことがわかる。つまり遺伝子型の集団内への伝播が非常

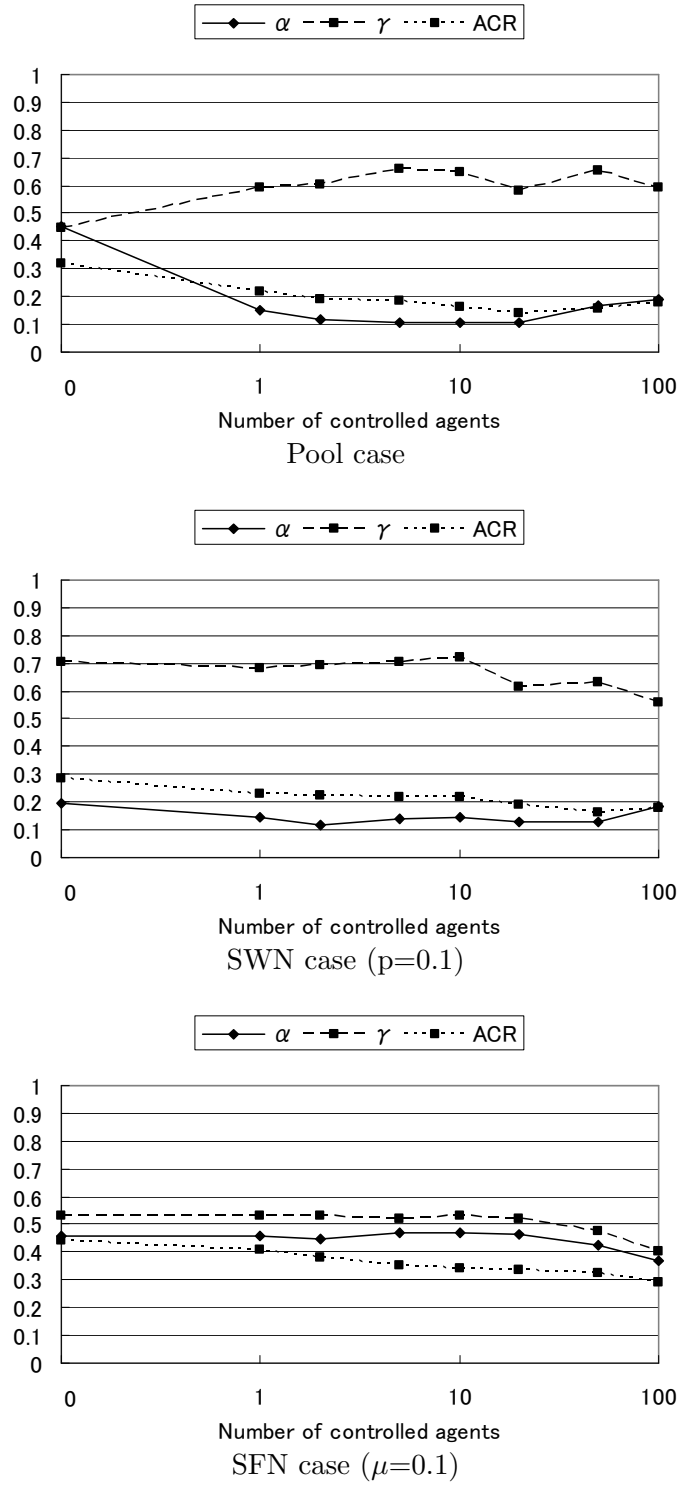


図 6.3. 強化学習パラメータ

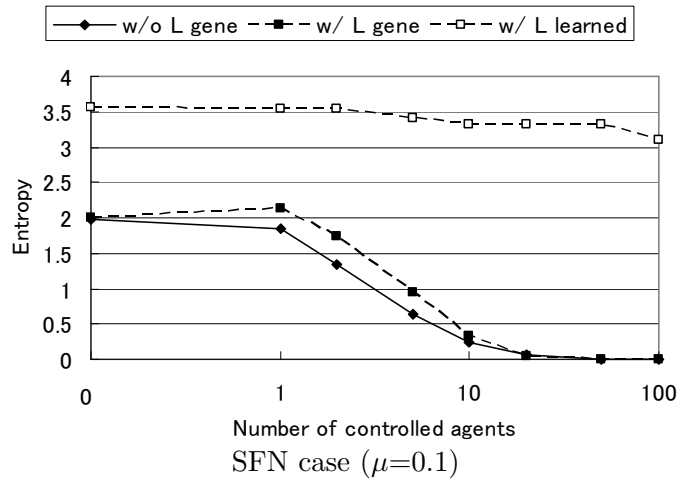
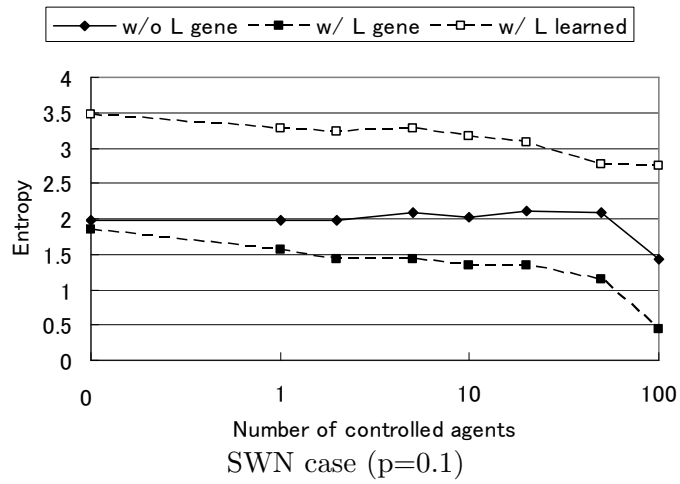
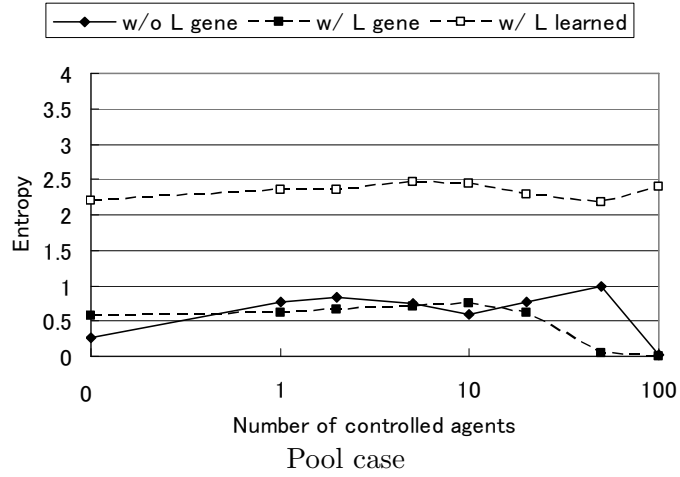


図 6.4. エントロピー

に強いということがわかる。学習パラメータの値自体は制御していない場合とあまり変化がなく、これは空間構造特有のものである。

6.4 空間構造と制御可能性

ここでは、以上の結果から得られた空間構造と集団の特性に関する知見を考察し、一般的に集団の制御を試みる場合にとり得る指針を検討する。

まず、制御主体にとって何が制御可能な対象であるか、によって指針は大きく変わる。

まず、制御の自由度が大きい場合、つまり空間構造を選択でき要素間のコミュニケーションを制限することができるような場合を考えてみる。その場合、エージェントの意思決定方法が不明ならば、スモールワールドネットワークが最も望ましい空間構造といえる。なぜなら、学習の有無に関わらず最も平均利得が高いからである。しかしながら、スモールワールドネットワークというネットワークを選択できるというその制御自体が最も大きな影響を与える。

次に、制御の自由度が小さい場合、すなわち空間構造自体は制御できない場合を考えてみる。これはある程度、要素の振る舞いが自由なシステムであれば一般的なケースであろう。その場合、ほとんどの集団においては要素のコミュニケーション自由度は比較的大きいだろう。その場合空間構造は、完全にアクセスが自由である状態は考えにくい、スケールフリーネットワークの性質を示すことが多いかもしれない。

ネットワークがスケールフリー構造を持つと仮定したとき、どのように集団の特性を向上させることができるか、という問いに対しては、少なくとも三つの方法が考えられる。

一つ目は、ノイズを減らす、という方法である。しかし、スケールフリーネットワークにおいてはあまり効果的ではない。ノイズにほとんど依存しないからである。

二つ目は、エージェントに対するアナウンスメントである。例えば対戦する相手の選択に対するリコメンデーションなどを行うことによって、ネットワークを操作するのに近い効果が得られるかもしれない。この方法はネットワークの変更を促すことにつながる。

三つ目は、提案手法である重要な位置への制御エージェントの埋め込みである。この方法には既に存在しているエージェントとの交渉という課題がある。しかし、成長するネットワークの場合、構造的に中心となるエージェントの性質を設定できる立場ならば有効であると言える。このとき、制御エージェントの戦略については一考の余地はある。スケールフリーネットワークケースでは、エントロピーが低いにもかかわらず平均利得が低かった理由は、TFT がノイズのある時には協力が続いているときにノイズにより裏切りに遷移する事があるからである。Pavlov 戦略の方がより良いかもしれない。

スケールフリーネットワークでは、戦略のエントロピーはランダムネスに強い影響を受けることがわかった。ランダムネスの増加による多様性の減少は別の問題を引き起こす可能性がある。一般的に、多様性がない場合には種として絶滅の危険性も高いと考えられるため、そのトレードオフも考慮する必要がある。

ノイズが全く存在しない場合には平均利得はほぼ 3 であることから、TFT がメジャーな戦

略であるにもかかわらず平均利得が低く押さえられているのはノイズが原因であるといえる。

ここで、前章と本章の結果をまとめると、表 6.2 のようになる。三つの空間構造と集団の傾向、学習傾向、提案手法の効果を示している。表を見るとわかるが基本的にはスモールワールドネットワークが協力的であり望ましい。しかし、第四章でも見たように、エージェントがネットワークの形成に関わるならば、スケールフリーネットワークの傾向があらわれるため比較的非協力的な傾向が出てくる。そこで本章で提案した手法をもし使えるならば、協力的な集団の方向へ誘導することができるかもしれない。

表 6.2. 空間構造と集団の性質

空間構造	集団の傾向	学習傾向	提案手法の効果
プール	非協力的	短期的学習	効果あり
スモールワールド	協力的	長期的学習	効果なし
スケールフリー	中間	短期的学習	効果あり

6.5 まとめ

本章では、集団の制御可能性について検討した。まず集団の制御方法を定性的に整理した。次に空間構造の特性を利用した方法を提案した。リンク数を多く持つエージェントを集中的に制御することにより、集団への影響を及ぼすというものである。提案方法について、実験によりその効果を観察した。その結果、集団の多様性とのトレードオフがあるものの、ある程度空間構造に偏りがある場合には効果があることが確認された。

第7章

他者意図の推定

7.1 はじめに

ゲーム理論では、完全に合理的な人間を仮定し、相手の手番を合理的に推測することにより、それに対して合理的な手番をとる。しかしながら、実際の人間は完全ではなく限定合理的であるため実際には正しくはない。

一方、これまでゲームをプレイするプレイヤーであるエージェントに関して、単純な刺激反射モデル、学習モデルを扱ってきた。基本的には、相手をほぼ環境と同等の存在として自分の置かれている状況を認識し、得られる利得が大きい戦略が生き残りやすかった。

しかし、現在の存在する我々人間を観察した場合、単にそれだけのもの以上の機能を使って意思決定をしていることに疑いはない。特に、実際に自分がゲーム的状况におかれたときに、相手のとるであろう手番を推測するという行動はごく自然に行われている。本論文ではこれまで、その相手の行動を考慮せず論を進めてきたが、ここで相手の意図を推測することによる効果を調べてみることにする。

相手の意図を推測するという行為に関し、これまで提出された一つの仮説に「心の理論」がある。心の理論という言葉はプレマックらによって初めて使われた [Premack 78]。自分の知っている状況ではなく、相手は相手が知っている状況からなにを考えるか、を知る能力とも言える。例えば有名な例として次のようなものがある。二人の登場人物 A,B が存在する場面において、A はボールをかごの中に入れる。次に A がその場から退出し、その間に B はボールを別の場所、例えば箱の中に移しておく。A が戻ってきたとき、A はボールを見つけるためにどこを探るか、と質問することによって心の理論を検査することができる。この問いに対して、かごの中、と答えれば他者の心を推測することができていることになる。一般に、4 歳程度になると、このように答えるようになるといわれている。

霊長類以外の種ではなかなか観察されにくいこのような機能が今現在ヒトに備わっていることに対する説明としてはマキャベリの知能仮説 [Byrne 88] がある。これは、集団生活による社会的関係の構築の必要性から、他者の意図を推測し政治的な振舞いが生まれたのではないかという仮説である。それをサポートする研究としては、集団サイズと脳の新皮質の関係を調べ

たもの [Dunbar 98] などがあり，この仮説に取り組む研究者も多い．近年の脳科学の発展により，他者の意図を感知する機能が脳に存在することを示すミラーニューロン [Rizzolatti 98] が発見され，相手の行動に反応して脳の一部が発火することがわかっている．

このように人間の脳に生得的に他者の意図を推測する機能が組み込まれている可能性は高い．また，その機能をもつ個体が進化的に有利であったから，そのような機能が残っていると考えられる．例えば，お互いが協力して何らかの目的を達成する場合には作業の効率化が期待できる．また，非協力的な状況においても協力の可能性を見出せるかもしれないからである．

最近では，計算機上でのモデルに関する研究も始まっている．協調的なタスクにおいて他者の意図を推定しつつ協力しながらタスクを解決するもの [長田 06] や，エージェント同士が干渉し合う独立な目的のタスクにおいて，他者の意図の推定の影響を調べたもの [高野 05] などがある．特に，何らかのタスクに対してエージェント同士が協力して解決するかたちの問題において他者の意図を推定することで協力の度合いが高まることは容易に推測され，マルチエージェントシステムなどの工学的な研究では特に利点が多いと考えられる．さらに，他者の意図を推測する場合には，再帰的レベルの問題もあり，“A think (B think (A think (...))))”のように推定における相手の心的内部モデルのレベルも話題とされることもあり，他人の思考の深さを探る試みなどもある [Ohtsubo 06] ．

少し異なるアプローチとしては，自己の感じる利得の中に他者の利得を組み込むというものもある [Fehr 99] ．例えば，ゲーム的状况において，自分の得られる利得を構成する関数の変数として他者の利得を含むことによって相手との協力が促進されるわけである．この場合，明示的に相手の意図を推定するのではなく，相手は利得が高い方が好ましく思うであろうという前提をもち，思考ではなく自己の中の価値観に組み込まれていることが仮定されていることになる．

また，協力に至るメカニズムとしては他に評判なども考えられるが，個体の識別はもちろんのこと様々な抽象的な概念を扱うため，高度に発達した機能を要求する．そのため，機能の発展上のステップとしては直接的に対峙した相手の意図を推定するものより後に出現するものと考えられる．

そこで，本章では非協力ゲームにおいて他者の意図を推測することによる効果がどのようなものかを調べることにする．他者の意図を推定することは，協力ゲームにおいても効果が高いと推測されるが，本研究ではここまで一貫して想定してきた協力より非協力ゲームが支配的な状況における効果に注目する．以降，まず，相手の手番を，繰り返しゲームにおけるこれまでの履歴から推測し，その推測を元に自分の手番を決定するモデルを導入する．次に計算機実験を行い協力の可能性を調査する．

7.2 他者意図を推定するエージェントモデル

図 7.1 に示す提案する他者意図推定エージェントモデルを説明する．ここでいうエージェントは繰り返し囚人のジレンマゲームをプレイするプレイヤーである．

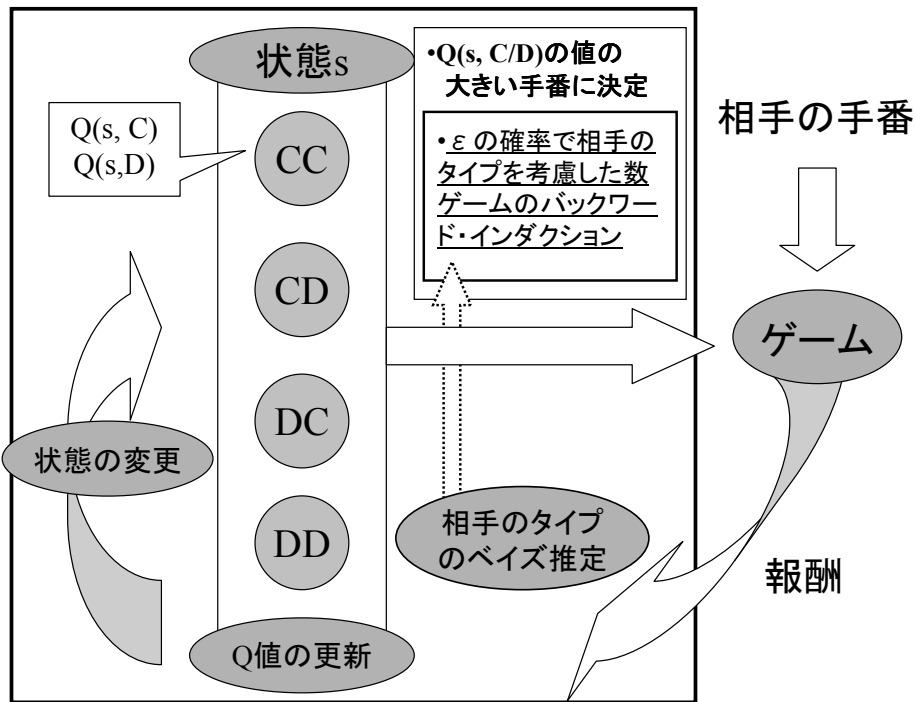


図 7.1. 他者の意図を推定するエージェントモデル

繰り返しゲームを行う際、エージェントは過去一回のゲームの記憶を保持し、前回の自分と相手の手番の組み合わせによって現在の状態を認識する。また、全ての状態ととり得る手番(協力 C/裏切 D)の組み合わせについて価値関数を持ち、任意の状態において手番の価値関数の最も大きなものを基本的には選択する。

ここまではそれぞれ既に提案した遺伝子モデル(価値関数が生得的に決定されゲーム中に変更しないもの)と、学習モデル(ゲームによって得られる利得を報酬として学習し価値関数に変更が生じるもの)である。

本章で提案する意図推定モデルは、上記の学習モデルの機能に加え、対戦相手の手番を推測しそれに対応した手番を選択する機構をもつ。機能としては、相手の手番を推測する部分と、それに対してどのように対応するか判断部の二部にわかれる。

推測部分については以下ようになる。基本的には相手の手番の観測が間違えることを考慮したベイズ推定である。エージェントは、繰り返しゲームにおけるそれぞれの対戦後に相手の手番情報を受け取る。その際に、自分の出した手番に関しては間違えることはないが、観測されたシグナル(相手の手番情報)が誤っている可能性は存在する。そこで、表 7.1 に示すように、実際に観測されたシグナルと実際の相手の手番が異なる確率をあらかじめ保持しておき、ある状態において相手をとるであろう手番の確率を予測していく。7.2 に相手の手番の予測確立を更新していく様子を示す。対戦開始時には、全ての状態において手番 C/D の選択確率は等しいが、このような仕組みによりゲームを続けるにつれ対戦相手のとるであろう手番の推測精度があがっていく。図におけるタイプはエージェントのとりうる戦略の違いを示しており、

繰り返しゲームにおいて遷移した状態が増えれば相手の戦略が予測できるようになる。

表 7.1. 推定テーブル

	シグナルが正しい場合	シグナルが間違っている場合
協力	$1-p$	p
裏切	$1-q$	q

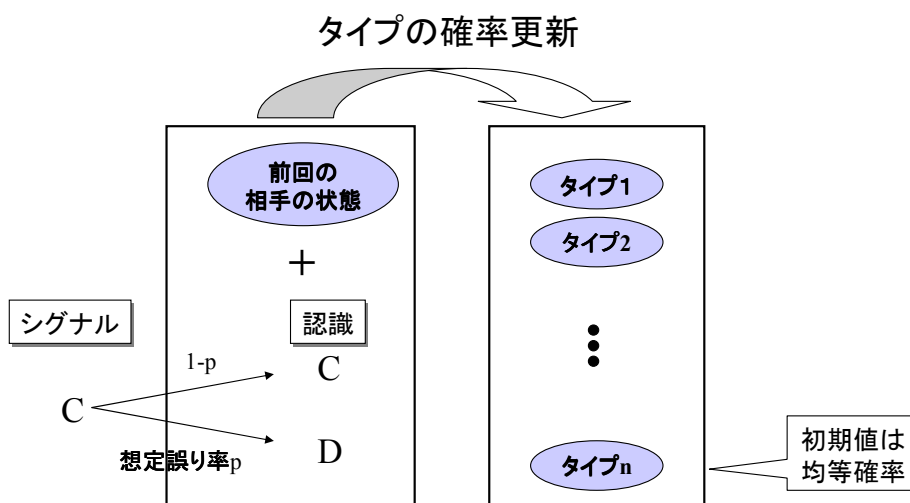


図 7.2. 推定の更新

次に、推測により得られた相手のとる手番の確率を利用し、自分がとる手番を決定する。7.3 にその様子を示す。決定基準は、 n ゲーム先の時点での期待利得値を最大化する手番である。全ての手番の組み合わせについて、 n ゲーム後の累積利得 v_1 から v_{2^n} をそれぞれ計算し、最大となる v_k に到達するための現在の手番を選択することになる。累積利得を計算する際には、現時点でのゲーム以外は将来得られる報酬であるため割引率を適用する。また、現時点から n ゲーム以内のそれぞれの対戦時点においては、あらゆる戦略と仮想的に対戦することになる。その際、先述の機構により推定した相手の戦略の確率を重み付けして計算する。これにより考えうる戦略のうち尤もらしい相手の戦略と対戦した場合に得られる利得が大きな重みで加算されることになる。このようにして、累積期待利得が最大となるよう現在の手番を決定する。

意図推定モデルでは、手番決定の際に、他者意図推定によるか、価値関数に従うかは遺伝子コーディングされている確率によって決定される。このパラメータは、遺伝子モデル、学習モデルにおいては、新たな探索のための ϵ に相当する。先述の二モデルにおいては、単にランダムに選択していたところを、相手の戦略にあわせて手番選択を行う点で異なる。

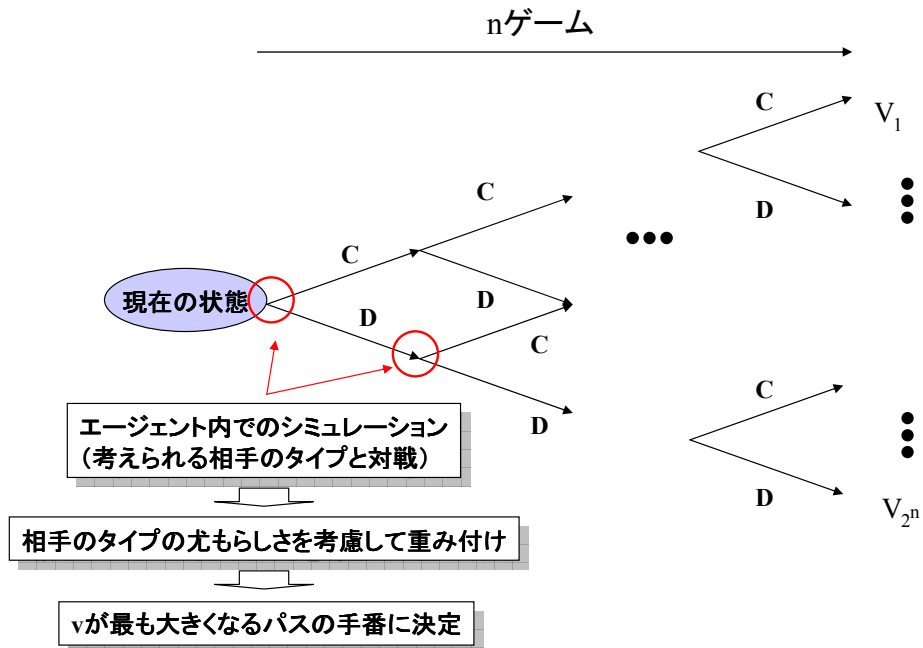


図 7.3. バックワードインダクションによる手番決定

7.3 シミュレーション実験

以上の他者の意図を推定するエージェントモデルの評価のため計算機シミュレーション実験を行った。なお、シグナルが誤って伝わる確率を示す (p, q) は $p = q$ と仮定し、 $p = q = 0.1$ としてシミュレーションを行った。また、提案したモデルにおいてパラメータ n は 3 と以降固定してある。表 7.2 にシミュレーション緒元を示す。

表 7.2. シミュレーション緒元

パラメータ	値
集団サイズ	400
対戦あたり繰り返しゲーム数	100
囚人のジレンマ利得パラメータ (T,R,P,S)	(5,3,1,0)
強化学習パラメータ (α, γ)	(0.4, 0.4)
バックワードインダクションの深さ	3
シグナル想定誤り率 (p, q)	(0.1, 0.1)

まず初期戦略に依存せず協力できる可能性を評価する。本論文では各エージェントは一次のメタ戦略を持つことを仮定しているため、32 通り (5 ビット) の戦略が存在しうる。まずその戦略それぞれについて網羅的に対戦を行った。

学習パラメータをそれぞれ $\alpha = 0.4$, $\gamma = 0.4$ として、手番を学習によらず他者意図推定を利用して決定する確率 ϵ を 0 から 0.2 刻みで 1 まで変動させたときに、各対戦ごとに 36 ポイントの評価が可能となる。

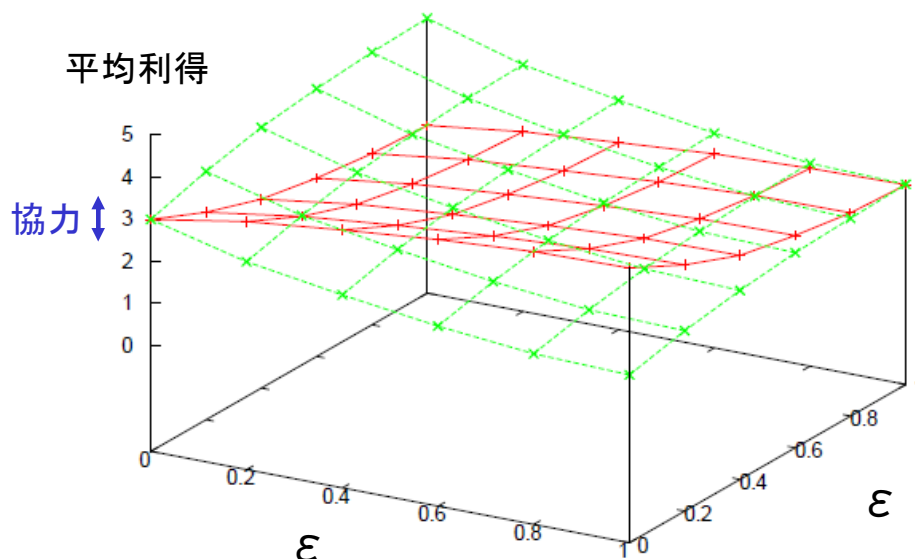


図 7.4. 対戦例

図 7.4 に対戦結果の例を示す。図には各エージェントの平均利得を示す二曲面が存在する。ノイズなしの条件下で、戦略 CCCCC 同士の対戦を行ったものが二曲面としてあらわれている。もし両エージェントが協力的であった場合には、囚人のジレンマの利得構造から 3 に近くなることが予想される。この例ではノイズ、初期戦略とも協力のための好条件が揃っているため、両エージェントの平均利得は 3 に近くなっており協力的であることがわかる。この図からは、両エージェントとも ϵ の値が大きくとも小さくとも双方が同じ程度の値であれば協力しやすいと言える。しかし、片方のエージェントのみが ϵ が大きい場合には、 ϵ の小さなエージェントは貪欲に利得をとるべく学習を行い、それに対して意図推定を行いがちなエージェントは対応できず比較的エージェント間の平均利得に差がでてきている。

図 7.4 は初期戦略の一つの組み合わせ例である。次に全ての初期戦略の組み合わせについて調べてみる。初期戦略は 32 通り存在するため、1024 (32*32) 通りの対戦組み合わせについて、上記のような対戦結果が観察される。統計的に観察するために、対戦の結果の平均利得が 3 ± 0.5 の範囲であれば「協力」と定義し、その割合がどのようになっているかを調べた。

図 7.5 に他者意図推定モデル同士の対戦の結果を示す。凡例はノイズをあらわしている。ノイズの発生確率があまりにも高い状態は現実的には考えにくいので、ここでは 0.2 を上限としている。横軸には上記定義の「協力」を満たしたポイント数、縦軸には「協力」を満たした対戦の割合を示す。ポイントが 0 の時には全ての対戦組み合わせが「協力」は 0 以上であるため 100% となり、その後徐々に減少していくこととなり、減少が小さい方が協力の度合いが高いといえる。例えば、ノイズがない場合に、各対戦 36 ポイント中全て協力する戦略の組み合わ

せ割合は 6.25% である。

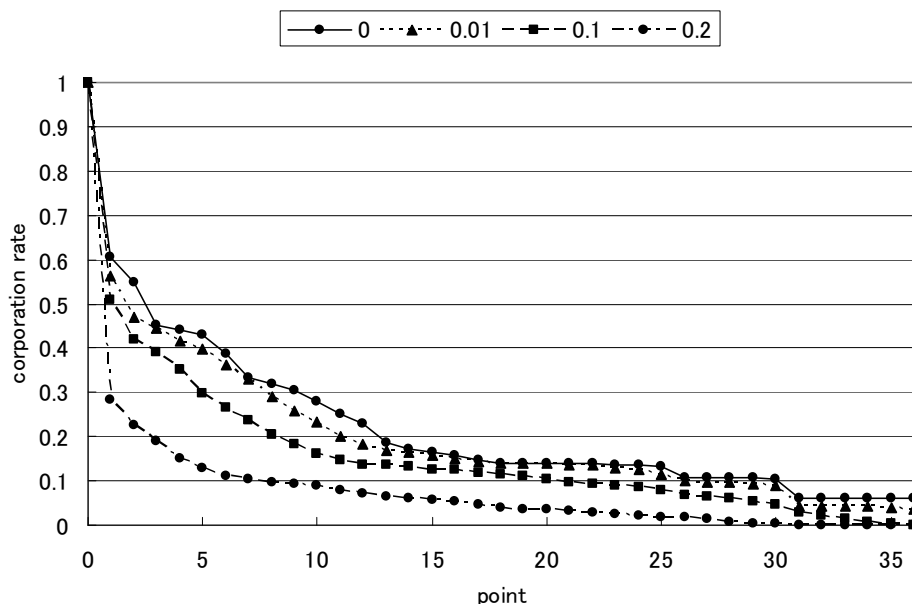


図 7.5. 協力率

ここではノイズの値が異なる 4 つのデータを示しているが、ノイズがない場合が当然最も協力的であり、ノイズの値が増加するにつれて協力現象が発生することが少なくなる。ノイズが 0.1 までは比較的曲線は似たような傾向を示しているが、0.2 となるとどのポイントでも協力できない場合が 70% 出現しその後なだらかに減少する。一方、ノイズが 0.1 以下であれば 50% から徐々に減少する。

7.3.1 異なるモデルとの対戦

図 7.5 では、他者意図推定モデル同士の対戦を見た。ここでは他のモデルとの対戦も含め、他者意図推定の機構がどのような効果を持つのかを調べる。

図 7.6 に対戦するエージェント種類の組み合わせを様々にした場合の、対戦における協力に至るケースの割合を示す。凡例は、bayes が意図推定モデル、rl が学習モデル、gene が遺伝子モデルであり対戦したエージェントのタイプを表す。縦軸は協力の度合いを示している。ただし、エージェントタイプ毎に作成される図 7.5 のような対戦結果に対して、特徴値として少なくとも m ポイントは協力しているポイントの相対比をあらわしている。例えば、図 7.5 の場合には、非常に協力的なので総じて高いが、ノイズが 0.2 の場合には 83%、それ以外の場合にはほぼ 100% となる。横軸はノイズで、ゲームにおいてエージェントが出した手番が逆転してしまう現象の発生確率をあらわしている。

図を見るとわかるように、明らかに意図推定モデルが関連する場合には協力する確率が高いことがわかる。特に意図推定モデル同士の対戦の場合、ノイズが増加するにつれ徐々に協力割合が低下するものの、協力の可能性が非常に高い。その他の二モデルとの対戦において、意

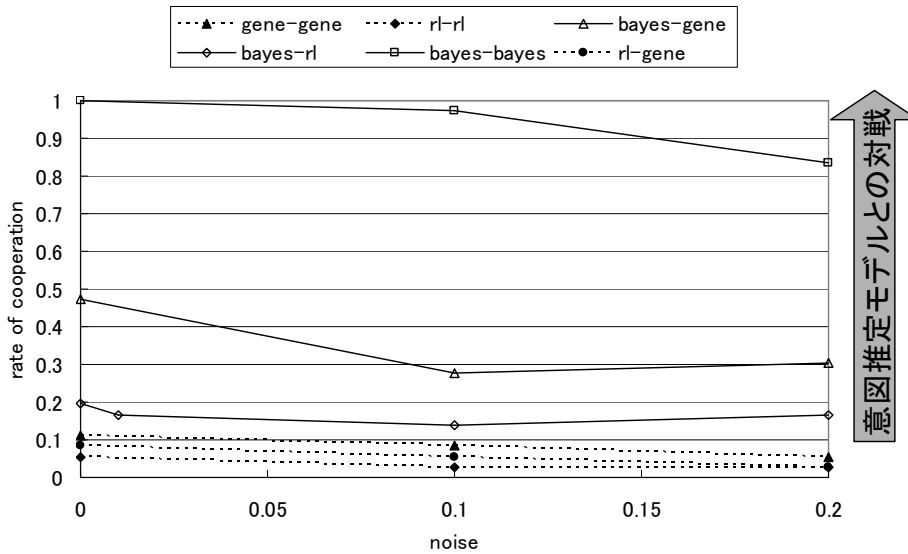


図 7.6. 協力に至るケースの割合

図推定モデルと対戦することにより、他の対戦相手の場合と比べて協力する傾向を示している。その際、学習モデルとの対戦に比較して、遺伝子モデルと意図推定モデルとの対戦の相性が良いのは、遺伝子モデルの単純な反応に対して意図の推定が容易であるからだと考えられる。反対に、学習モデルは対戦相手から貪欲に利得を奪おうと常に試みるため、意図の推定が困難であり意図推定モデルでも対応が難しいことを示している。この傾向は、学習モデルが関連する対戦全てにあてはまり、特に学習モデル同士の対戦において協力の割合が非常に低くノイズのない状態でもなかなか協力できないことがわかる。

7.3.2 推定精度

次に、他者の手番のシグナルの想定誤り率について調べた。図 7.7 に、想定誤り率を 0, 0.1, 0.2 にした場合を示している。縦軸は、全ての対戦組み合わせにおいて、推定して手番を決定する確率を変えていくとき、0.5 以上の確率において協力が行われた対戦の割合を示している。図を見るとわかるように、横軸のノイズ発生確率が増加するにつて、協力の割合が減少するが、想定誤り率によって振る舞いが異なることがわかる。特に全く誤りを想定しない場合には、そのほかの場合に比べ協力の割合が顕著に小さい。それに対し、0.1, 0.2 の確率でシグナルに誤りがあると想定している場合にはノイズに対してある程度の耐性がある。実験では現実を考えて、ノイズを 0.2 までしか試していないが、それ以下では想定誤り率も 0.2 と少し大きめにしておくほうが柔軟に対応できていることがわかる。

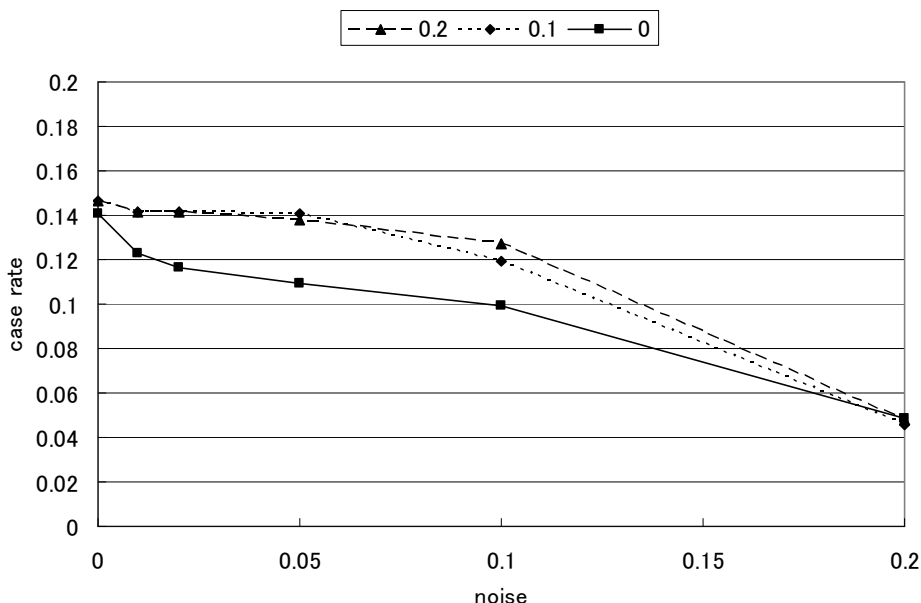


図 7.7. 想定シグナル誤り率による差異

7.3.3 進化的安定性

ここでこの他者の意図を推測する戦略の安定性を調べる．進化的なゲームにおける安定性の指標として ESS(Evolutionary Stable Strategy)[Maynard Smith 82] がある．ESS の定義は次のようなものである．

戦略 s_i (ただし $i \in N$) が存在するとき，ESS である戦略 s_i は次の条件を満たす．

$$e(s_i, s_i) > e(s_{-i}, s_i) \tag{7.1}$$

または

$$e(s_i, s_i) > e(s_{-i}, s_i) \text{ かつ } e(s_i, s_{-i}) > e(s_{-i}, s_{-i}) \tag{7.2}$$

ただし， s_i 以外の戦略を s_{-i} とし， $e(a, b)$ は利得関数で a が b との対戦で得る利得を表す．

ここで前節までに調べた，反射，学習，意図推定モデルの対戦を調べてみると，この条件を満たすことがわかる．そのため，この三タイプの集団においては，進化的に安定であることがわかる．

実際に，遺伝モデル，学習モデル，意図推定モデルを混在させて，集団を進化させてみたところ，図 7.8 のように確かに意図推定モデルが生き残ることがわかった．

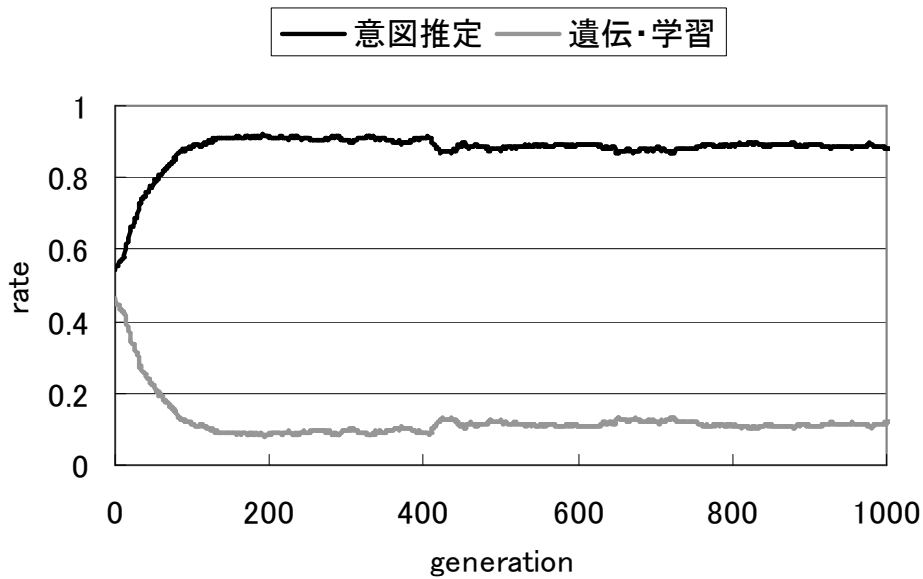


図 7.8. 進化の過程例

7.4 まとめ

本章では、他エージェントの意図を推測しゲームを行うエージェントを提案し、その効果を調べた。

まず、対戦相手のエージェントの戦略を推定し、後方帰納法に基づき最も利得の得られる行動を選択するモデルを提案し、実験によりその効果を確かめた。その結果、反射的なモデル、学習するモデルと比較して著しく協力する可能性が高いことがわかった。また、ノイズのある状況では相手の手番が間違っている可能性があることを考慮して相手の意図を推定する法が協力しやすいことがわかった。

囚人のジレンマでは、ゲームの利得構造から非協力を選択することが合理的とされている。しかし、以上の結果から、相手の意図を推定することによって、双方にとって最適ではないナッシュ均衡から抜け出すことが期待できることがわかった。

第8章

結論

8.1 本研究の結論

本研究では、根底に創発現象の根本の法則を見出すということを大きな目的として、空間構造における囚人のジレンマゲームという一つの具体的な領域に限定し知見を得ることを目的としてきた。

第二章では、一般的なシステムにおいてシステムを構成する要素が意思決定する枠組みとなるゲームについて種々の概念を整理した。また、関連する分野について述べ、本研究の枠組みでそれらを統合する視野を持つことができることを示した。

第三章では、空間構造上において囚人のジレンマ・ゲームをプレイするエージェント集団の基礎的な振る舞いを考察し、これまで生態学の分野で確認されていた現象がゲームをプレイするエージェントについても同様に確認されることがわかった。

第四章では、空間構造が変化する場合について、エージェント間の接続を操作可能であるエージェントを提案し、どのような空間構造になるかをシミュレーション実験によって確かめた。その結果、空間構造はスケールフリー性が強く出てくることがわかった。特に、他のエージェントとの接続に際し全くコストがかからない条件下ではごく少数のエージェントがその他の全てのエージェントと接続する構造になることがわかった。

第五章では、空間構造が学習するエージェントに及ぼす影響について調べた。生物の進化を考えれば、反射的な行動について次の段階として学習が行われることが知られている。そこで次に強化学習を行いゲームに有効な戦略を学習していくエージェントを提案し、エージェントの学習パラメータを観察することで、空間構造がエージェントの学習に与える影響を調査した。その結果、エージェント同士の関係性が強ければ、将来得られる報酬を重視した学習プロセスをとることがわかった。ただし、学習は一般には協力を妨げる傾向があることもわかった。

第六章では、集団の制御可能性について検討する。空間構造を利用した制御方法を提案し、実験によりその効果を調べた。提案する手法は、空間構造上において多くの他のエージェントと接続されているエージェントの戦略を固定するというものである。その手法では、空間構造

を持たない集団とスケールフリーネットワーク上の集団については効果が認められた．現実的な集団内の空間構造を考慮すると，スケールフリー性の高い構造において有効な手法であるといえる．

第七章では，反射・学習行動に加え他エージェントの意図を推測しゲームを行うエージェントを提案する．これまでのエージェントに対する意図を推測するエージェントの強さを調べ，意図を推測する機能の効果を調べる．多エージェントの意図を推測することによって，協力の可能性が劇的に高まり，進化の過程においても安定的であることがわかった．

第二章で詳しく述べたように，多数の要素から構成されるシステムにおいて，要素間の関係性をゲーム理論的に定義して考えるというアプローチは，複雑なシステムを考える上で有効である．その枠組みにおいて限定された領域として，本論文を通して，集団内におけるエージェント同士の囚人のジレンマゲームについて考察してきた．特に，空間構造が集団とエージェントに与える影響に注目した．エージェントは，反射的に行動するもの，学習を行うもの，他のエージェントの意図を推測するものを扱っている．

以上の内容について，ここまで細かくは様々な知見を示してきたが，本研究の貢献をここでまとめると，次のようになる．

- 他の関連する領域を含め一般的なシステムについて統一的な見方を示したこと
- 空間構造がエージェントの学習プロセスに与える影響を明らかにしたこと
- 他エージェントの意図を推測する機構は進化的に安定であると数値的に示したこと

8.2 今後の課題

本研究に関しての今後の課題としては，次のような点が挙げられる．進化的なプロセスを考える場合に実際にはどのようなゲームが支配的であったかという点が非常に重要である．本研究では，囚人のジレンマがそれにあたるとの仮定を暗黙のうちにおいて論を進めてきた．実証的な研究としての，支配的なゲームの種類は重要であろう．また，ゲームの種類を変えた場合に集団の振る舞いがどのように変化するかを調べることも価値があると考えられる．また応用を考える場合，集団の特性の制御に関しては今後の発展が期待される．

最初に述べたように，創発現象の根本的法則を見出すということ大きな目的としてきた．しかしながら，本論文のスコープは空間構造上における囚人のジレンマゲームに絞っており，創発現象の一般法則への接続を達成することはできなかった．ここで，その目的に向かうべく今後の課題を記しておく．

第二章で述べたが，ニュートン力学に対応する統計力学のようなもの考えるにせよ，そもそも要素が従う法則の整理が必要であると考えられる．慣性の法則，作用・反作用の法則，運動の法則というようなシンプルな法則とはならないにせよ，どのような法則がありうるかという検討はこの分野に大きく検討するのではないかと思われる．もしかしたら，それらのルールの種類

類によってある程度集団の挙動を予測することができるようになるかもしれない。

次に、非協力ゲームから協力ゲームへの接続である。既存の研究では、主に非協力ゲームと協力ゲームをそれぞれ独立に扱っている。もちろん囚人のジレンマゲームのような非協力ゲームにおいて、手番として協力を選択しやすくする機構に関する研究は存在するが、ゲーム理論における協力ゲームのような枠組みにまで接続するものはない。確かに、ナッシュプログラムのように、理論的に両者の間をつなぐのはかなりの困難であると思われる。しかし、人の意思決定プロセスについて何らかの仮定をおくことで、変わってくるのではないかとも思われる。構成論的アプローチの利点は、理論的には扱いにくい非合理的な行動の仮定をおくことができることであり、このアプローチで様々な仮定を検討するというのも一つの方法であると考えられる。

ここまで応用に比較して基礎研究的な事項であったが、実際に社会に応用しやすい知見を考えてみると、私見では協力の際の組織の大きさを理論的に求めるのは一定の価値があるのではないかと考える。現在の社会では個人の力で何事かを成し遂げるのはなかなか困難である。そのため、会社組織のように集団で協力して目的を達成するべく組織を構成している。しかし、必ずしも適切な規模を考えてのものではなく、往々にして過去からの慣習などによって決定されている。組織の規模を決定する原因は様々に存在すると考えられるが、取引コストと組織によるメリットが支配的なものだと思われる。実際の組織への応用を考える上では、これらの値の測定が重要となってくるだろう。

以上、いくつか今後の課題として挙げた。創発的な現象の基本的原理を明らかにする研究は近年始まったばかりであり今後の発展が期待される。

謝辞

本研究を行うにあたり，御指導を賜りました石塚満教授に心より深謝致します．テーマ選択から研究の実行まで一連のプロセスを温かく見守って頂きました．適切な御指導を頂き，なんとか一つの形として研究を纏め上げることができました．ここに，厚く感謝し，御礼申し上げます．

また，中間審査，本審査で有益なコメントをしていただいた先生方に感謝致します．研究室メンバーにおいては，特に助手の土肥助手，同期の西村君に感謝致します．最後に，これまで様々な面で支えてくださった友人達と家族に心より感謝致します．

参考文献

- [Akerlof 85] Akerlof, G. A. and Yellen, J. L.: A Near-Rational Model of the Business Cycle, with Wage and Price Inertia, *The Quarterly Journal of Economics*, Vol. 100, (1985)
- [有田 02] 有田 隆也 : 人工生命, 医学出版 (2002)
- [Axelrod 85] Axelrod, R.: *The Evolution of Cooperation*, Basic Books (1985)
- [Bala 00] Bala, V. and Goyal, S.: A Noncooperative Model of Network Formation, *Econometrica* (2000)
- [Barabasi 99] Barabasi, A. and Albert, R.: Emergence of scaling in random networks, *Science*, Vol. 286, No. 5439, pp. 509–512 (1999)
- [Bertalanffy 73] Bertalanffy, von L.: *General System Theory*, Penguin University Books (1973), 長野敬 太田邦昌 訳, 一般システム理論 その基礎・発展・応用, みすず書房, 1973
- [Byrne 88] Byrne, R. W. and Whiten, A.: *Machiavellian Intelligence*, Clarendon Press (1988), 藤田和生 友永雅己 山下博志訳, マキャベリの知性と心の理論の進化論 ヒトはなぜ賢くなったか, ナカニシヤ出版, 2004
- [Coleman 98] Coleman, J.: *Foundations of Social Theory*, Belknap Press (1998), 久慈利武訳, コールマン 社会理論の基礎, 青木書店, 2004
- [Deguchi 98] Deguchi, H.: Agent Based Approach for Social Complex Systems: Management of Constructed Social World, in *Community Computing and Support Systems*, Vol. LNCS 1519, pp. 62–77, Springer (1998)
- [Doya 00] Doya, K.: Complementary roles of basal ganglia and cerebellum in learning and motor control, *Current Opinion in Neurobiology*, Vol. 10, (2000)
- [Dunbar 98] Dunbar, R. I. M.: The Social Brain Hypothesis, *Evolutionary Anthropology*, Vol. 6, No. 5, pp. 178–190 (1998)
- [Epstein 96] Epstein, J. M. and Axtell, R.: *Growing Artificial Societies: Social Science from the Bottom Up (Complex Adaptive Systems)*, Mit Press (1996), 服部正太 木村香代子 訳, 人工社会 複雑系とマルチエージェント・シミュレーション, 構造計画研究所, 1999
- [Fehr 99] Fehr, E. and Schmidt, K. M.: A Theory of Fairness, Competition, and Cooperation, *Quarterly Journal of Economics*, Vol. 114, (1999)
- [Frank 93] Frank, R. H., Gilovich, T., and Regan, D. T.: Does studying economics inhibit

- cooperation?, *Journal of Economic Perspectives*, Vol. 7, (1993)
- [Freeman 79] Freeman, L. C.: Centrality in social networks conceptual clarification, *Social Networks*, Vol. 1, No. 3, pp. 215–239 (1979)
- [Girvan 02] Girvan, M. and Newman, M. E. J.: Community structure in social and biological networks, *Proc. Natl. Acad. Sci. USA*, Vol. 99, pp. 7821–7826 (2002)
- [Granovetter 73] Granovetter, M. S.: The strength of weak ties, *American Journal of Sociology*, Vol. 78, No. 6, pp. 1360–1390 (1973)
- [Hardin 68] Hardin, G.: The Tragedy of Commons, *Science*, Vol. 162, pp. 1243–1248 (1968)
- [Harsanyi 88] Harsanyi, J. C. and Selten, R.: *A general theory of equilibrium selection in games*, MIT Press, Cambridge, Mass. (1988)
- [Hingston 04] Hingston, P. and Kendall, G.: Learning versus evolution in iterated prisoner's dilemma, in *Proceedings of the Congress on Evolutionary Computation, CEC2004*, Vol. 1, pp. 364–372, Portland, Oregon (2004)
- [Hobbes 51] Hobbes, T.: *Leviathan*, Oxford: Clarendon Press (1651), 永井道雄 宗片那義 訳, リヴァイアサン, 世界の名著 28 ホッブズ, 1979
- [石淵 00] 石淵 久生, 中理 達生, 中島 智晴: 空間型繰返し囚人のジレンマゲームにおける隣接プレーヤ間での信頼関係のモデル化, 電子情報通信学会論文誌 D-I, Vol. J83-D-I, No. 10, pp. 1097–1108 (2000)
- [和泉 03] 和泉 潔: 人工市場 市場分析の複雑系アプローチ, 森北出版 (2003)
- [Jensen 94] Jensen, M. C. and Meckling, W. H.: The Nature of Man, *Journal of Applied Corporate Finance*, Vol. 7, (1994)
- [Kahneman 79] Kahneman, D. and Tversky, A.: Prospect Theory: An Analysis of Decision under Risk, *Econometrica* (1979)
- [金子 03] 金子 守: ゲーム理論と蒟蒻問答, 日本評論社 (2003)
- [金子 05] 金子 守: ゲーム理論における合理性と限定合理性, 数理化学, Vol. 503, pp. 76–83 (2005)
- [川人 96] 川人 光男: 脳の計算理論, 産業図書 (1996)
- [木村 02] 木村 邦博: 大集団のジレンマ, ミネルヴァ書房 (2002)
- [Klemm 02a] Klemm, K. and Eguíluz, V. M.: Growing scale-free networks with small-world behavior, *Physical Review E*, Vol. 65, No. 5, pp. 057102–1–057102–4 (2002)
- [Klemm 02b] Klemm, K. and Eguíluz, V. M.: Highly clustered scale-free networks, *Physical Review E*, Vol. 65, No. 3, pp. 036123–1–036123–5 (2002)
- [Knight 21] Knight, F. H.: *Risk, Uncertainty And Profit*, Houghton Niffin Co. (1921), 奥隅 栄喜訳, 危険・不確実性および利潤, 文雅堂銀行研究社, 1959
- [Lessig 99] Lessig, L.: *Code: And Other Laws of Cyberspace*, Basic Books (1999), 山形浩生 柏木亮二 訳, コード, 翔泳社, 2001

- [Lindgren 94] Lindgren, K. and Nordahl, M. G.: Evolutionary dynamics of spatial games, *Physica D*, Vol. 75, pp. 292–309 (1994)
- [Marr 82] Marr, D.: *VISION : A Computational Investigation into the Human Representation and Processing of Visual Information*, W H Freeman & Co (1982)
- [Masuda 03] Masuda, N. and Aihara, K.: Spatial prisoner’s dilemma optimally played in small-world networks, *Physics Letters A*, Vol. 313, pp. 55–61 (2003)
- [松本 96] 松本 元 : 愛は脳を活性化する, 岩波書店 (1996)
- [Maturana 80] Maturana, H. and Varela, F.: *Autopoiesis and Cognition: the Realization of the Living*, Springer (1980), 河本英夫 訳, オートポイエーシス生命システムとは何か, 国文社, 1991
- [Maynard Smith 82] Maynard Smith, J.: *Evolution and the Theory of Games*, Cambridge University Press (1982)
- [Milgram 67] Milgram, S.: The Small World Problem, *Psychology Today*, Vol. 1, No. 1, pp. 60–67 (1967)
- [長田 06] 長田 悠吾, 石川 悟, 大森 隆司, 森川 幸治 : 他者意図の推定に基づく協調行動の計算モデル化, 人工知能学会 2006 全国大会 (2006)
- [Nakamaru 97] Nakamaru, M., Matsuda, H., and Iwasa, Y.: The Evolution of Cooperation in a Lattice-Structured Population, *Journal of Theoretical Biology*, Vol. 184, pp. 65–81 (1997)
- [中島 06] 中島 秀之 : 構成的情報学と AI, 人工知能学会論文誌, Vol. 21, No. 6, pp. 502–513 (2006)
- [Nash 51] Nash, J.: Non-Cooperative Games, *The Annals of Mathematics*, Vol. 54, No. 2, pp. 286–295 (1951)
- [Nash 53] Nash, J.: Two-Person Cooperative Games, *Econometrica*, Vol. 21, No. 1, pp. 128–140 (1953)
- [Neumann 44] Neumann, J. V. and Morgenstern, O.: *Theory of Games and Economic Behavior*, Princeton University Press (1944)
- [Newman 04] Newman, M. E. J. and Girvan, M.: Finding and evaluating community structure in networks, *Physical Review E*, Vol. 69, p. 026113 (2004)
- [Nowak 92] Nowak, M. A. and May, R. M.: Evolutionary games and spatial chaos, *Nature*, Vol. 359, pp. 826–829 (1992)
- [Nowak 93] Nowak, M. A. and Sigmund, K.: A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner’s Dilemma game, *Nature*, Vol. 364, pp. 56–58 (1993)
- [Ohtsubo 06] Ohtsubo, Y. and Rapoport, A.: Depth of reasoning in strategic form games, *Journal of Socio-Economics*, Vol. 35, No. 1, pp. 31–47 (2006)
- [Ohtsuki 06] Ohtsuki, H., Hauert, C., Lieberman, E., and Nowak, M. A.: A simple rule for evolution of cooperation on graphs and social networks, *Nature*, Vol. 441, pp. 502–505

- (2006)
- [岡田 96] 岡田 章：ゲーム理論, 有斐閣 (1996)
- [Oliveira 05] Oliveira, R. and Spencer, J.: Connectivity Transitions in Networks with Super-Linear Preferential Attachment, *Internet Mathematics*, Vol. 2, (2005)
- [Premack 78] Premack, D. and Woodruff, G.: Does the Chimpanzee Have a Theory of Mind?, *The Behavioral and Brain Sciences*, Vol. 4, pp. 515–526 (1978)
- [Rapoport 66] Rapoport, A. and Guyer, M.: A Taxonomy of 2 x 2 games, *General Systems*, Vol. 11, No. 1-3, pp. 203–214 (1966)
- [Rawls 71] Rawls, J.: *A theory of justice*, Harvard University Press, Cambridge (1971), 矢島鈞次 訳, 正義論, 紀伊國屋書店, 1979
- [Rilling 02] Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., and Kils, C. D.: A neural basis for social cooperation, *Neuron*, Vol. 35, pp. 395–405 (2002)
- [Rizzolatti 98] Rizzolatti, G. and Arbib, M. A.: Language within our grasp, *Trends Neuroscience*, Vol. 21, pp. 188–194 (1998)
- [Sandholm 96] Sandholm, T. W. and Crites, R. H.: Multiagent reinforcement learning in the Iterated Prisoner's Dilemma, *Biosystems*, Vol. 37, No. 147–166 (1996)
- [Simon 76] Simon, H. A.: From substantive to procedural rationality, in Latsis, S. ed., *Method and Appraisal in Economics*, pp. 129–148, Cambridge University Press, Cambridge (1976)
- [Snyder 78] Snyder, G. H.: *Conflict Among Nations: Bargaining Decision Making and System Structure in International Crises*, Princeton Univ Press (1978)
- [Sutton 98] Sutton, R. S. and Barto, A. G.: *Reinforcement Learning*, MIT Press (1998)
- [鈴木 99] 鈴木 一功監修, グロービス・マネージメント・インスティテュート編：MBA ゲーム理論, ダイアモンド社 (1999)
- [Suzuki 03] Suzuki, R. and Arita, T.: Evolutionary Analysis on Spatial Locality in N-Person Iterated Prisoner's Dilemma, *International Journal of Computational Intelligence and Applications*, Vol. 3, No. 2, pp. 177–178 (2003)
- [Suzuki 04] Suzuki, R. and Arita, T.: Interactions between Learning and Evolution: Outstanding Strategy generated by the Baldwin Effect, *Biosystems*, Vol. 77, No. 1-3, pp. 57–71 (2004)
- [高野 05] 高野 雅典, 加藤 正浩, 有田 隆也：心の理論における再帰のレベルの進化に関する構成論的手法に基づく検討, *認知科学*, Vol. 12, No. 3, pp. 221–233 (2005)
- [友野 06] 友野 典男：行動経済学 経済は「感情」で動いている, 光文社 (2006)
- [Vázquez 03] Vázquez, A.: Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations, *Physical Review E*, Vol. 67, p. 056104 (2003)
- [Watts 98] Watts, D. J. and Strogatz, S. H.: Collective dynamics of 'small-world' net-

works, *Nature*, Vol. 393, pp. 440–442 (1998)

[Watts 99] Watts, D. J.: *Small Worlds*, Princeton University Press (1999)

[山岸 98] 山岸 俊男：信頼の構造 心と社会の進化ゲーム, 東京大学出版会 (1998)

[松本 06] 松本 光崇：N 人囚人のジレンマゲームにおける規範内部化と協調の関係, 人工知能学会論文誌, Vol. 21, (2006)

関連発表文献

著書

1. Masahiro Ono. IPD Game based on the Spatial Structure. In *Interaction and Emergent Phenomena in Societies of Agents*, MIT press, (to appear).

学術論文

1. 小野真裕, 石塚満. 囚人のジレンマゲームにおけるネットワーク構造の影響. 人工知能学会論文誌 (submitted).

国際会議

1. Masahiro Ono and Mitsuru Ishizuka. Prisoner's dilemma game on network. In *Proceedings of the Eighth Pacific-Rim International Workshop on Multi-Agents, PRIMA2005*, pp. 9–22, Kuala Lumpur, Malaysia, Sep. 2005.
2. Masahiro Ono and Mitsuru Ishizuka. Learning and evolution affected by spatial structure. In *Proceedings of PRICAI 2006*, pp. 651–660, Guilin, China, Aug. 2006.
3. Masahiro Ono and Mitsuru Ishizuka. The controllability of the population characteristics of ipd game players on the spatial structure. In *AAAI Fall Symposium Technical Report FS-06-05 "Interaction and Emergent Phenomena in Societies of Agents"*, pp. 119–126, Arlington, VA, Oct. 2006.

研究報告

1. 小野真裕, 石塚満. 囚人のジレンマゲームにおけるネットワーク構造の影響. 大域ディペンダブル情報基盤シンポジウム, Feb. 2005.
2. 小野真裕, 石塚満. ネットワーク上での囚人のジレンマゲーム. 第11回創発システムシンポジウム, P01, Aug. 2005.
3. 小野真裕, 石塚満. 囚人のジレンマゲームにおける空間上のエージェントの進化・学習. ネットワーク生態学 2006 シンポジウム, March 2006.
4. 小野真裕, 石塚満. 囚人のジレンマゲームにおける空間上のエージェントの進化・学習. 大域ディペンダブル情報基盤シンポジウム, March 2006.

全国大会

1. 小野真裕, 石塚満. 囚人のジレンマゲームにおける空間上のエージェントの進化・学習. 情報処理学会全国大会, 7B-3, March 2006.
2. 小野真裕, 石塚満. クラシファイア・システムによるレート制御方式. 電子情報通信学会全国大会, B-6-53, March 2005.
3. 小野真裕, 石塚満. ネットワーク上での囚人のジレンマゲーム. 人工知能学会 全国大会, 2F2-01, June 2005.
4. 小野真裕, 石塚満. 囚人のジレンマゲームにおけるネットワーク構造の影響. 情報処理学会全国大会, 3J-4, March 2005.